Organization of ribosomal RNA genes from the footrot pathogen *Dichelobacter nodosus*

Sharon La Fontaine† and Julian I. Rood

Author for correspondence: Julian I. Rood. Tel: +61 3 9905 4825. Fax: +61 3 9905 4811. e-mail: julian.rood@med.monash.edu.au.

Department of Microbiology, Monash University, Clayton, Victoria 3168, Australia

Southern hybridization analysis revealed that there were three *rrn* loci within the genome of *Dichelobacter nodosus*, the causative organism of ovine footrot. These loci (*rrnA, rrnB* and *rrnC*) were isolated on recombinant lambda clones, and comprised 16S, 23S and 5S rRNA genes closely linked in that order. Sequence and primer extension analysis revealed the presence of putative genes encoding tRNA^•ë^ and tRNA^A^ within the 16S–23S spacer region, as well as a number of potential regulatory features. These elements included a single promoter, which was mapped upstream of the 16S rRNA gene and which was similar to *Escherichia coli* consensus promoter sequences, an AT-rich upstream region, a GC-rich motif that may be involved in stringent control, leader and spacer antitermination sequences, sites for ribonuclease processing, and a putative factor-independent terminator sequence. Potential open reading frames (ORFs) were identified within the regions flanking the *rrn* loci, with identical copies of the 3′ terminal ORF present downstream of each rRNA operon. Determination of the complete sequence of the 5S rRNA gene, and derivation of the 5S rRNA secondary structure, further substantiated the 16S rRNA-based placement of *D. nodosus* within the gamma division of the Proteobacteria.

**Keywords:** *Dichelobacter nodosus*, footrot, *rrn* operon, rRNA genes, phylogeny

**INTRODUCTION**

Since ribosomes constitute the translational machinery of the cell, ribosomal RNA (rRNA) is vital to cellular growth, function and survival. Consequently, the primary, secondary and tertiary structures of rRNA molecules have been highly conserved during evolution (Gutell, 1992). The organization of rRNA genes (*rrn*) from bacteria representing several different phylogenetic groupings has been examined (Krawiec & Riley, 1990; Menke *et al.*, 1991). These studies have enabled comparisons to be made with respect to bacterial *rrn* gene organization and the sequence elements which govern the synthesis, regulation and processing of rRNA molecules, and have led to the suggestion that *rrn* organization may provide additional genetic evidence for the coherence of groupings defined on the basis of phylogenetic analyses (Menke *et al.*, 1991).

The genes encoding the rRNA molecules of *Escherichia coli* have been most extensively characterized and are organized as transcriptional units, each with the gene order 16S–23S–5S (Jinks-Robertson & Nomura, 1987; Srivastava & Schlessinger, 1990). This order and linkage of rRNA genes is a common, though not universal, feature of bacterial *rrn* organization (Krawiec & Riley, 1990; Srivastava & Schlessinger, 1990; Menke *et al.*, 1991). Transcription of the seven *E. coli* *rrn* operons proceeds from multiple promoters which are differentially regulated, is subject to growth-rate and stringent control (Jinks-Robertson & Nomura, 1987; Wagner, 1994), and is accompanied by antitermination mechanisms to prevent premature transcription termination (Berg *et al.*, 1989). In addition, two classes of 16S–23S rRNA intergenic spacer regions exist, one carrying tRNA^•ë^ and tRNA^A^ genes (*rrnA, D* and *H*), and the other carrying a tRNA^G^ gene (*rrnB, C, E* and *G*) (Jinks-Robertson & Nomura, 1987). Additional 5S rRNA and/or tRNA genes are present at the distal end of some *E. coli* *rrn* operons.
Dichelobacter nodosus is a member of the Cardiobacteriaceae (Dewhirst et al., 1990), and is the essential causative agent of ovine footrot (Stewart, 1989), a debilitating disease of sheep which continues to afflict the Australian sheep and wool industry. This bacterium is a slow-growing, fastidious, obligately anaerobic, Gram negative, rod-shaped organism, for which there is no system for genetic manipulation. Determination of the *D. nodosus* 16S rRNA sequence led to the elucidation of its true phylogenetic status (Dewhirst et al., 1990; La Fontaine & Rood, 1990), and to improved methods of detection using 16S rRNA-based PCR analysis (La Fontaine et al., 1993). To obtain a greater understanding of the inherent properties of *D. nodosus*, a more detailed analysis of its rRNA genes was undertaken. This report describes the isolation and characterization, by Southern hybridization, sequence and primer extension analysis, of the three *rrn* loci from *D. nodosus* strain A198, and provides an insight into the putative regulatory elements that control rRNA synthesis in *D. nodosus*.

**METHODS**

**Bacterial strains, plasmids and bacteriophages.** The *D. nodosus* strain used was the reference strain A198 (VC/S1001). *E. coli* strains DH5α [F'endorAI hdR17 (r6 m2) thi-1 λ recA1 gyrA96 relA1 proAB lacYIΔΔΔ ΔZAM15] (Bethesda Research Laboratories) and LE392 [F supF′ supE44 hsdR15 (r6 m2) galK2 galK22 trpR 55 metB lacYI] (Sambrook et al., 1989) were used for transformation with recombinant plasmids and for bacteriophage infection, respectively. Recombinant plasmids were derived from the vectors pUC18 (Yanisch-Perron et al., 1985), pBlue-script II KS (Stratagene) and pWSK29 (Wang & Kushner, 1991). Recombinant bacteriophages were derived from strain A198 genomic lambda libraries constructed in the vectors EMBL3 and LambdaGEM-12 (Katz et al., 1994).

**Media and culture conditions.** *D. nodosus* was cultured on hoof (Thomas, 1958), trypticase-argonine-serine (TAS) (Skerman, 1975), or Eugonagar (BBL) agar plates, and in either TAS broth or Eugonbroth (BBL). Cultures were incubated at 37°C for 3-4 d (for plate cultures) or 48 h (for broth cultures), in an atmosphere containing 10% H2 and 10% CO2 in N2. *E. coli* DH5α derivatives were cultured on 2YT (Miller, 1972) solid and liquid media, while strain LE392 was cultured on solid and liquid L + G medium [LB medium (Sambrook et al., 1989) with the addition of 10 mM Tris/HCl pH 7.5, 1 mM MgCl2 and 0.2% (w/v) glucose] supplemented with 1% (w/v) maltose. Ampicillin (100 μg ml−1) was added where required.

**Preparation of nucleic acids and general molecular techniques.** Total genomic DNA from *D. nodosus* was isolated from broth cultures as previously described (Anderson et al., 1984). Plasmid DNA was isolated by the methods of Birnboim & Doly (1979) and Morelle (1989). DNA from recombinant bacteriophages was prepared from plate lysates as described by Sambrook et al. (1989). DNA fragments were excised from 0.8-1.2% (w/v) agarose gels and purified using a GeneClean Kit (BIO 101), as recommended by the manufacturer. Oligonucleotides were chemically synthesized in an automatic DNA synthesizer (model 381A, Applied Biosystems). The cloning, analysis and manipulation of DNA, which included ligation reactions, transformation of plasmid DNA, restriction analysis, and agarose gel electrophoresis, was carried out essentially as described by Sambrook et al. (1989). Total *D. nodosus* RNA was prepared from plate cultures as previously described (La Fontaine et al., 1993), or from broth cultures using TRISOLV (Biotex Laboratories), as described by the manufacturer, with the inclusion of a 5 min incubation at 55°C after resuspending cells in the TRISOLV reagent. In addition, the isolated RNA was treated with RNase-free DNase I (RQ1, Promega; 1·0 unit) at 37°C for 30 min, followed by a second round of extraction with TRISOLV. Individual RNA molecules were isolated after electrophoresis and excision of individual bands from a 1·2% (w/v) low-melting-temperature agarose (SeaPlaque, FMC Bioproducts) gel. Purification involved melting the gel slice in an equal volume of distilled H2O (dH2O) at 70°C for 15 min, followed by two extractions with phenol/chloroform/isoamyl alcohol (25:24:1, by vol.) and one with chloroform/isoamyl alcohol (24:1). RNA was precipitated from the aqueous phase as described by Baylis & Bibb (1988a).

**Labelling of nucleic acid probes.** DNA fragments were labelled with [α-32P]-dATP (~400 Ci mmol−1 (~14·8 TBq mmol−1); Amersham) using a GigaPrime DNA Labelling Kit (Biacore), or with digoxigenin (DIG)-11-dUTP (Boehringer-Mannheim), using protocols recommended by the manufacturer. DNA probes were also generated by the incorporation of DIG-11-dUTP during the polymerase chain reaction (PCR) as follows. Each 100 μl PCR mixture contained approximately 200-500 ng template DNA, 1 μM of each oligonucleotide primer, 180 μM of each deoxynucleotide triphosphate (dNTP; Promega), 2 μl of DIG DNA Labelling Mixture (10× concentrated; Boehringer-Mannheim), 2-5 units of Taq DNA polymerase and 1× reaction buffer (Boehringer-Mannheim). The PCR mixture was placed in a thermal cycler (FTS-1 Thermal Sequencer, Corbett Research) and the temperature was cycled to 95°C for 1 min, 49°C for 1 min and 72°C for 1 min, for 30-35 cycles. The labelled PCR fragment was excised after electrophoresis through a 1·2% (w/v) low-melting-temperature agarose gel and boiled for 10 min in dH2O prior to use. Oligonucleotides were 5′ end-labelled with [α-32P]-dATP (>5000 Ci mmol−1 (>185 TBq mmol−1); Amersham) and T4 polynucleotide kinase (Promega) as described by Keller & Manak (1989). Unincorporated nucleotides were separated from labelled probes using NACS Prepac columns (Bethesda Research Laboratories). Labelled rRNA probes were prepared by hydrolysis as described by Arrand (1985) and then end-labelled with [α-32P]-dATP (>5000 Ci mmol−1) and T4 polynucleotide kinase.

**Southern hybridization analysis.** Digested DNA was fractionated by electrophoresis through 0·8% (w/v) agarose gels and transferred to nitrocellulose by the method of Southern (1975). Genomic DNA was also prepared and digested within agarose blocks as described by Smith & Cantor (1987), and subjected to pulsed-field gel electrophoresis (PFGE) through a 0·8% (w/v) agarose gel using a CHEF-DR II apparatus (Bio-Rad). Electrophoresis was carried out at 150 V at 14°C using a ramped switch time of 0·5-10 s over 10 h. The DNA was transferred to nitrocellulose as before, or to nylon membranes (Hybond-N+, Amersham) under alkaline conditions using recommended protocols (Bio-Rad). Hybridization with [α-32P]-dATP-labelled DNA probes was carried out as already described (Katz et al., 1991). For DIG-11-dUTP-labelled DNA fragments, the manufacturer’s protocols were used for the hybridization and washing of filters. In experiments which employed the DIG-11-dUTP-labelled PCR fragment as a probe, hybridization was carried out at 60°C, and the highest stringency post-hybridization washes were carried out in 0·1 SSC, 0·1% (w/v) SDS at 60°C. Hybridization with [α-32P]-dATP-labelled RNA probes was carried out at 50°C in 50% (w/v) deionized formamide, 2× Denhardt’s solution (Sambrook et al., 1989), 5× SSC, 0·2% (w/v) SDS, 10 mM EDTA and...
0.5 mg salmon sperm DNA ml⁻¹. Washes were carried out in 2 x SSC, 0.5% (w/v) SDS, twice for 5 min at room temperature, and in 0.1 x SSC, 0.5% (w/v) SDS, twice for 10 min at 45 °C. Conditions described elsewhere (La Fontaine et al., 1993) were employed for the hybridization of oligonucleotide probes to Southern blots.

Screening of genomic lambda libraries. Bacteriophage infection of the E. coli host LE392 was carried out as described by Sambrook et al. (1989). Recombinant bacteriophage DNA from plaques within the E. coli lawn were transferred to nitrocellulose filters (Schleicher and Schuell) as described by Sambrook et al. (1989), and hybridized with [α-³²P]dATP-labelled DNA probes under conditions described previously (Katz et al., 1991), or with DIG-11-dUTP-labelled probes as recommended by the manufacturer.

DNA sequence and primer extension analysis. Sequencing of double-stranded DNA templates was achieved using a T7 DNA polymerase sequencing kit (Pharmacia) and the recommended protocols. Primer extension experiments using approximately 50 μg of total D. nodosus RNA extracted from broth cultures, and a commercially available Primer Extension System (Promega), were carried out according to the manufacturer’s instructions. Primer extension products were analysed alongside DNA sequence ladders prepared using the appropriate templates and primers.

Computer analysis of the nucleotide and deduced amino acid sequences involved comparisons with the sequences in the nucleotide (GenBank and EMBL) and protein (SWISS-PROT, PIR and GenPept) sequence databases using the fastA (Pearson & Lipman, 1988) and BLAST (Gish & States, 1993) programs. The analysis of rRNA sequences was facilitated using data and programs within the Ribosomal Database Project (RDP) (Maizel et al., 1994) and version 3.5e of PIR and GenPept (Phylogeny Inference Package) (Felsenstein, 1989).

RESULTS AND DISCUSSION

Copy number and linkage of D. nodosus rRNA genes, and isolation of the rrl loci

The sequence of the D. nodosus 16S rRNA molecule (Dewhirst et al., 1990), and the isolation of the distal two-thirds of a D. nodosus 16S rRNA gene on the recombinant plasmid pJ1R284 (La Fontaine & Rood, 1990), was previously reported. However, it was determined subsequently that the portion of the pJ1R284 insert that was located downstream of the 16S rRNA gene was devoid of rRNA-encoding sequences and did not constitute a contiguous region of the D. nodosus genome (data not shown). To isolate intact rrl loci from D. nodosus, genomic lambda libraries derived from strain A198 were screened with a pJ1R284-derived probe containing the cloned portion of the 16S rRNA gene. The bacteriophage AR33 was one of six positive recombinants that were identified after screening approximately 1600 EMBL3-derived recombinant plaques, and was characterized in detail. By a combination of Southern hybridization with oligonucleotide and rRNA probes (data not shown), and sequence analysis, 16S, 23S and 5S rRNA genes were localized within the cloned insert (Fig. 1). The region of AR33 DNA carrying the three rRNA genes was designated rrlA, in accordance with accepted nomenclature (Nomura et al., 1977).

Internal probes derived from each of the rRNA genes within AR33 (Fig. 1) were used to probe genomic DNA digests. With each probe, three hybridizing bands were evident in the NsiI, HpaI and EcoRV digests (Fig. 2). In addition, the HpaI hybridization profiles obtained with each probe were identical, as were the EcoRV profiles. These results showed that there were three copies of each of the rRNA genes within the strain A198 genome, and indicated a close linkage of the individual rRNA genes within each locus.

Further screening of approximately 1400 LambdaGEM-12-derived recombinant plaques led to the isolation of the remaining two rrn loci, rrnB and rrnC, on AR38 and AR104, respectively. Restriction mapping and Southern hybridization analysis were used to localize the rRNA genes within these recombinants (Fig. 1). Although the regions containing the 16S, 23S and 5S rRNA genes on AR33, AR38 and AR104 shared considerable restriction identity, the flanking regions had different restriction profiles, confirming that different rrl loci had been isolated. Additional hybridization experiments confirmed that all of the rRNA genes had been isolated, that they were isolated on contiguous inserts, and that these inserts were derived from the strain A198 genome (data not shown).

Nucleotide sequence analysis of the rrl leader regions – identification of putative regulatory elements, the transcriptional start site and the 5' mature 16S rRNA terminus

To identify the regulatory elements involved in the synthesis and processing of the rRNA molecules in D. nodosus, the regions flanking the rRNA genes from each of AR33, AR38 and AR104 were subcloned and their nucleotide sequences were determined. For rrlA, the sequence of 485 bp upstream of the 16S rRNA gene was determined on both strands, and that of an additional 124 bases was obtained for one strand only. For rrlB and rrlC, the sequence of 465 and 540 nucleotides, respectively, was determined on one strand upstream of the 16S rRNA gene. The sequences of rrlA, rrlB and rrlC were identical for 247 nucleotides upstream of the sequence encoding the mature 16S rRNA molecule. Thereafter, the sequence of rrlA diverged from those of rrlB and rrlC, which remained identical for an additional 60 nucleotides (Fig. 3).

Within the rrl leader regions, several putative regulatory elements were identified. Two potential promoter sequences (P1 and P2), which had —10 and —35 elements with similarity to the E. coli σ⁷₀-containing RNA polymerase (RNAP) consensus binding sites (Rosenberg & Court, 1979), were identified (Fig. 3). However, based on the suboptimal 14 bp spacing between the P2 —10 and —35 elements, and their location relative to the start of the mature 16S rRNA sequence, it was unlikely that P2 would be functional. In addition, primer extension analysis with oligo-
nucleotide #1154 (Fig. 3) confirmed that $rnn$ transcription initiated from P1, with the T residue located eight bases downstream of the $-10$ hexamer (Fig. 4a). The existence of what appeared to be only a single $rnn$ promoter in $D. \text{nodosus}$ was in contrast to the presence of multiple $rnn$ promoters in $E. \text{coli}$ (Wagner, 1994). Prior to this study, only one other $D. \text{nodosus}$ promoter, that of the fimbrial operon, had been mapped (Hobbs et al., 1991). The $fim$ promoter elements resembled those which are bound by $\sigma^70$-containing RNAP and which are characteristically regulated by an activator protein (Kustu et al., 1989). Therefore, the $rnn$ promoter is the first $D. \text{nodosus}$ promoter to be mapped which shows similarity to elements that are recognized by the major sigma factor, $\sigma^70$ (Rosenberg & Court, 1979).

The GC-rich region (GCGCACC) located immediately 3' of the $-10$ hexamer of P1 (Fig. 3), was similar to the discriminator motif (GCGC) that is characteristic of stringently controlled promoters, which include the $rnn$
Fig. 3. Nucleotide and deduced amino acid sequences of the regions upstream of rRNA, rrbB and rrcC. Amino acid sequences are shown below the corresponding nucleotide sequences. Stop codons are indicated by an asterisk. Dashes represent unsequenced regions, uppercase letters indicate double-stranded sequence and lowercase letters indicate single-stranded sequence. Nucleotides that could not be determined are denoted by an underline. The GenBank accession numbers for the rRNA, rrbB and rrcC leader sequences are U26680, U26681, and U26682, respectively.

Premature termination of rRNA transcription is prevented by the existence of antitermination mechanisms similar to those of lambdoid phages (Berg et al., 1989). Regions equivalent to the nai loci (boxes A, B and C) that are involved in the bacteriophage lambda antitermination system (Friedman & Gottesman, 1983), have been identified within the rRNA promoter regions of many bacteria (Berg et al., 1989), and were also present downstream of the D. nodosus rRNA P1 promoter (Fig. 3). The putative D. nodosus box A sequence had seven of the nine consensus nucleotides defined for box A sequences ([C/T]GCTT[T/-]A) (Friedman & Gottesman, 1983). A putative box B element, which is not a conserved sequence but a region of hyphenated dyad symmetry, was identified 5′ to the box A element (Fig. 3), a location that is common to the rRNA antitermination system (Berg et al., 1989). The D. nodosus box B element had the potential to form two possible hairpin loops with either a five- or a 15-nucleotide GA-rich loop that is characteristic of both phage (Friedman & Gottesman, 1983) and rRNA box B regions (Berg et al., 1989). Twenty-one nucleotides were used for primer extension experiments are underlined. The 10–35 hexamers of the P1 promoter are boxed, as are putative box A and box C sequences. The region of box B dyad symmetry is indicated by arrows above the sequence. Other regions of dyad symmetry that may participate in the formation of secondary structures are indicated by broken arrows below the sequence. Nucleotides that are likely to be involved in the formation of a processing stem with a region downstream of the 16S rRNA sequence are in bold type and shaded. The potential P2 promoter sequences are also indicated. The GenBank accession numbers for the rRNA, rrbB and rrcC leader sequences are U26680, U26681, and U26682, respectively.
downstream of the box A sequence, a sequence was identified that resembled the phage and *E. coli* rrr box C sequence, which consists of an alternating GT-rich pattern [GG(T/C)GT(G/A)(T/C)G] (Friedman & Gottesman, 1983) (Fig. 3).

Primer extension analysis with oligonucleotide #115 resulted in a consistent, albeit somewhat heterogeneous, product which spanned several nucleotides (Fig. 4b) and represented the 5′ end of the mature 16S rRNA molecule. The sequence of the *D. nodosus* 16S rRNA was previously reported from nucleotide position 2, which was shown to be an A residue (Dewhirst *et al.*, 1990). Primer extension analysis indicated that the sequence immediately adjacent to this nucleotide was ... TTTTGG (Fig. 3) and that the 5′ nucleotide of the mature 16S rRNA (M) could be localized to one of the three T nucleotides immediately preceding the G (Fig. 4b). A longer exposure of the autoradiograph revealed the presence of a primer extension product of 216 nucleotides (Fig. 4c), which was consistently observed. The nucleotide corresponding to this band may represent the 5′ end of a premature (P) form of the 16S rRNA. This nucleotide occurred within a sequence of 30 nucleotides (Fig. 3), an inverted form of which was present downstream of the 16S rRNA gene (Fig. 5a).

**Identification of tRNA genes within the 16S–23S rRNA spacer**

The nucleotide sequence of the 16S–23S rRNA intergenic spacer region, the 5′ and 3′ termini of the 23S rRNA gene, the 23S–5S rRNA intergenic spacer, the 5S rRNA gene and 300 bp downstream, was determined for both strands of *rrn A* (Fig. 5), but for only one strand of *rrn B* and *rrn C* since their sequences were identical. The 3′ nucleotides of the *D. nodosus* 16S rRNA molecule were determined previously (Dewhirst *et al.*, 1990), while the 5′ and 3′ termini of the 23S rRNA gene were assigned by comparison with published sequences within the Ribosomal Database Project (Maidak *et al.*, 1994). The 16S–23S spacer region contained putative tRNA^Ile^ and tRNA^Ala^ genes (Fig. 5a) which had 62–89 % and 52–98 % sequence identity, respectively, with equivalent genes from other organisms. These sequences were able to fold into the compact cloverleaf structure characteristic of tRNA molecules, and contained anticodons specifying Ile and Ala within the anticodon arms of the respective molecules. The presence of tRNA genes within the 16S–23S spacer regions is another common, but not universal, feature of rRNA operons (Krawiec & Riley, 1990), and in many cases there appears to be a bias towards the presence of
Signals for the processing of rRNA precursors

Within the *D. nodosus* rRNA loci, the sequences encoding the mature 16S and 23S rRNA molecules were bordered by sequences with the potential to form double-stranded tandemly arranged tRNA<sup>16s</sup> and tRNA<sup>23s</sup> genes (Loughney et al., 1982; Jinks-Robertson & Nomura, 1987; Nakagawa et al., 1992; Gill et al., 1994; Minnick et al., 1994). Although the significance of the association of these particular tRNA genes with *rrn* spacers remains unclear, perhaps they represent an ancestral feature of rRNA operons and a lack of selective pressure over time has resulted in non-uniformity with respect to the presence of tRNA<sup>16s</sup> and tRNA<sup>23s</sup> genes in rRNA operons within and between organisms. Between the *rrn<sup>16s</sup>loci* and the 23S rRNA genes, additional putative box A and box B elements were identified, which also are common among bacterial *rrn* spacer regions (Berg et al., 1989).

The 23S–5S rRNA spacer region (Fig. 5b) did not show any significant similarity to sequences within the nucleotide databases. The mature termini of the 5S rRNA gene were assigned by comparison with other bacterial 5S rRNA gene sequences, and a putative factor-independent transcription termination signal (T) \[AG = -29 \text{ kcal mol}^{-1} (121 \text{ kJ mmol}^{-1})\], calculated by the method of Tinoco et al. (1973)] was identified seven nucleotides downstream from the 3′ terminus (Fig. 5b).

**Sequence and structural analysis of the 5S rRNA gene**

The *D. nodosus* 5S rRNA gene was approximately 120 bp in length. Using a secondary structure model for 5S rRNA sequences that was derived from comparative sequence analysis, biochemical and chemical studies (Specht et al., 1990), a secondary structure for the *D. nodosus* 5S rRNA was deduced from its gene sequence (data not shown). This structure, comprising five helical segments, A to E, connected by loops a to e, conformed to the minimal model of secondary structure proposed for an *ss* rRNA, and included the variation proposed for the minimal model of eubacterial 5S rRNA (Specht et al., 1990). In addition, the *D. nodosus*-derived structure contained features in common with those of other *Proteobacteria*, which included an A:C base pair at position 11:109 and a three-base loop e, the latter common to gamma group structures.

Comparison of the *D. nodosus* 5S rRNA sequence with that from organisms of the beta and gamma groups of the *Proteobacteria* showed that the *D. nodosus* sequence shared 61–67% and 78–82% sequence identity with members of these groups, respectively. As expected from these results, *D. nodosus* clustered with other gamma group bacteria within 5S rRNA-based phylogenetic trees (data not shown). Therefore, comparative analysis of the deduced primary and secondary structures of the 5S rRNA from *D. nodosus* has provided independent confirmation of its phylogenetic position within the gamma division of the *Proteobacteria*, which was initially determined by 16S rRNA-based phylogenetic analyses (Dewhirst et al., 1990).
stem structures (Fig. 6). These structures are yet another conserved feature of rRNA operons, and contain the primary processing sites for RNase III in *E. coli*, or its equivalent in other organisms (Srivastava & Schlessinger, 1990). The 9S precursor containing the 5S rRNA sequence is cleaved by RNase E, and the resultant molecule, as well as the pre-16S and pre-23S rRNAs, are acted upon by additional endo- and exoribonucleases to produce the mature termini (Srivastava & Schlessinger, 1990). In *D. nodosus* 30 nucleotides flanking the 16S rRNA sequence (Figs 3 and 5a) had the potential to participate in stem formation (Fig. 6a), in contrast to the approximately 45 stem-forming nucleotides flanking the *E. coli* 16S rRNA sequence (Srivastava & Schlessinger, 1989). These bases in both *E. coli* (Srivastava & Schlessinger, 1989) and *D. nodosus* (Figs 3 and 6) included the box C sequence. Cleavage by *E. coli* RNase III occurs within this sequence (Srivastava & Schlessinger, 1989), and in *D. nodosus*, a primer extension product obtained using oligonucleotide #115 (Fig. 4c) was consistent with cleavage within the putative box C sequence.

As in *E. coli* (Srivastava & Schlessinger, 1990), the putative 23S rRNA processing stem included nucleotides at the 5' and 3' mature termini (Figs 5 and 6b). Terminal nucleotides of the 5S rRNA sequence also had the potential to participate in the formation of a short double-stranded stem (Fig. 6b). The point(s) of cleavage that would lead to the generation of premature and mature 23S and 5S rRNA molecules were not determined. Although the specificity of *E. coli* RNase III is likely to involve secondary and tertiary structural features of the RNA (Srivastava & Schlessinger, 1990), and despite the sequence differences in the stem structures associated with *D. nodosus* rRNA molecules compared with those of *E. coli*, there were regions of limited sequence similarity which, in *E. coli*, corresponded to regions surrounding the sites of RNase III cleavage (Fig. 6).

**Analysis of putative ORFs flanking the rrn loci**

The nucleotide and deduced amino acid sequences of the regions located upstream and downstream of the *rrn* loci were compared with the sequences in the nucleotide and protein sequence databases. The C-terminal ends of

---

**Fig. 6.** Potential processing stems for *D. nodosus* rRNA molecules. Lines denote standard base pairs and filled circles denote G:U base pairs. Filled arrows indicate the mature termini (M) of the rRNA molecules as identified for the 16S rRNA, and as proposed for the 23S and 5S rRNAs. (a) Processing stem for 16S rRNA and potential secondary structures that may form from sequences within the *rrn* leader region. Box C within the stem is shaded and the point of cleavage within box C is indicated by an arrow. The *E. coli* sequence surrounding the RNase III cleavage site within the stem-forming sequence that flanks the 16S rRNA 3' terminus (Srivastava & Schlessinger, 1989), is also shown. *E. coli* nucleotides that are identical with those in the *D. nodosus* sequence are in uppercase bold type. (b) Putative processing stem for 23S and 5S rRNAs. The *E. coli* sequence surrounding the RNase III cleavage site within the stem-forming sequence that flanks the 23S rRNA 5' terminus is also shown. *E. coli* nucleotides that are identical with those in the *D. nodosus* sequence are in uppercase bold type.
potential ORFs, designated ORFA and ORFC, were identified upstream of rrnA and rrnC (Fig. 3), respectively, and identical copies of a putative ORF, ORFE, were located downstream of all three rrn loci (Fig. 5b).

ORFA had 37.0-7% amino acid sequence identity, over an 81 amino acid region, with a hypothetical protein, ORFY, of unknown function from Pseudomonas aeruginosa (Hobbs et al., 1993). The latter ORF is located within the pilS/pilR locus of the E. coli genome, and the clpB gene in P. aeruginosa. ORFC had 48-57% amino acid sequence identity with DNA binding proteins from Bacillus subtilis (ORF1 of the comE operon: Hahn et al., 1993), and from Homo sapiens (OBP-1 and OBP-2; Zhang & Nonoyama, 1994), respectively.

ORFs have been identified upstream of rRNA operons from other bacteria, including the Streptomyces coelicolor rrnD operon (Baylis & Bibb, 1988b), an rRNA operon from Thiothrix ferrooxidans (Bakshi et al., 1990), and the E. coli rrnB operon (Brosius et al., 1981), with the clpB gene located upstream of the E. coli rrnG operon (Squires et al., 1991). There does not appear to be any functional significance as to which genes or ORFs reside in close proximity to rRNA operons but, as observed for D. nodosus ORFA and ORFC, neither the E. coli rrnB, S. coelicolor rrnD, nor the T. ferrooxidans rRNA operons were followed by any obvious factor-independent transcription termination signals. In D. nodosus, termination of transcription of ORFA and ORFC may be factor-independent, or transcription may continue uninterrupted into the rRNA operons, as occurs with the ORF upstream of rrnB from E. coli (Boros et al., 1983).

Conclusions

Analysis of the D. nodosus rRNA genes revealed that there were three copies of each rrn locus within the D. nodosus strain A198 genome, each comprising closely linked genes encoding the 16S, 23S and 5S rRNAs, with tRNA^Ala^ and tRNA^Thr^ genes located between the 16S and 23S rRNA genes. In D. nodosus, transcription of each rrn locus appeared to originate from a single E. coli-like promoter that was mapped upstream of the 16S rRNA gene. This gene order is similar to that determined for many other bacteria, although it is not universal (Krawiec & Riley, 1990; Srivastava & Schlessinger, 1990; Menke et al., 1991).

Based on the restriction identity between the rRNA coding regions, and the sequence identity between the sequenced regions of all three rrn loci, it is concluded that the three D. nodosus rRNA operons have arisen through precise duplications, and that the multiple copies of ORFE arose as a result of its duplication along with the rRNA operons. Although restriction site differences between the three loci were observed downstream of the sequenced region of ORFE (Fig. 1), further sequence analysis will be required to determine the full extent of the identity, and the point of divergence, between the three loci. The apparent lack of heterogeneity between rrn loci may reflect the absence of sequence-associated drift, or it may indicate that the sequences within and surrounding the rRNA and tRNA genes are functionally important.

Although there is no correlation between rRNA gene copy number and genome size, an observed trend is that slow-growing organisms have fewer rRNA gene copies than more rapidly growing bacteria (Krawiec & Riley, 1990). Although rrn gene copy number must influence growth rate to a large extent, other factors such as genome size and complexity, which in turn determine an organism's physiological and metabolic capabilities, and perhaps rrn promoter strength and efficiency, must also be important. D. nodosus is a slow-growing bacterium with complex growth requirements, and it is not known if all three of its rRNA operons are required to support the maximum growth capacity of this organism. In addition, compared with fast-growing bacteria such as E. coli (Wagner, 1994) and Clostridium perfringens (Garner et al., 1991) which have multiple rrn promoters that are, or appear to be, differentially regulated, the probable presence of a single functional rrn promoter in D. nodosus may imply a less complex, albeit less efficient, system for the synthesis of rRNA in this bacterium, and may be a factor contributing to its slow growth rate.

Acknowledgements

We thank Stephen Billington and Volker Haring for helpful discussions and assistance with computing, Joan Sloan, Pauline Howarth and Tracey Warner for their most valuable technical assistance, and Eric Moses for kindly providing the D. nodosus strain A198 genomic lambda libraries. We acknowledge the strong support of the Australian Wool Research and Development Corporation. S.L. was the recipient of an Australian Wool Corporation Postgraduate Research Scholarship.

References


organization and primary structure of a ribosomal RNA operon

Bacteroides nodostls

S. Bacteroides nodostls


Dewhirst, F. E., Paster, B. J., La Fontaine, S. & Rood, J. I. (1990). Transfer of *Kingella* indologenes (Snell and Lapage 1976) to the genus *Suttonella* gen. nov. as *Suttonella indologenes* comb. nov.; transfer of *Bacteroides nodosus* (Beveridge 1941) to the genus *Dichelobacter* gen. nov. as *Dichelobacter nodosus* comb. nov.; and assignment of the genera *Cardobacterium*, *Dichelobacter*, and *Suttonella* to *Cardiobacteriaceae* fam. nov. in the gamma division of *Proteobacteria* based on 16s ribosomal ribonucleic acid sequence comparisons. *Int J Syst Bacteriol* 40, 426–433.


Smith, C. L. & Cantor, C. R. (1987). Purification, specific frag-


Received 22 May 1995; revised 21 November 1995; accepted 27 November 1995.