Genetic characterization of the $\beta$-glucuronidase enzyme from a human intestinal bacterium, *Ruminococcus gnavus*

Diane Beaud, Patrick Tailliez and Jamila Anba-Mondoloni

INTRODUCTION

Glucuronidation is a major detoxification process and converts a large number of xenobiotic and endogenous substances into more-hydrophilic metabolites (Tephly & Burchell, 1990; Gueraud & Paris, 1998). Glucuronide formation is catalysed by uridine diphosphate glucuronosyl transferases (EC 2.4.1.17), and the major glucuronide synthesis activity is found in the liver (Dutton, 1978). Some of the glucuronides are secreted through the biliary route into the intestine. They are poorly reabsorbed into the bloodstream and are efficiently eliminated from the body unless they are hydrolysed by intestinal $\beta$-glucuronidase enzymes. $\beta$-Glucuronidase activity has been detected both in intestinal tissues and in intestinal bacteria. The greatest part of the activity in the caecum and in the large intestine of rats has, however, been attributed to bacterial enzymes (Rod & Midveldt, 1977).

$\beta$-Glucuronidase activity increases the enterohepatic circulation of toxic compounds. It thus plays a major role in the generation of toxic and carcinogenic metabolites which may promote tumour formation at different sites, including the large bowel (McBain & Macfarlane, 1998; Arimochi et al., 1999). Results from two different studies (Hawkesworth et al., 1971; Kim et al., 1998) suggest that $\beta$-glucuronidase activity can be considered a cancer-risk biomarker.

On the other hand, $\beta$-glucuronidases are also capable of selectively activating low-toxicity glucuronide prodrugs into highly cytotoxic agents at the tumour site, leading to a better anti-tumour effect and a reduction of systemic toxicity (De Graaf et al., 2002; Chen et al., 2003). In addition, these enzymes may play a beneficial role by releasing aglycone residues with protective effects, such as lignans, flavonoids, ceramide and glycyrrhetinic acid. That is liberated by the hydrolysis of the corresponding glucuronides.

$\beta$-Glucuronidase activity (encoded by the gus gene) has been characterized for the first time from *Ruminococcus gnavus* E1, an anaerobic bacterium belonging to the dominant human gut microbiota. $\beta$-Glucuronidase activity plays a major role in the generation of toxic and carcinogenic metabolites in the large intestine, as well as in the absorption and enterohepatic circulation of many aglycone residues with protective effects, such as lignans, flavonoids, ceramide and glycyrrhetinic acid, that are liberated by the hydrolysis of the corresponding glucuronides. The complete nucleotide sequence of a 4537 bp DNA fragment containing the $\beta$-glucuronidase locus from *R. gnavus* E1 was determined. Five ORFs were detected on this fragment: three complete ORFs (ORF2, gus and ORF3) and two partial ORFs (ORF4 and ORF5). The products of ORF2 and ORF3 show strong similarities with many $\beta$-glucoside permeases of the phosphoenolpyruvate : $\beta$-glucoside phosphotransferase systems (PTSs), such as *Escherichia coli* BglC, *Bacillus subtilis* BglP and *Bacillus halodurans* PTS Enzyme II. The product of ORF5 presents strong similarities with the amino-terminal domain of *Clostridium acetobutylicum* $\beta$-glucosidase (BgIa). The gus gene product presents similarities with several known $\beta$-glucuronidase enzymes, including those of *Lactobacillus gasseri* (69 %), *E. coli* (61 %), *Clostridium perfringens* (50 %) and *Staphylococcus aureus* (58 %). By complementing an *E. coli* strain in which the uidA gene encoding the enzyme was deleted, it was confirmed that the *R. gnavus* gus gene encodes the $\beta$-glucuronidase enzyme. Moreover, it was found that the gus gene was transcribed as part of an operon that includes ORF2, ORF3 and ORF5.

$\beta$-Glucuronidase activity increases the enterohepatic circulation of toxic compounds. It thus plays a major role in the generation of toxic and carcinogenic metabolites which may promote tumour formation at different sites, including the large bowel (McBain & Macfarlane, 1998; Arimochi et al., 1999). Results from two different studies (Hawkesworth et al., 1971; Kim et al., 1998) suggest that $\beta$-glucuronidase activity can be considered a cancer-risk biomarker.

Glucuronidation is a major detoxification process and converts a large number of xenobiotic and endogenous substances into more-hydrophilic metabolites (Tephly & Burchell, 1990; Gueraud & Paris, 1998). Glucuronide formation is catalysed by uridine diphosphate glucuronosyl transferases (EC 2.4.1.17), and the major glucuronide synthesis activity is found in the liver (Dutton, 1978). Some of the glucuronides are secreted through the biliary route into the intestine. They are poorly reabsorbed into the bloodstream and are efficiently eliminated from the body unless they are hydrolysed by intestinal $\beta$-glucuronidase enzymes. $\beta$-Glucuronidase activity has been detected both in intestinal tissues and in intestinal bacteria. The greatest part of the activity in the caecum and in the large intestine of rats has, however, been attributed to bacterial enzymes (Rod & Midveldt, 1977).

$\beta$-Glucuronidase activity increases the enterohepatic circulation of toxic compounds. It thus plays a major role in the generation of toxic and carcinogenic metabolites which may promote tumour formation at different sites, including the large bowel (McBain & Macfarlane, 1998; Arimochi et al., 1999). Results from two different studies (Hawkesworth et al., 1971; Kim et al., 1998) suggest that $\beta$-glucuronidase activity can be considered a cancer-risk biomarker.

On the other hand, $\beta$-glucuronidases are also capable of selectively activating low-toxicity glucuronide prodrugs into highly cytotoxic agents at the tumour site, leading to a better anti-tumour effect and a reduction of systemic toxicity (De Graaf et al., 2002; Chen et al., 2003). In addition, these enzymes may play a beneficial role by releasing aglycone residues with protective effects, such as lignans, flavonoids, ceramide and glycyrrhetinic acid. These molecules are active against tumours, platelet aggregation, viral infection, and allergic or inflammatory responses. Some of them have antioxidant effects (Kim et al., 1998; Jenab et al., 1999; Schmelz et al., 1999; Akao, 2000; Kim et al., 2000; Jeong et al., 2002; Kohno et al., 2002).
$\beta$-Glucuronidase activity has been detected among bacterial genera belonging to the dominant human intestinal microbiota, such as *Bacteroides*, *Bifidobacterium*, *Eubacterium* and *Ruminococcus* (Hawkesworth & Hill, 1971; McBain & Macfarlane, 1998; Akao, 1999, 2000). Among bacteria, genes encoding $\beta$-glucuronidase have been described for *Escherichia coli*, *Lactobacillus gasseri* and *Staphylococcus* sp. (Jefferson et al., 1986; Jefferson, 1989; Russel & Klaenhammer, 2001), and identified in *Clostridium perfringens*, *Staphylococcus Aureus* and *Thermotoga maritina*, for which whole genome sequences are available (Jefferson et al., 1986; Nelson et al., 1999; Russel & Klaenhammer, 2001; Shimizu et al., 2002). Previous studies have shown that $\beta$-glucuronidase is found in some bacteria of the gastrointestinal tract, such as *Eubacterium* and *Bacteroides* (Hill et al., 1971; McBain & Macfarlane, 1998; Akao, 1999, 2000). However, no $\beta$-glucuronidase gene has been described in these bacteria and in other strict anaerobes from the digestive ecosystem.

Here, we describe the identification of a new $\beta$-glucuronidase gene, *gus*, from *Ruminococcus gnavus* strain E1. *R. gnavus* E1 is a Gram-positive, strictly anaerobic strain isolated from the dominant human intestinal microbiota of a healthy donor (Ramare et al., 1993; Dabard et al., 2001). This strain is capable of expressing very high $\beta$-glucuronidase activity in culture. The nucleotide and deduced amino acid sequences of this gene are described. Additionally, its genetic environment is presented and its expression in *R. gnavus* E1 was investigated.

**METHODS**

**Bacterial strains and growth conditions.** Bacterial strains used in this study are presented in Table 1. Anaerobic strains were grown in an anaerobic chamber at 37°C in reduced BHI broth (Difco) supplemented with 5 g yeast extract l$^{-1}$ (Difco) and 5 mg haemin l$^{-1}$ (Sigma-Aldrich). *E. coli* strains were grown aerobically at 37°C in Luria–Bertani medium (Difco). *E. coli* transformants were selected by adding 100 μg ampicillin ml$^{-1}$ (Sigma-Aldrich) to the LB medium.

**DNA manipulations.** Total cellular DNA was extracted from a 10 ml overnight culture. The bacterial cells were harvested by centrifugation. The cell pellet was suspended in 0.5 ml TES buffer (10 mM Tris, 1 mM EDTA, 25% sucrose, pH 8.0) supplemented with lysozyme (40 mg ml$^{-1}$) and mutanolysin (50 U ml$^{-1}$), and incubated for 2 h at 37°C to perform cell lysis. A Nucleospin Tissue kit (Macherey-Nagel) was used to purify the nucleic acids. Restriction enzymes and ligase were purchased from Gibco-BRL (Life Technologies SARL). PCR reactions were performed in a 50 μl volume by using the TaKaRa Ex Taq kit (Takara Shuzo).

**DNA sequencing and analysis.** Nucleotide sequences were determined on both strands using the ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction kit (Perkin-Elmer) and an ABI Prism 377 DNA automated sequencer (Perkin-Elmer). Nucleotide sequences were assembled with the Staden program (Hawkesworth et al., 1971) and analysed with GCG software (Wisconsin package version 10.3, Accelrys Inc., San Diego, CA). Criteria applied for identifying putative genes were the length of ORFs (>30 codons) and translation initiation signals (start codons ATG, TTG, GTG, ATG or ATC) preceded by a potential ribosome-binding site (RBS).

The GenBank accession number for this sequence is AY307023.

**Construction of plasmids.** PCR fragments of the *gus* gene region encoding $\beta$-glucuronidase in *R. gnavus* strain E1 were obtained using several pairs of primers (Table 2) and were ligated into the pGem-T vector (Promega). Ligation mixtures were introduced by electroporation into the *E. coli* ΔuidA::kan$^R$ strain (L91), described below, and selected on LB agar plates containing 100 μg ampicillin ml$^{-1}$.

**Construction of an *E. coli* strain deleted for the gene *uidA* encoding the $\beta$-glucuronidase.** The TG1 strain, which has an active $\beta$-glucuronidase, was used to produce an isogenic derivative strain inactivated for this enzymic activity, as described by Datsenko and Wanner (2000). TG1 ΔuidA::kan$^R$ (named L91) was

---

**Table 1. Bacterial strains and plasmids**

<table>
<thead>
<tr>
<th>Strains and plasmids</th>
<th>Characteristics</th>
<th>Reference or source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>E. coli</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S123</td>
<td></td>
<td>Fujiawa &amp; Mori (1996)</td>
</tr>
<tr>
<td>TG1</td>
<td></td>
<td>Gilson (1984)</td>
</tr>
<tr>
<td>L91</td>
<td></td>
<td>This work</td>
</tr>
<tr>
<td><em>R. gnavus</em> E1</td>
<td></td>
<td>Ramare <em>et al.</em> (1993)</td>
</tr>
<tr>
<td><em>Bacteroides</em> sp. F1</td>
<td>Human faecal microflora</td>
<td>Our collection</td>
</tr>
<tr>
<td><em>C. perfringens</em> S79</td>
<td>Human faecal microflora</td>
<td>Fujiawa &amp; Mori (1996)</td>
</tr>
<tr>
<td><em>C. paraputrificum</em> 217.59</td>
<td>Rat faecal microflora</td>
<td>Our collection</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pGem-T</td>
<td>Amp$^R$, M13 ori pBR322 ori, linear T-overhang vector</td>
<td>Promega</td>
</tr>
<tr>
<td>pDB5</td>
<td>pGem-T plasmid containing 46/37 fragment</td>
<td>This work</td>
</tr>
<tr>
<td>pDB6</td>
<td>pGem-T plasmid containing 46/37 fragment</td>
<td>This work</td>
</tr>
<tr>
<td>pDB10</td>
<td>pGem-T plasmid containing 20/37 fragment</td>
<td>This work</td>
</tr>
<tr>
<td>pDB11</td>
<td>pGem-T plasmid containing 20/37 fragment</td>
<td>This work</td>
</tr>
<tr>
<td>pDB12</td>
<td>pDB10 with an internal Ndel deletion in the <em>gus</em> gene</td>
<td>This work</td>
</tr>
</tbody>
</table>
Table 2. Oligonucleotides used in this study

<table>
<thead>
<tr>
<th>Primer (orientation)</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 (R)</td>
<td>5'-CCCCAGTGAAATCCTCTCCCA-3'</td>
</tr>
<tr>
<td>3R (F)</td>
<td>5'-TGGGAGAATGTTCTGACCGG-3'</td>
</tr>
<tr>
<td>4R (R)</td>
<td>5'-CGGATGTTCTCCTGTGTAACG-3'</td>
</tr>
<tr>
<td>20 (F)</td>
<td>5'-GATTGGCATCTGGGCGATTG-3'</td>
</tr>
<tr>
<td>23 (R)</td>
<td>5'-ATCATTCTCTGATCACTGC-3'</td>
</tr>
<tr>
<td>25 (R)</td>
<td>5'-CAATGCCGCAATGCAAATG-3'</td>
</tr>
<tr>
<td>26 (F)</td>
<td>5'-GGGAAATCAGTCAATGGTT-3'</td>
</tr>
<tr>
<td>31 (R)</td>
<td>5'-CGATATAATTCGGAATATTC-3'</td>
</tr>
<tr>
<td>35 (R)</td>
<td>5'-CCATTCTCCAGCTGCGGCTG-3'</td>
</tr>
<tr>
<td>36 (F)</td>
<td>5'-CGGTTCTGCTCTGTTAGATG-3'</td>
</tr>
<tr>
<td>37 (R)</td>
<td>5'-ACTGATATCGTGTATTCTCCT-3'</td>
</tr>
<tr>
<td>38 (F)</td>
<td>5'-CAGGTTATTTGACAGCTATG-3'</td>
</tr>
<tr>
<td>46 (F)</td>
<td>5'-GAGACATCTGCTAAGCTGGT-3'</td>
</tr>
<tr>
<td>P1 (F)/domain 1</td>
<td>5'-TTGAYTYYTYTTAYAYTG-3'</td>
</tr>
<tr>
<td>P2 (R)/domain 3</td>
<td>5'-AARTNGCAARATTCAC-3'</td>
</tr>
<tr>
<td>P3 (F)/domain 2</td>
<td>5'-CAYTAYCTTGATGGCARGA-3'</td>
</tr>
<tr>
<td>P4 (R)/domain 2</td>
<td>5'-TCYTNCNGCRTANGRTARTG-3'</td>
</tr>
</tbody>
</table>

Enzyme assays. β-Glucuronidase activity in cell extracts was measured using a spectrophotometer to quantify the rate of release of p-nitrophenol (λ = 400 nm) from p-nitrophenyl-β-D-glucuronide (PNPG) (Sigma-Aldrich). Cultures (equivalent to 10^9 cells) were washed in 1 ml phosphate buffer (0.1 M sodium phosphate, pH 6-8). The cell pellet was resuspended in 1-6 ml of the same buffer. After adding 500 mg of 106 μm diameter glass beads (Sigma-Aldrich), cell suspensions were subjected to two cycles of 40 s at power 4.5 m s^-1 on a FastPrep FP120 instrument (BI0 101-Q, B10Gene), with 1 min on ice between cycles. The pellets, beads and cell debris were removed by centrifugation. The cell extracts were collected and kept temporarily on ice until the start of the assays. Protein concentrations were determined using the Folin-Lowry method (Lowry et al., 1951). For each assay, the cell extracts were diluted to a quarter in 3 ml of phosphate buffer (0.1 M sodium phosphate, pH 6-8). These diluted cell extracts were warmed to 37 °C and 1-7 ml of 0.5 mM PNPG was added. At appropriate time intervals, usually 5, 10, 15, 30 and 45 min, the reaction was stopped by adding 1-6 ml of 125 mM Na_2CO_3 to 300 μl of the reaction mixture. Optical density was measured at 400 nm. One unit of activity was defined as 1 nanomole of p-nitrophenol liberated per minute per milligram of protein. Each value presented is the mean of the results from at least three independent measurements.

Southern hybridization. To identify the longest genomic fragment carrying the gus gene and to select the enzyme to use for inverse PCR, R. gnavus E1 DNA was digested with EcoRV, PstI, NcoI, BglII and Smal in independent reactions. Digested genomic DNA (10 μg) was separated by electrophoresis through a 1 % agarose gel (Tebu, Le-Perray-en-Yvelines, France) and transferred by capillary blotting to a Hybond-N+ nylon membrane (Amersham Biosciences). Membranes were hybridized with ECL-labelled oligonucleotides (Amersham Biosciences) complementary to the coding strand of the gus gene of strain E1. Signal detection was carried out as recommended by the supplier.

Inverse PCR. EcoRV-digested DNA (10 μg) of R. gnavus E1 was circularized with a ligase kit, as recommended by the supplier (for enzyme selection, see Methods, Southern hybridization). PCR was performed with the 3R/4R primers, and the product obtained was cloned in pGem-T and sequenced using universal primers appropriate to the vector (#1211F and #1233Rev, Biolabs).

RNA extraction. RNA extractions were performed following the instructions in the High Pure RNA Kit (Roche Diagnostics). Bacteria (~2×10^9 cells) were briefly disrupted, in the presence of phenol, 10 % SDS, 3 M sodium acetate (pH 5-2) and glass beads (Sigma-Aldrich) by shaking using the Fast Prep centrifuge FP120 (BIO 101-Q, B10Gene), as described above. Cell lysates were obtained after a 15 min centrifugation and treated to isolate total RNA. The concentration and quality of the RNA were determined by measuring the A260 and A280 with a BIOphotometer (Eppendorf) and by agarose gel electrophoresis. Trace amounts of DNA in RNA samples were removed by treatment with DNase I (Invitrogen SARL).

RT-PCR analysis. cDNA synthesis was performed by reverse transcription of 2 μg RNA primed with 500 ng random primer hexamer (New England Biolabs). The reaction was carried out at 37 °C for 50 min with M-MLV Reverse Transcriptase (Ambion), as recommended by the supplier. A dilution (1:20) of the cDNA solution obtained was used for PCR amplification. PCR reactions were performed with specific oligonucleotides. As a positive control, additional reactions were performed by using genomic DNA as template. PCR amplification with non-reverse-transcriptase-treated RNA made it possible to verify the absence of DNA contamination.

RESULTS

Identification of β-glucuronidase gene fragments from different bacteria belonging to digestive microbiota

Protein sequences of β-glucuronidases of E. coli, Staphylococcus sp. and L. gasseri were aligned, and 40–50 % similarity was found (Fig. 1). Three conserved domains (d1–d3) were identified and used to design the four degenerate primers P1 to P4 (Table 2, Fig. 2). Previous studies in our laboratory have mentioned the presence of active β-glucuronidase in different species isolated from the digestive microbiota (unpublished data). We chose five species to start the genetic characterization of this activity. For each of the five bacterial strains analysed (E. coli S123, R. gnavus E1, Bacteroides sp. F1, C. perfringens S79 and Clostridium paraputrificum 217.59), PCR amplifications were performed using the primer pairs P1/P2, P1/P4 and P2/P3 (Table 2, Fig. 2). A 540 bp fragment between domains d1 and d2 was successfully obtained by PCR with oligonucleotides P1 and P4 from the genomic DNA of the five bacterial strains studied. PCR products were cloned in the pGem-T vector and sequenced. The sequence obtained with Bacteroides sp. F1 showed 65 % homology to the gene encoding the β-galactosidase of Caldicellulosiruptor lactoaceticus. The sequence obtained for C. perfringens S79 was identical to the β-glucuronidase gene sequence of C. perfringens strain.
13 (Shimizu et al., 2002), the genome of which had just become available in the TIGR database. For C. paraputridum 217.59, the nucleotide sequence obtained was similar, with 98 % identity to the β-glucuronidase gene of C. perfringens. The sequence obtained with E. coli S123 shared 90 % homology to the β-glucuronidase gene of E. coli already described. The 540 bp sequence from R. gnavus E1 shared 64 % homology with the gene encoding β-glucuronidase of L. gasseri, 61 % with that of C. perfringens and 57 % with that of E. coli. The R. gnavus E1 strain is the only species for which the gene encoding β-glucuronidase was unknown.

**Analysis of the R. gnavus β-glucuronidase locus**

The 540 bp fragment was used as a probe in Southern hybridizations to identify the longest genomic fragment carrying the gus gene. One unique EcoRV fragment of 5 kb was shown to hybridize to this probe. This fragment was further sequenced after inverse PCR with 3R/4R oligonucleotides (Table 2). The DNA sequence covered 4537 bp with a GC content of 39.2 %. This 4537 bp fragment comprised three complete ORFs, named ORF2, gus and ORF3, and two incomplete ORFs, named ORF4 and ORF5 (Fig. 3).

Sequence analysis revealed that ORF2 to ORF5 were transcribed in the same orientation, while ORF4 was transcribed in the opposite orientation. All these ORFs were preceded by a putative RBS and initiated by an ATG start codon. By sequence homology against protein databases, the gene encoding the β-glucuronidase was named gus. The corresponding protein has 603 amino acids. The gus gene product shows 69 % similarity to the L. gasseri β-glucuronidase and 61 % similarity to the E. coli β-glucuronidase encoded by the uidA gene, while it shows 59 % and 58 % similarity, respectively, to the putative gene products from C. perfringens and S. aureus for which complete genomes are available. The putative product of ORF2 corresponds to a protein of 464 amino acids which has conserved domains located in the amino-terminal part of many permeases of PTSs. This family of protein transporters is usually composed of three domains: IIA, IIB and IIC. ORF2 has 58 % homology with the EIIB–IIC domains of the Bacillus halodurans sugar phosphotransferase Enzyme II (EII) and 60 % similarity to E. coli BglF. The ORF3 product has 60 % similarity to domain IIA, and partially to the IIC domain, of EII BglP of B. subtilis, and 59 % similarity to the E. coli BglF domains found in the carboxy-terminal of EII. The peptide deduced from ORF4 has no homology with any protein in the databases while, surprisingly, the partial product of ORF5 has clear similarity with the first 57 amino acids of the C. acetobutylicum β-glucosidase (bgIA) which is normally found in PTSs (Fig. 3).

**β-Glucuronidase activity**

To determine the temporal expression of β-glucuronidase, the enzyme activity of R. gnavus strain E1 was measured during anaerobic growth (Fig. 4). Maximum β-glucuronidase activity was detected during the stationary phase.

**Complementation of E. coli strain L91 with the gus gene of R. gnavus**

To confirm that the gus gene of R. gnavus encodes the β-glucuronidase, complementation of an E. coli mutant inactivated in the uidA gene was tested. Different constructions performed with the R. gnavus β-glucuronidase locus in E. coli L91 were screened on agar plates with the substrate X-Gluc (Fig. 5a). E. coli strain L91, containing the pDB10 plasmid carrying the gus gene alone, gave blue

---

**Fig. 1.** Alignment of β-glucuronidase amino acid sequences of *Escherichia coli* (GenBank s69414), *Staphylococcus* sp. (GenBank AF354044) and *Lactobacillus gasseri* (GenBank AF305888). Identical residues are indicated by an asterisk, conserved residues by a colon and semi-conserved residues by a dot. The three consensus domains are indicated by boxes in bold. Arrows symbolize the P1, P2, P3 and P4 degenerate primers used in this study.

**Fig. 2.** Schematic diagram of the gus locus of *Ruminococcus gnavus* E1. Arrows symbolize primers used for PCR amplification and (3R/4R) for inverse PCR in order to clone and characterize the gus locus.
colonies, whereas \textit{E. coli} strain L91, containing the pDB12 plasmid carrying a truncated \textit{gus} gene, gave white colonies. To determine whether ORF2 and/or ORF3 play any role in the level of expression of the \(\beta\)-glucuronidase, the plasmids pDB5, pDB6 and pDB11 were introduced into \textit{E. coli} strain L91, and \(\beta\)-glucuronidase activity was measured. The \(\beta\)-glucuronidase activity was higher when \textit{gus} was expressed with ORF2 or ORF3 than with \textit{gus} alone (Fig. 5a). These results suggest that ORF2 and ORF3 could be involved in related mechanisms, such as substrate transport. Moreover, these complementations in \textit{E. coli} indicate that the machinery of transcription and translation of a Gram-negative bacterium is able to recognize Gram-positive expression signals.

Transcription analysis of the \textit{gus} locus

Northern analysis was first performed on total \textit{R. gnavus} RNA to determine the size of the mRNA \textit{gus} gene. More than 90 \(\mu\)g of total RNA was loaded on an agarose gel to identify any discrete band using a \(^{32}\text{P}\)-random-labelled PCR (23/26) fragment. No clear result was obtained, even with a large quantity of total RNA. We therefore checked whether \textit{gus} is part of an operon including ORF2 and ORF3 by using RT-PCR experiments on RNA extracted from strain E1 in the late-exponential growth phase. All the specific RT-PCR products corresponding to ORF2, the \textit{gus} gene, ORF3 and ORF5 were correctly amplified, suggesting that these ORFs were efficiently transcribed. Moreover,
RT-PCR fragments overlapping ORF2 and the gus gene, and between the gus gene and the downstream part of ORF3 through to ORF5, were also obtained (Figs 5b and 6), suggesting that the mRNA of gus was co-transcribed from ORF2 to ORF5. In the absence of reverse transcription, no PCR amplification was obtained, demonstrating that no DNA contamination occurred in the RNA extract. These results suggest that the potential promoter directing expression of the β-glucuronidase is probably located upstream from ORF2 and is recognized by the transcriptional machinery of E. coli. Careful examination of the nucleotide sequence upstream from ORF2 did not reveal any potential promoter, in spite of the presence of a potential catabolic repression element box (CRE box) 170 nucleotides upstream from the start codon of ORF2.

**DISCUSSION**

In this work, we report the identification and cloning of the gus gene of R. gnavus strain E1 by sequence homology. Its activity was confirmed by complementation of an E. coli ΔuidA::kanR mutant that we constructed. Moreover, the three amino acids described as belonging to the active site of this family of enzymes are also present in the Gus product (Islam *et al.*, 1999). R. gnavus E1 is a Gram-positive, strictly anaerobic strain isolated from the human intestinal microbiota of a healthy donor. This species has been reported to be one of the predominant bacteria of the human large bowel (Suau *et al.*, 1999). Previous studies have shown that β-glucuronidase activity is found in some inhabitants of the gastrointestinal tract, such as *Eubacterium* sp, and *Bacteroides* (Hawkesworth *et al.*, 1971; McBain & Macfarlane, 1998; Akao, 1999, 2000), but up to now, no β-glucuronidase gene has been described in them. Despite the physiological importance of β-glucuronidase to human health, only the genetic elements encoding the β-glucuronidase enzymes of *E. coli* and *L. gasseri*, two subdominant intestinal bacteria, have been identified and studied (Jefferson *et al.*, 1986; Wilson *et al.*, 1987).
1992; Russel & Klaenhammer, 2001). The discovery of this new bacterial gus gene in a predominant bacterium of the large bowel will provide a new context in which to study the effects of bacterial β-glucuronidase on gastrointestinal health and disease.

The genetic organization of the R. gnavus gus gene was surprising, as the gus gene is inserted inside genes involved in PTSs. This organization has nothing in common with any other Gram-positive or Gram-negative PTS that has been described, nor with other β-glucuronidases previously studied, such as the E. coli uidA gene (Wilson et al., 1992) or the L. gasseri gusA gene (Russel & Klaenhammer, 2001). In E. coli, the gusA gene belongs to the uidRABC operon, in which uidR encodes the transcriptional repressor while uidB encodes the transporter, and a membrane-associated protein (uidC) is involved in the glucuronide transport system (Wilson et al., 1992; Liang et al., 2005). In L. gasseri, unlike in E. coli, the gusA gene is transcribed as a monocistronic unit (Russel & Klaenhammer, 2001). The DNA surrounding the L. gasseri gusA gene does not reveal any gene encoding an obvious transport protein, indicating that L. gasseri uses either a specific transporter located elsewhere on the chromosome or an alternative transporter for glucuronide uptake (Russel & Klaenhammer, 2001). However, in R. gnavus strain E1, the gus gene is transcribed as part of an operon of at least four ORFs. Due to this genetic organization, and taking into account the results of complementation studies in E. coli, the gus-flanking ORF2 and ORF3 are good candidates for transport proteins. ORF2 shows homology with the membrane-spanning domain IIC of the enzyme EII of the PTSs of E. coli and B. subtilis. ORF3 has two potential start codons preceded by a potential RBS. The longest product of ORF3 is a 195 amino acid protein and the shortest is a 167 amino acid protein. Both products show homology with the B. subtilis bglP product. This unusual organization of the β-glucuronidase locus suggests that this enzyme could be active on phospho-β-glucuronide.

One can suppose that the gene encoding this activity has been inserted inside the PTS operon by horizontal gene transfer. However, the GC content of the gus gene, 39-9%, does not show any divergence from the PTS genes in which it is included (GC content 37.7–41.8%). The GC composition of the gus locus is not significantly lower than that of R. gnavus E1 chromosomal DNA (43%). This trait suggests that this insertion is not a recent event or that this fragment comes from another organism having a similar GC content.

The regulation of uidA expression has been well studied in E. coli. This expression is controlled at the transcriptional level by the specific repressor UidR (Novel & Novel, 1976) and by catabolic-responsive elements (Jefferson et al., 1986). The repressor blocks the transcription of uidABC in the absence of methylglucuronide. Expression is induced by a variety of glucuronides and is subject to catabolic repression via cAMP. In L. gasseri, the regulation of β-glucuronidase expression does not involve catabolic repression or glucuronide induction, even in the presence of a potential repressor upstream from the gene encoding the glucuronidase. In R. gnavus, the structural organization of the gus gene embedded in the IIB, IIC and IA domains of the well-described PTS EII would suggest a co-regulation of β-glucuronidase activity by catabolic repression. As the activity was higher during the stationary phase, it is possible that activation of this gene product depends on sugar availability. As no chemically defined media were available for R. gnavus growth, it was not possible to study what kind of sugar will not induce catabolic repression in this species. In low-GC Gram-positive bacteria where catabolic repression has been described, the genes are regulated by the catabolic control protein CcpA. CcpA is a global regulator that binds to catabolic-responsive elements (CREs) in promoter regions to control carbon catabolic repression (Stülke & Hillen, 1999). The genes regulated by CcpA share a common box upstream from the start codon called the CRE box (TGNANCNT-NNCA). A potential box (TGAGAAGGTTAACA) was found 170 bp upstream from the ORF2 start codon, suggesting that β-glucuronidase expression is submitted to catabolic repression, as in E. coli. Because of the limited extent of the R. gnavus DNA sequence, this CRE box has not yet been described, and as neither genetic tools nor chemically defined media are available, further experiments are needed to check our hypothesis on the regulation of β-glucuronidase expression in this species. In order to obtain a better insight into the regulation of R. gnavus β-glucuronidase expression, studies are currently being carried out using the Gram-positive bacterium Lactococcus lactis, for which genetic tools and chemically defined media are available. Additionally, the gus gene can be used to construct a reporter gene vector specific to Gram-positive bacteria.

ACKNOWLEDGEMENTS

This work was supported by the INRA and the Île-de-France region and is part of a PhD thesis to be presented by D. B. We thank P. Langella, J. Doré, S. Rabot, M. Flores and G. Corthier for their advice during the preparation of the manuscript.

REFERENCES

Dabard, J., Bridonneau, C., Phillipie, C. & 8 other authors (2001). Ruminococcin A, a new lantibiotic produced by a Ruminococcus


