Primary and secondary structures of rRNA spacer regions in enterococci

Afaf Naïmi, Geneviève Beck and Christiane Branlant

Author for correspondence: Christiane Branlant. Tel: +33 3 83 91 20 91. Fax: +33 3 83 91 20 93. e-mail: cbranlant@legg.u-nancy.fr

Laboratoire d’Enzymologie et de Génie Génétique, URA CNRS 457, Université de Nancy I, Bld des Aiguillettes, BP 239-54506 Vandœuvre-Lès-Nancy Cedex, France

The 16S–23S and 23S–5S rRNA spacer DNA regions (spacer regions 1 and 2, respectively) from Enterococcus faecalis, Enterococcus faecium, Enterococcus hirae, Enterococcus durans and Enterococcus mundtii were amplified by PCR. Their nucleotide sequences were established and a secondary structure model showing the interaction between the two spacer regions was built. Whereas lactococci and Streptococcus sensu stricto are characterized by a single type of spacer region 1, the enterococci show a high degree of variability in this region; thus the spacer regions 1 with and without tRNA^Ala^ were characterized. However, as shown for lactococci and Streptococcus sensu stricto, the tRNA^Ala^ gene does not encode the 3’-terminal CCA trinucleotide. A putative antitermination signal is found downstream from the tRNA^Ala^ gene. Based on comparison with Lactococcus lactis and Streptococcus thermophilus, a double-stranded processing stem is proposed. In E. hirae, one of the three different types of spacer region 1 contains no tRNA^Ala^, but displays a 107 nt insertion that forms a long stem-loop structure. A similar insertion (115 nt in length) was found in E. faecium and base compensatory mutations preserve the ability to form the long stem-loop structure. Such insertions may correspond to mobile intervening sequences, as found in the 23S rRNA coding sequences of some Gram-negative bacteria. The spacer regions 1 and 2 from the three subgroups of streptococci were compared, and except for the tRNA^Ala^ gene and the double-stranded processing sites, little similarity was found, which opens large possibilities for future development of DNA-based typing methods.

Keywords: 16S–23S rRNA spacer, 23S–5S rRNA spacer, tRNA^Ala^, enterococci

INTRODUCTION

The enterococci are increasingly recognized as common causes of infection that may become difficult to treat because of both inherent and acquired antibiotic resistance (Hall et al., 1992). Enterococci are characterized by their ability to grow at 10 and 45 °C, in 6·5% NaCl at pH 9·6, and to hydrolyse esculin in the presence of 40% bile (Mundt, 1986). However, several of the recently described species do not conform to these phenotypic characteristics (Devriese et al., 1993). Hence, the development of rapid and reliable DNA-based methods for the identification of Enterococcus isolates is of medical importance.

Based on 16S rRNA nucleotide sequence determination and nucleic acid hybridization, the previous Bergey’s classification of Gram-positive bacteria (Hardie, 1986) has been revised, and enterococci, lactococci and Streptococcus sensu stricto were found to be three genetically distinct subgroups (Kirpisch-Bälz et al., 1982; Ludwig et al., 1985; reviewed by Schleifer & Kilpper-Bälz, 1987). According to these nucleic acid analyses, the genus Enterococcus was also divided into ‘species groups’ (Williams et al., 1991). The ‘faecium group’ consists of the Enterococcus durans, Enterococcus faecium, Enterococcus hirae and Enterococcus mundtii species, the ‘avium group’ of the Enterococcus avium, Enterococcus malodoratus and Enterococcus pseudovaum species and the ‘gallinarum group’ of the species pair Enterococcus casseliflavus and Enterococcus gallinarum. Other species, such as Enterococcus faecalis, were found to form distinct lineages. In spite of this great progress in the classification of enterococci, only
limited efforts have been made to develop DNA-based typing methods for their rapid identification. The main achievement was the design of two 20-mer oligonucleotides as 23S-rRNA-targeted probes for the identification of *E. faecium* and *E. faecalis* (Betzl et al., 1990, Beimfohr et al., 1993).

The 16S–23S spacer region (spacer region 1) from the bacterial rRNA genetic loci (*rrn*) is an important tool for the development of DNA-based typing methods because it shows a significant degree of variation in length and sequence from one species to the other (Barry et al., 1991; van der Giessen et al., 1994; review by Gurtler & Stanisich, 1996). In addition a given species may be characterized by a defined polymorphism of spacer region 1 and this may also be used for rapid characterization of bacterial species (Jensen et al., 1993; Whiley et al., 1995). The most direct and certainly the most rapid method of visualizing the polymorphic character of *rrn* spacer regions is PCR amplification of these spacer regions (Jensen et al., 1993). In the case of enterococci, only the spacer region 1 of two species, *E. hirae* and *E. faecalis* were studied, and a polymorphism was found for both of them. Two types of spacer regions 1 differing by the presence or the absence of a tRNA<sup>Ala</sup> gene were identified, by direct cloning of genomic DNA fragments in the case of *E. hirae* (Sechi & Daneco-Moore, 1993), and by PCR amplification for *E. faecalis* (Hall, 1994). Whereas the absence of a tRNA<sup>Ala</sup> gene in spacer region 1 is accompanied by a size reduction in *E. faecalis* (Hall, 1994), this is not the case in *E. hirae* (Sechi & Daneco-Moore, 1993). This polymorphism of the spacer region 1 in enterococci is in contrast with the observation of an identical spacer region 1 sequence (Nour et al., 1995) in the five or six *rrn* operons of *Lactococcus lactis*, *Streptococcus thermophilus* and *Streptococcus salivarius* (Pébay et al., 1992; Le Bourgeois et al., 1995). A similar number of *rrn* operons was proposed for *E. hirae* (Sechi & Daneco-Moore, 1993). The question is whether the polymorphism observed for the six operons of *E. hirae* (Sechi & Daneco-Moore, 1993) is a general feature of enterococci.

Based on detailed studies made on the rRNA maturation process of *Escherichia coli* (review by Apirion & Miczak, 1993), both spacer region 1 and the 23S–5S rRNA spacer region (spacer region 2) play a crucial role in pre-rRNA maturation. Studies on *Esch. coli* showed that base-pair interactions formed by the leader region upstream of 16S rRNA and spacer region 1 on the one hand, and by spacer regions 1 and 2 on the other hand, generate the two primary double-stranded processing sites (dsPS1 and dsPS2), that are cleaved by RNase III (Apirion & Miczak, 1993). Although only a limited number of *rrn* spacer regions have been characterized (reviewed by Gurtler & Stanisich, 1996), the presence of long double-stranded regions that close the 16S and 23S rRNA sequences and contain the primary processing sites seem to be a general feature of bacterial pre-rRNAs, they are not always found in archaeobacteria (Potter et al., 1995). A model of the interaction between spacer regions 1 and 2 that is supported by results of S1 nuclease digestion was proposed for *L. lactis* (Chiaruttini & Milet, 1993), and we recently proposed a model of interaction for the spacer regions 1 and 2 of *S. thermophilus* and *S. salivarius* (Nour et al., 1995). To establish a model of interaction between spacer regions 1 and 2 for the third genus of streptococci, namely *Enterococcus*, the nucleotide sequence of a spacer region 2 from enterococci had to be determined. Determination of an enterococcal spacer region 2 nucleotide sequence is also of interest for the study of SS rRNA maturation. Indeed, RNase E, which is involved in SS rRNA maturation in *Esch. coli*, has a recognition site within spacer region 2 (5'-ACAGAAUUUG-3') (reviewed by Apirion & Miczak, 1993).

Table 1. EMBL database accession numbers for the intergenic rRNA spacer regions established in this paper

<table>
<thead>
<tr>
<th>Strain</th>
<th>EMBL no.</th>
<th>Spacer region*</th>
<th>Length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. faecalis</em> ATCC 19433&lt;sup&gt;T&lt;/sup&gt;</td>
<td>X87186</td>
<td>1, tRNA&lt;sup&gt;Ala&lt;/sup&gt;</td>
<td>328</td>
</tr>
<tr>
<td><em>E. faecalis</em> ATCC 19433&lt;sup&gt;T&lt;/sup&gt;</td>
<td>X87182</td>
<td>1</td>
<td>226</td>
</tr>
<tr>
<td><em>E. faecalis</em> ATCC 19433&lt;sup&gt;T&lt;/sup&gt;</td>
<td>X87183</td>
<td>2</td>
<td>96</td>
</tr>
<tr>
<td><em>E. durans</em> ATCC 19432&lt;sup&gt;T&lt;/sup&gt;</td>
<td>X87177</td>
<td>1, tRNA&lt;sup&gt;Ala&lt;/sup&gt;</td>
<td>330</td>
</tr>
<tr>
<td><em>E. durans</em> ATCC 19432&lt;sup&gt;T&lt;/sup&gt;</td>
<td>X87178</td>
<td>1</td>
<td>277</td>
</tr>
<tr>
<td><em>E. durans</em> ATCC 19432&lt;sup&gt;T&lt;/sup&gt;</td>
<td>X87179</td>
<td>2</td>
<td>92</td>
</tr>
<tr>
<td><em>E. faecium</em> ATCC 19434&lt;sup&gt;F&lt;/sup&gt;</td>
<td>X87180</td>
<td>1</td>
<td>344</td>
</tr>
<tr>
<td><em>E. faecium</em> ATCC 19434&lt;sup&gt;F&lt;/sup&gt;</td>
<td>X87181</td>
<td>2</td>
<td>94</td>
</tr>
<tr>
<td><em>E. faecium</em> LMA631</td>
<td>X87190</td>
<td>1</td>
<td>344</td>
</tr>
<tr>
<td><em>E. faecium</em> LMA631</td>
<td>X87191</td>
<td>2</td>
<td>94</td>
</tr>
<tr>
<td><em>E. hirae</em> ATCC 8043&lt;sup&gt;T&lt;/sup&gt;</td>
<td>X87184</td>
<td>1</td>
<td>232</td>
</tr>
<tr>
<td><em>E. hirae</em> ATCC 8043&lt;sup&gt;T&lt;/sup&gt;</td>
<td>X87185</td>
<td>2</td>
<td>94</td>
</tr>
<tr>
<td><em>E. mundtii</em> ATCC 582</td>
<td>X87187</td>
<td>1, tRNA&lt;sup&gt;Ala&lt;/sup&gt;</td>
<td>331</td>
</tr>
<tr>
<td><em>E. mundtii</em> ATCC 582</td>
<td>X87188</td>
<td>1</td>
<td>230</td>
</tr>
<tr>
<td><em>E. mundtii</em> ATCC 582</td>
<td>X87189</td>
<td>2</td>
<td>89</td>
</tr>
</tbody>
</table>

* 1, 16S–23S rRNA spacer DNA; 2, 23S–5S rRNA spacer DNA.
Hence, determination of the nucleotide sequence of spacer regions 1 and 2 of a series of enterococci with a study of their polymorphism was expected to bring new insight on the structure and the evolution of the pre-rRNA spacer region in streptococci and to be a prerequisite for the future development of DNA-based methods for the identification of enterococci.

In this paper, we describe an exhaustive analysis of the spacer regions 1 and 2 from E. durans, E. faecium, E. hirae, E. mundtii and E. faecalis using PCR and direct sequence analysis of the amplification products. For each species, an RNA secondary structure model showing the interaction between the two spacer regions 1 and 2 was built. The results obtained are discussed in the light of the present knowledge on rRNA and tRNA processing. The possible utilization of the established sequences for the development of rapid DNA-based typing methods is also discussed.

METHODS

Bacterial strains and cultures. The enterococcal strains used in this study are listed in Table 1. All strains were stored at −20 °C in Brain-Heart Infusion (BHI) medium (Diagnostics Pasteur) in the presence of 15 % (v/v) glycerol. Cultures were grown at 37 °C in BHI medium for E. faecalis, E. durans, E. faecium ATCC 19434T, E. hirae and E. mundtii and in lac/MRS (De Man et al., 1960) liquid medium for E. faecium LMA631. For E. mundtii ATCC 582, BHI medium was supplemented with 10 % (v/v) horse blood as recommended by the Pasteur Institute strain distribution service. Culture purity was checked on nutrient agar (10 g tryptone l−1, 5 g yeast extract l−1, 5 g NaCl l−1, 15 g agar l−1). E. mundtii was spread on Colombia agar (Diagnostic Pasteur).

Isolation of DNA and PCR. Enterococcal genomic DNA was prepared from early stationary-phase cells according to Hill et al. (1990) with modifications to temperature and duration of incubations as previously described (Nour et al., 1995). PCR amplifications were performed as described by Nour et al. (1995). The DNA fragments encoding the 16S–23S rRNA spacer region (spacer region 1) were amplified using primer combination 1: primer 514, 5’-TGGATCACCTCCTTCTTCTGTT-3’ (sequence of the 18 nt at the 3’-end of Streptococcus pneumoniae 16S rRNA; Bacot & Reeves, 1991) and primer 554, 5’-ACCGTGTAGCGTTAGTCGCT-3’, which is complementary to a highly conserved 23S rRNA sequence (positions 1–25 of the Esch. coli 23S rRNA; Branlant et al., 1981). The DNA fragments encoding the 23S–5S rRNA spacer region (spacer region 2) were amplified using primer combination 2: primer 601, 5’-CCTTACTTGTGTTTGCAG-3’ (sequence of the S. thermophilus 23S rRNA between positions 2876 and 2900; Ludwig et al., 1989). Gels were stained with ethidium bromide and DNA was visualized by UV fluorescence. Elution was according to the methods of Sambrook et al. (1989). The complete sequences of the amplified fragments were determined on both strands as described by Nour et al. (1995). Oligonucleotides 514, 554, 601 and 667 were used as primers for sequence analysis. The annealing reaction was for 3 min at 94 °C in the annealing buffer (50 mM NaCl, 10 mM MgCl2, 40 mM Tris/HCl, pH 7.5). The annealing mixture was immediately cooled on ice for 10 min and used for elongation in the presence of T7 DNA polymerase (Pharmacia) (Tabor et al., 1987).

Computer sequence analysis. The computer program CLUSTAL w (Thompson et al., 1994) was used for sequence alignment and the Alscript program (Borton, 1993) for representation of sequence similarities. The Mfold program of the Genetics Computing Group (GCG, University of Wisconsin, USA) software version 8 was used for prediction of RNA secondary structure.
Fig. 2. Nucleotide sequence comparison of enterococcal spacer regions 1. All the enterococcal palindromic sequences al/a2 and the two pairs of directly inverted sequences (bl/b2 and dl/d2) are boxed. The borders of the region that differs in the LS1 and SS1 fragments sequenced in this study were aligned with the CLUSTAL W program (Thompson et al., 1994). The LS1 spacer regions with (LS1A) and without (LS1B) tRNA, previously sequenced for E. hirae ATCC 9790 (Sechi & Dano-Moore, 1993), are included in the comparison. The 16S and 23S rRNA sequences at the extremities of the amplified fragments were excluded from the comparison. Similarities were represented by the Alscript program (Borton, 1993). When at a given position the same nucleotide is found in at least three of the aligned sequences, it is shaded in all the sequences where it is present. Sequence heterogeneities are indicated by the usual nomenclature (M for A or C, R for A or T). The borders of the region that differs in the LS1 and SS1 fragments are denoted by a vertical line marked by a sub on the left and by a broken line marked by sub on the right. The tRNA sequence is represented by a horizontal dashed line. The sequences involved in the putative dsPS1 and dsPS2 stems are overlined. The pairs of palindromic sequences a1/a2 and the two pairs of directly inverted sequences (bl/b2 and dl/d2) are boxed.
RESULTS

Length polymorphism of the enterococcal spacer region 1

The type strain was selected for *E. faecalis*, *E. durans*, *E. faecium* and *E. hirae*, and the ATCC strain 582 was selected for *E. mundtii* (Table 1). For *E. faecalis* and *E. faecium*, a second strain was introduced in the collection for evaluation of the intraspecific degree of conservation. Primer combinations 1 and 2 (Methods) were used for PCR amplification of the *rrn* spacer regions 1 and 2, respectively. For all the species studied, primer combination 2 generated a unique amplification product of about 100 bp (Fig. 1b). A larger number of fragments were amplified with primer combination 1, and their lengths varied from one species to the other (Fig. 1a). One or two amplification products of about 300 bp were obtained for all the strains (LS1, for large spacer region 1). A smaller fragment of about 220 bp was found for *E. faecalis*, *E. mundtii* and *E. hirae* and a fragment with an intermediate size (about 270 bp) for *E. durans*. These fragments were denoted SS1, for small spacer region 1. Finally, very large fragments varying in number, size and intensity were obtained for most of the species, except *E. faecium* LMA631. It is not clear whether these large fragments (500–1000 bp) corresponded to amplified rDNA, since the yield of some of them was significantly lower than those of the amplified fragments LS1 and SS1. Similar observations were made by Hall (1994) for *E. faecalis* ATCC 19433. Thus, only the LS1 and SS1 fragments were sequenced. For *E. hirae*, the two LS1 fragments were too close to be resolved. They were not sequenced since according to their sizes, they corresponded to the two spacer regions 1 that were previously cloned and sequenced for *E. hirae* ATCC 9790 (Sechi & Danco-Moore, 1993). *E. mundtii* also had a pair of LS1 fragments, one of which was largely prominent and this was sequenced. DNA fragments amplified with primer combination 2 were also sequenced; these are denoted S2 (for spacer region 2).

Nucleotide sequences of enterococcal spacer regions 1 and 2

None of the SS1 fragments contained a tRNA gene. In contrast, all LS1 fragments, except for *E. faecium*, contained a tRNA<sup>A^1</sup> gene. The nucleotide sequences for the SS1 and LS1 fragments are aligned in Fig. 2, together with the nucleotide sequences of the two LS1s sequenced for *E. hirae* ATCC 9790 (Sechi & Danco-Moore, 1993). Sequences from DNA fragments amplified with primer combination 2 are aligned in Fig. 3.

For both spacer regions, no sequence differences were detected between *E. faecalis* ATCC 19433 and ATCC 5430, and the sequences that we established for the SS1 and LS1 fragments of these two strains are in perfect agreement with those published by Hall (1994) for ATCC 19433. This author also studied two clinical isolates (617 and 805). In the LS1 fragment, Hall (1994) found four single-base differences between strain ATCC 19433 and the two clinical isolates. At one of these four positions (position 85 in Fig. 2), an A/G heterogeneity was suspected in *E. faecalis* ATCC 19433, and our sequencing data confirmed the presence of this heterogeneity. We also detected two micro-heterogeneities in the LS1 fragment of *E. mundtii* ATCC 582 (positions 71 and 72 in Fig. 2). The LS1 fragments of *E. faecium* ATCC 19434 and LMA631 differed at four positions, at one of these positions (315 in Fig. 2), a C residue was read on the sequencing gel for *E. faecium* LMA631, whereas both C and U residues were detected at this position for *E. faecium* ATCC 19434. The three other differences (positions 233, 281 and 295 in Fig. 2) are single-base substitutions between the two strains. Micro-heterogeneity seems to be a specific feature of the LS1. No micro-heterogeneity was detected in the SS1 or S2 fragments. We should point out that detection of two different nucleotides at a given position was not the result of sequencing or PCR artefacts, since the two complementary nucleotides were also detected upon sequencing the complementary DNA strand, and the

---

**Fig. 3.** Nucleotide sequence comparison of enterococcal spacer region 2. The sequences established from the DNA fragments amplified with the primer combination 2 were aligned with the CLUSTAL W program. Similarities were represented with the Ascript program using the same rules as in Fig. 2. The nucleotide sequences involved in formation of the putative dsPS2 stem are overlined and the putative RNase E recognition sites in *E. faecium*, *E. hirae* and *E. mundtii* are boxed.
Fig. 4. For legend see facing page.
As a final comment on strain sequence differences, we should point out that we found two differences between enterococci in the LS1 fragment indicates that at least two single-base deletions and mutations in the SS1 spacer region from the same E. hirae strain are indicated: the 28 nt sequence in place of the tRNA and bordering sequences, the 107 nt insert, the rearrangement of stem-loop H2 and the single-base deletions (Δ) and mutations (circled nucleotides). The single-base deletions and mutations in the SS1 spacer region from E. hirae ATCC 8043 are indicated by Δ and nucleotides in squares, respectively. The 115 nt insert from E. faecium ATCC 19434 is shown on the left, differences in the LMA631 strain are represented by circled nucleotides. (b) Part of the LS1/spacer region 2 interaction of E. mundtii including the dsPS2 stem. Nucleotides conserved in all the sequences established are indicated by capital letters. The effects of sequence heterogeneity on the stem-loop structure upstream of the tRNA^Ala^ is shown in the lower inset (LS1*) on the left (the two variant nucleotides are circled). The putative A box of the antitermination signal and the putative RNase E site are indicated as in (a). The extremities of the segment substituted by a 4 nt sequence in the SS1 spacer region are indicated by arrows marked with sub, for substitution. The two stem-loop structures of the segment that replace the tRNA^Ala^ and bordering sequence in the E. durans SS1 spacer region are shown in the upper inset on the left. The nucleotide sequences of the dsPS2 stems of L. lactis (LI; Chiaruttini & Millet, 1993), S. thermophilus, S. salivarius (StSS; Nour et al., 1995) and B. subtilis (Bs; Ogasawara et al., 1983) are shown in the three insets on the right, respectively. Nucleotides identical to those in enterococci are indicated by capital letters.

Comparison of the SS1 and LS1 fragments for each species (Fig. 2) shows that they differ by the substitution of the tRNA^Ala^ gene and bordering sequences in LS1 with a short nucleotide sequence in SS1. Whereas 3, 4 and 8 bp sequences replaced the tRNA^Ala^ gene in the E. faecalis, E. mundtii and E. hirae ATCC 8043 SS1 fragment, respectively, a longer sequence was found for E. durans (55 bp) (Fig. 2) and this explains the greater length of the E. durans SS1 fragment. The unexpected

**Fig. 4.** Scheme for the secondary structure of the enterococcal pre-rRNA region between the dsPS1 stem and SS rRNA. A secondary structure model was established for each enterococcal species studied on the basis of the structure proposed for the corresponding pre-rRNA region of L. lactis (Chiaruttini & Millet, 1993) and S. thermophilus (Nour et al., 1995) and by use of the GCG program Mfold. (a) Interaction between the LS1A spacer region of E. hirae ATCC 9790 and S2 spacer region of E. hirae ATCC 8043. Nucleotides are numbered starting from position 1 of each spacer region. Capital letters represent nucleotides conserved in the five enterococcal species studied. Extremities of 16S and SS rRNAs are indicated by vertical arrows. The 23S rRNA and the tRNA^Ala^ are schematically represented. The putative A box of the antitermination signal and the putative RNase E site are indicated. All the differences found in the LS1B spacer region without tRNA from E. hirae ATCC 8043 and ATCC 19434 and E. faecium ATCC 19434 are indicated by Δ and nucleotides in squares, respectively. The 115 nt insert from E. faecium ATCC 19434 is shown on the left, differences in the LMA631 strain are represented by circled nucleotides. (b) Part of the LS1/spacer region 2 interaction of E. mundtii including the dsPS2 stem. Nucleotides conserved in all the sequences established are indicated by capital letters. The effects of sequence heterogeneity on the stem-loop structure upstream of the tRNA^Ala^ is shown in the lower inset (LS1*) on the left (the two variant nucleotides are circled). The putative A box of the antitermination signal and the putative RNase E site are indicated as in (a). The extremities of the segment substituted by a 4 nt sequence in the SS1 spacer region are indicated by arrows marked with sub, for substitution. The two stem-loop structures of the segment that replace the tRNA^Ala^ and bordering sequence in the E. durans SS1 spacer region are shown in the upper inset on the left. The nucleotide sequences of the dsPS2 stems of L. lactis (LI; Chiaruttini & Millet, 1993), S. thermophilus, S. salivarius (StSS; Nour et al., 1995) and B. subtilis (Bs; Ogasawara et al., 1983) are shown in the three insets on the right, respectively. Nucleotides identical to those in enterococci are indicated by capital letters.

same relative levels of the two alternative nucleotides were reproducibly observed on independently amplified DNA fragments. The presence of sequence heterogeneity in the LS1 fragment indicates that at least two rrr operons have a long version of the spacer region 1. As a final comment on strain sequence differences, we should point out that we found two differences between the spacer regions 2 of E. faecium ATCC 19434 and E. faecium LMA631 (positions 15 and 43 in Fig. 3).
length of the *E. faecium* ATCC 19434 and LMA631 and *E. hirae* ATCC 9790 spacer regions 1 without tRNA gene is explained by the presence of a 115 bp insertion in the *E. faecium* ATCC 19434 and LMA631 LS1 fragments without tRNA and a 107 bp insertion in the *E. hirae* ATCC 9790 LS1 fragment without tRNA, as compared to the SS1 fragments without tRNA in other species. These insertions are located at the same position in the two species (Fig. 2). Interestingly, the four nucleotide sequence differences between the LS1 fragments of the two *E. faecium* strains studied are located in the 115 bp insertion. The length, the position and also the sequence of the insert are highly conserved from *E. hirae* to *E. faecium* (Fig. 2). The situation for *E. hirae* is highly complex: it contains (i) *rrn* operons without the 107 bp insert, (ii) *rrn* operons without the tRNA gene and with the 107 bp insert, and (iii) *rrn* operons without the 107 bp insert but with a tRNA gene. As two high-molecular-mass DNA fragments were amplified in high yield for *E. faecium* ATCC 19434 (Fig. 1a), they may correspond to spacer regions 1 with a tRNA gene. However, no large fragment was detected for *E. faecium* LMA631, which suggests that this strain contains no spacer region 1 with a tRNA gene.

In addition to the tRNA gene, an 8–12 nt sequence on the 5’-side of the tRNA and a 20–25 nt sequence on the 3’-side of the tRNA, are absent in spacer regions 1 without the tRNA gene (Fig. 2). At the RNA level, the tRNA 3’ bordering sequence forms a stem–loop structure (Fig. 4). For each of the species studied, except for the substitution of the tRNA gene and bordering sequences, the short and long versions of the spacer region 1 had the same sequence. The only exception was *E. hirae*, but the observed differences may be due to the fact that the SS1 is from *E. hirae* ATCC 8043 and the LS1s are from *E. hirae* ATCC 9790.

In all the species studied, the tRNA gene insertion was bracketed by a pair of directly inverted sequences (5’-TACT-3’/5’-TCAT-3’, denoted b1/b2 in Fig. 2) and a pair of palindromic sequences (5’-CAGTTTTT-3’/5’-AAAACGT-3’, denoted a1/a2 in Fig. 2). Interestingly, partial mirrors of the a1/a2 palindromic sequences were found at the junctions between the constant regions and the substitution area (Fig. 2). Finally, in spite of mutations (*E. durans*) or heterogeneity (*E. mundtii*), a conserved pentanucleotide sequence, 5’-CTCAA-3’, was always found immediately upstream of the substitution area (Fig. 2). Some of these peculiar features may be related to the insertion/deletion events that generated the two kinds of spacer region 1. Two directly inverted sequences (5’-AATTTT-3’/5’-TTTTAA-3’, denoted d1/d2 in Fig. 2) also bracketed the 115 bp insertion in *E. faecium* and the 107 bp insertion in *E. hirae*.

The percentage similarities between the spacer regions 1 of the five species studied were calculated taking into consideration only the segments that were common to the LS1 and SS1 spacer regions. As the various spacer regions 1 sequenced for *E. hirae* differed by point mutations, each of them was compared individually to spacer regions 1 of other species (Table 2). In agreement with previous results on 16S tRNA (*Williams et al.*, 1991), the highest levels of similarity were observed for *E. durans* and *E. hirae* (92–96.9%), *E. durans* and *E. faecium* (96.3%) and *E. faecium* and *E. hirae* (from 90 to 96%), the lowest for *E. faecalis* compared with the four other species (82.4–91.9% similarity). The degree of nucleotide sequence variation for spacer region 2 was higher than that for spacer region 1 (Table 2). However, as for spacer region 1, the lowest scores of similarities were found when *E. faecalis* was compared to the species from the ‘faecium group’ (52.8–67% similarity), which is in accord with the idea that *E. faecalis* constitute a different phylogenetic branch.

**Secondary structure of spacer regions 1 and 2**

Using the established sequences, we built secondary structure models showing the interaction between spacer regions 1 and 2. The secondary structure models that we built contain a putative 23S tRNA double-stranded processing site (dsPS2) and, when present, the tRNA gene. (Fig. 4). The nucleotide sequence of the stem containing the putative dsPS2 site is highly conserved in the five enterococcal species (Figs 2, 3 and 4) and shows a strong sequence conservation when compared to the *L. lactis* (Chiaruttini & Milet, 1993), *S. thermophilus* and *S. salivarius* (Nour et al., 1995) dsPS2 stem (Fig. 4b). The sequences that join the extremity of the 23S tRNA to the dsPS2 stem can be folded in a similar way in the five enterococci. Four lateral stem–loop structures denoted H1–H4 can be formed and their nucleotide sequence is highly conserved (Fig. 4a).

None of the sequence variations that we detected between two isolates of a given species or two operons of a given isolate disturb the proposed secondary structure. For instance, the three differences detected between the *E. hirae* ATCC 8043 SS1 and the *E. hirae* ATCC 9790 LS1 fragments concern single-stranded nucleotides (Fig. 4a, squared nucleotides). A larger number of differences are found between the two large spacer regions 1 (LS1A and LS1B) of *E. hirae* ATCC 9790; they also do not disturb RNA secondary structure (Fig. 4a, circled nucleotides). For instance, two successive A residues in the stem–loop structure H2 are mutated into C residues in the *E. hirae* LS1B spacer region with the 107 nt insert, but the possibility of forming a stem–loop structure is conserved (Fig. 4a, inset LS1B). Two heterogeneities in the *E. mundtii* LS1 fragment are located in the stem–loop structure that precedes the tRNA gene, and stem–loop structures with similar stabilities can be formed with the two alternative sequences (Fig. 4b, inset LS1*”* circled nucleotides). The spacer regions 2 (S2) from the two *E. faecium* strains ATCC 19434 and LMA631, differ at two positions (Fig. 3). One mutation concerns a bulged nucleotide, the other one results in the replacement of a UA pair by a UA pair.

Antitermination signals were proposed to contribute to transcription efficiency of the *rrn* operons (reviewed by Berg et al., 1989). They consist of a B box that forms a
Table 2. Similarities of the rRNA spacer regions 1 (boxed) and rRNA spacer regions 2 (not boxed) of enterococci

<table>
<thead>
<tr>
<th>Species</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. <em>E. bira</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SS1</td>
<td>96-9</td>
<td>96-0</td>
<td>91-5</td>
<td>88-7</td>
<td></td>
</tr>
<tr>
<td>LS1A</td>
<td>96-0</td>
<td>92-9</td>
<td>90-2</td>
<td>86-5</td>
<td></td>
</tr>
<tr>
<td>LS1B</td>
<td>92-0</td>
<td>90-2</td>
<td>88-4</td>
<td>82-4</td>
<td></td>
</tr>
<tr>
<td>2. <em>E. dur</em></td>
<td>86-6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. <em>E. faec</em></td>
<td>97-8</td>
<td>87-7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. <em>E. munit</em></td>
<td>73-0</td>
<td>69-0</td>
<td>73-0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. <em>E. faec</em></td>
<td>59-1</td>
<td>67-0</td>
<td>64-8</td>
<td>52-8</td>
<td></td>
</tr>
</tbody>
</table>

stem–loop structure, followed by an A box. Putative antitermination signals were detected in the spacer region 1 of various Gram-positive bacteria, namely *Bacillus subtilis* (Ogasawara et al., 1983), *Streptomyces ambofaciens* (Pernodet et al., 1989), *L. lactis* (Chiaruttini & Milet, 1993), and mycobacteria (Harasawa et al., 1992). The A box proposed for the spacer region 1 of *B. subtilis* has the sequence 5'-TGTTCTTGGAAA-3' and it overlaps the dsPS2 stem (Berg et al., 1989). A sequence differing at only one position 5'-TGTTCCATTGAAA-3' is found at a similar position in all enterococcal spacer regions 1 studied (Fig. 4). In the LS1 spacer region, the stem–loop structure on the 3'-side of the tRNA^Ala^ is located at a correct distance to behave as a B box. The presence of a strong antitermination signal upstream of the tRNA^Ala^ gene may be required to overcome pauses of the RNA polymerase at the highly G+C-rich helical structures of the tRNA^Ala^.

A sequence, 5'-ACAAAAUUUA-3', very similar to the *Esch. coli* RNAse E recognition site (5'-ACAGAAUUUG-3' (Apriion & Miczak, 1993) is found upstream of the *E. faecium*, *E. bira* and *E. muniti* SS rRNA gene (Figs 3 and 4). Surprisingly, no counterparts were found in the *E. dur* and *E. faecalis* spacer regions 2 (Fig. 3).

The 115 and the 107 bp insertions that are found in the *E. faecium* and *E. bira* LS1, respectively, form long stem–loop structures (Fig. 4a). The inserts of the two species are located at the same position within spacer region 1 in an internal loop. They show 76-6% similarity and have mutations that compensate each other to preserve helix formation (Fig. 4a). Since the four differences between the two *E. faecium* strains studied are located in internal loops or the terminal loop, there seems to be a strong selection pressure to preserve the helical structure of the insert.

The spacer region 1 with a tRNA^Ala^ from *E. dur* and *E. faecalis*, *L. lactis*, *S. thermophilus* and *S. pneumoniae* were compared (Fig. 5a). The most conserved region is the tRNA gene. The two segments involved in the putative dsPS1 and dsPS2 stems are also highly conserved. Except for these segments and the 5'-terminus, the similarity observed was at the lower limit of significance. The enterococcal LS1 spacer region is larger than the spacer regions 1 of *L. lactis* and *Streptococcus sensu stricto*. The *E. dur* and *E. faecalis* spacer region 2 were also compared to the spacer region 2 of *L. lactis* and *S. thermophilus* (Fig. 5b). The degree of similarity is very poor, only a part of the sequence involved in the putative dsPS2 stem is conserved.

**DISCUSSION**

Based on the nucleotide sequence determination of the spacer regions 1 and 2, we propose a general secondary structure model showing the interaction of these two spacer regions in enterococci. This model is supported by the observation that sequence differences between species, strains or operons either compensate each other to preserve RNA secondary structure or do not affect RNA secondary structure. Our study reveals a high degree of plasticity of the enterococcal spacer region 1: most of the enterococci studied had long and short versions of the spacer region 1 that differ by the absence or the presence of a tRNA^Ala^ gene. Among streptococci, this feature seems to be unique to the *Enterococcus* genus. Indeed, spacer regions 1 without tRNA^Ala^ were found neither in *L. lactis* (Chiaruttini & Milet, 1993) nor in species from the *Streptococcus sensu stricto* genus (Bacot & Reeves, 1991; Nour et al., 1995; Whiley et al., 1995). The co-existence of spacer regions 1 with and without tRNA was already observed for *B. subtilis* (Loughney et al., 1982). However, in this case two tRNA genes, namely tRNA^Ala^ and tRNA^Il^, are present in the long version of the spacer region 1. According to the results obtained for *E. faecium* LMA631, some *E. faecium* strains may contain no tRNA gene in the spacer region 1. A complete absence of tRNA gene in all spacer regions 1 of a given species is not a frequent situation, although it has been described for mycobacteria (Ji et al., 1994; van der Giessen et al., 1994).

The enterococcal LS1 has the characteristic feature found for the spacer region 1 of other streptococci: a unique tRNA^Ala^ gene that does not encode the 3'-terminal CCA sequence. Until now, such a feature had been found only in streptococci (Bacot & Reeves, 1991; Chiaruttini & Milet, 1993; Nour et al., 1995), lactobacilli (Pittet & Hottinger, 1989) and some archaeobacteria (reviewed by Bacot & Reeves, 1991). In Gram-positive bacteria, post-transcriptional addition of the 3'-terminal CCA of tRNA has been reported previously: in a *B. subtilis* tRNA gene cluster, three out of 21 tRNA
genes (Ala, Glu and Leu) do not encode the 3'-terminal sequences that border the tRNA\textsuperscript{Ala} gene are highly variable among streptococci (Fig. 5a). Only a 5'-UUA/G-3' sequence on the 3'-side of the tRNA\textsuperscript{Ala} is conserved.

Our comparative analysis of SS1 and SS1 showed pairs of palindromic and/or directly inverted sequences that may be related to the genetic events responsible for the presence of the two forms of spacer region 1 (Fig. 2). A 5'=-CTCAA-3' sequence that always precedes the area of sequence divergence probably plays an important role in the process. Based on the identity of the nucleotide sequences of the SS1 and LS1 spacer region in a given strain, except for the substitution area, an homogenization of the nucleotide sequences from the SS1 and

---

Fig. 5. Nucleotide sequence conservation of streptococcal spacer regions 1 and 2. Multiple sequence alignment of the spacer regions 1 (a) and 2 (b) from one species of the 'E. faecium group' (E. durans, this paper), E. faecalis (this paper and Hall, 1994), L. lactis (Chiaruttini & Milei, 1993), S. thermophilus (Nour et al., 1995) were made with the CLUSTAL W program. Similarities are represented by grey shading as in Fig. 2. In (a) the sequence of the S. pneumoniae spacer region 1 (Bacot & Reeves, 1991) is included in the comparison. The sequences involved in the putative dsPS1 and dsPS2 stems are overlain. The tRNA\textsuperscript{Ala} and stem-loop structures H1, H2, H3 and H4 from enterococci (Fig. 4) are indicated.

---

832
LS1 spacer regions by frequent DNA recombination events may be postulated.

Another example of the plasticity of the enterococcal spacer region 1 is the insertion of a 115 nt stem–loop structure in the *E. faecium* spacer region 1 and of a 107 nt stem–loop structure in one of the long versions of the *E. hirae* spacer region 1 (Fig. 4a). As the inserted sequence folds into a long stem–loop structure, the insertion does not alter the overall structure formed by spacer regions 1 and 2 (Fig. 4a). One open question is whether insertion of such stem–loop structures is an ancient or a recent event. Our observation of the 115 bp insertion, in two *E. faecium* strains with very different origins, which has two sequence differences in spacer region 2 and four in spacer region 1, favours the hypothesis of an ancient event. However, according to the secondary structure they form and their occurrence in only one type of *E. hirae* *rrn* operons, the observed insertions share some common features with the mobile intervening sequences found in the 23S rRNA coding sequence of some Gram-negative bacteria (reviewed by Konkel et al., 1994). A larger collection of *E. faecium* and *E. hirae* species will have to be studied to get a more definite answer. The high similarities of *E. hirae* and *E. faecium* spacer regions 1 and 2 (Table 2) and the presence of insertions suggest a close phylogenetic relationship between these two species.

Based on their poor nucleotide sequence conservation in streptococci, the spacer regions 1 and 2 are suitable targets for the design of genus-specific oligonucleotide probes. For instance, the sequence from positions 320 to 357 and from positions 371 to 386 (referring to the alignment in Fig. 5a) that display no significant homologies with the spacer regions 1 of lactococci and *Streptococcus sensu stricto* may be used as targets for the design of enterococci-specific DNA probes. If the inserts in *E. faecium* and *E. hirae* prove to be general, their sequences may be good targets for the design of species-specific DNA probes, and based on the alignment in Figs 2 and 3, several sequences from the spacer regions 1 and 2 could be tested as targets for the design of species-specific oligonucleotide probes. Another alternative for species identification may be PCR typing, since patterns of DNA fragments amplified with primer combination 1 differ markedly among species.

**ACKNOWLEDGEMENTS**

A.N. was a fellow of the Franco–Algerian Cooperation Organization. This work was supported by the French Centre National de la Recherche Scientifique and the French Ministère de la Recherche et de l’Enseignement supérieur. G. Mulliert is strongly thanked for his help in sequence alignment, Professor G. Lefebre is thanked for providing us with the *E. faecium* strain LMA631 and for helpful discussion. K. Harthmuth is thanked for his help in the establishment of phylogenetic trees and for reading this manuscript. J. Bayeul and E. Habermacher are thanked for their technical assistance.

**REFERENCES**


Hall, L. M. C. (1994). Are point mutations or DNA rearrangements responsible for the restriction fragment length polymorphisms that are used to type bacteria? *Microbiology* 140, 197–204.


Harasawa, R., Uemori, T., Asada, K., Kato, I. & Shiragami, N.


Received 24 April 1996; revised 24 October 1996; accepted 19 November 1996.