Comparison of *Bacteroides thetaiotaomicron* and *Escherichia coli* 16S rRNA gene expression signals

Matthew D. Mastropaolo,†‡ Mary L. Thorson† and Ann M. Stevens

Department of Biological Sciences, Virginia Tech, 219 Life Sciences 1, Washington Street, Blacksburg, VA 24061-0910, USA

There are barriers to cross-expression of genes between *Bacteroides* spp. and *Escherichia coli*. In this study, a *lux*-based reporter system was developed for *Bacteroides* and used to compare the promoter structure and function of a *Bacteroides thetaiotaomicron* 4001 (BT4001) 16S rRNA promoter with those of *E. coli* in vivo. Analysis of the BT4001 sequences upstream of the 16S rRNA gene revealed the same overall structure known for *E. coli* 16S rRNA promoters in that there were two promoters separated by ~150 bp. However, the BT4001 16S rRNA promoter contains the proposed *Bacteroides* −7 and −33 consensus sequences instead of the *E. coli* −10 and −35 consensus sequences. The biological activity of various configurations of the BT4001 16S rRNA promoter was analysed. Experiments pairing the BT4001 16S rRNA promoter with an *E. coli* RBS, and vice-versa, confirmed that gene expression between the two species is restricted at the level of transcription. In *Bacteroides*, a difference in translation initiation also appears to limit expression of foreign genes.

INTRODUCTION

*Bacteroides* spp. make up a large portion of the microbial flora in the human colon, and some species are the primary cause of anaerobic infections in humans (Werner, 1974). They are as phylogenetically distinct from Gram-negative *γ*-proteobacterial enteric bacteria as they are from Gram-positive bacteria. There are barriers that impede successful expression of genes from *Escherichia coli* and other Gram-negative species in *Bacteroides* (Smith et al., 1992). It is thought that this difference in gene expression lies in part at the level of transcription initiation and is due to the sequences within the promoter region itself. In *Bacteroides*, two conserved regions analogous to the −10 and −35 housekeeping consensus sequences in *E. coli* have been found, but their sequences are significantly different (Fig. 1a). One region, TTTG, is centred at −33, and the other, TAnnTTTTT, is centred at −7 (Bayley et al., 2000). Interestingly, a promoter consensus sequence similar to that of *Bacteroides* has also been found in *Flavobacterium hibernum*, another member of the phylum *Bacteroidetes*, with TTG (−33) and TAnnTTTTT (−7) (Chen et al., 2007b).

†These authors contributed equally to this work.

‡Present address: Center for Adaptation Genetics and Drug Resistance, Department of Molecular Biology and Microbiology, Tufts University School of Medicine, 136 Harrison Ave, Boston, MA 02111, USA.

Abbreviations: BT4001, *Bacteroides thetaiotaomicron* 4001; CAT, chloramphenicol acetyltransferase; RLU, relative light units.

A supplementary table, listing primers used in this study, is available with the online version of this paper.

In this study, the structure and function of a *Bacteroides thetaiotaomicron* 4001 (BT4001) 16S rRNA promoter was compared with that of an *E. coli* 16S rRNA promoter from the *rrnB* operon. In *E. coli*, transcription starts from two promoters (P1 and P2) that are separated by approximately 150 bp, each containing −10 and −35 sites. P1 is located ~350 bp upstream, while P2 is ~200 bp upstream, from the 5′ end of the mature 16S rRNA. Three antitermination sequences located downstream of P2 are thought to allow efficient transcription of long, highly structured untranslated rRNA. *E. coli* rRNA synthesis is regulated by the collective response of the dual promoters to growth rate (Gourse et al., 1996), amino acid starvation (Cashel et al., 1996; Gafny et al., 1994) and rRNA gene dose (Gourse et al., 1996). P1 is considered the stronger of the two promoters since it accounts for most of rRNA transcription at all but the slowest growth rates (Sarmientos & Cashel, 1983). It has been suggested that P2 promoters are inhibited by transcription from P1 (Gafny et al., 1994). P2 promoters are low-level constitutive promoters, responsible for most of the rRNA expression at low growth rates when transcription from P1 is reduced (Murray et al., 2003; Sarmientos & Cashel, 1983). P1 of *E. coli* contains sequences upstream of its −35 hexamer, including an UP element (Rao et al., 1994; Ross et al., 1993) and Fis-binding sites (Ninnemann et al., 1992) that are responsible for strong stimulation of promoter activity, but are not required for growth rate regulation. P2 also contains an UP element; however, it has a much smaller effect on promoter activity than that exerted by the P1 UP element (Ross et al., 1993). Both P1 and P2 contain a G+C rich
observed that (1) E. coli transcription between in vitro study evaluated the structure of Microbiology 2684 M. D. Mastropaolo, M. L. Thorson and A. M. Stevens et al. in vitro. A previous polymerase with E. coli E. coli not promote expression of both holoenzymes (Vingadassalom et al., 1995; Murray et al., 2003). Another in vitro study examined the differences in expression. These reporters will be useful tools for the Bacteroides–Cytophaga–Flavobacterium group.

**METHODS**

**Bacterial strains, media and growth conditions.** E. coli strains JM109 (Yanisch-Perron et al., 1985), DH5α (Hanahan, 1983) and S17-1 (Simon et al., 1983) were grown aerobically in Luria–Bertani (LB) broth at 37 °C with shaking (250 r.p.m.) and the appropriate antibiotics required for maintenance of the plasmids, 100 µg ampicillin ml⁻¹ (Ap100), 10 µg tetracycline ml⁻¹ (Tc10), 25 µg chloramphenicol ml⁻¹ (Cm25) or 100 µg kanamycin ml⁻¹ (Kan100). BT4001 (Shoemaker et al., 1996) was grown anaerobically at 37 °C in 10 ml culture tubes containing trypticase yeast extract glucose broth (TYG; Holdeman & Moore, 1975), 10 µg erythromycin ml⁻¹ (Em10) and 200 µg gentamicin ml⁻¹ (Gm200). Stock tubes of B. thetaiotaomicron were kept in 7 ml PRAS chopped meat broth (Remel) for up to 2 weeks.

**Isolation of a BT4001 16S rRNA promoter.** As this work was initiated prior to publication of the genome sequence (Xu et al., 2003), an inverse PCR method was used to obtain the sequence of the region upstream of the 16S rRNA promoter. Primers 89R and 599F (Supplementary Table S1) were then designed to amplify the putative BT4001 16S rRNA promoter from chromosomal DNA. The resulting PCR fragment was cloned into pGEM-T (Promega; pMLT2). Primers MLT2B and MLT2A (Supplementary Table S1) were used to add flanking PstI and BamHI sites to the end of the cloned PCR product (pMLT7).

**Development of a bioluminescent gene reporter.** Two reporters were constructed, pMMD and pMMT series vectors (Fig. 2), containing the entire lux operon (luxCDABE) or just containing region. A previous in vitro study examined the differences in transcription between E. coli and Bacteroides. It was observed that (1) Bacteroides fragilis σ^70 factor could bind E. coli core RNA polymerase but did not allow transcription initiation, (2) purified E. coli core RNA polymerase plus σ^70 was able to bind Bacteroides promoters but not E. coli promoters, and (3) Bacteroides core RNA polymerase with E. coli σ^70 could form a complex and promote expression of both E. coli and Bacteroides promoters, but with less efficiency than the native holoenzymes (Vingadassalom et al., 2005). Another in vitro study evaluated the structure of σ^70 and its interactions with DNA (Vingadassalom et al., 2007).

Besides differences in promoter regions and sigma factor specificity, there is also a difference in the RBS between E. coli and Bacteroides. A canonical E. coli RBS is AGGAGGU (Fig. 1b; Weaver, 1999). Although the ribosomal machinery does not demand a precise distance, the optimal spacing between the Shine–Dalgarno sequence and the transcription initiation site is required for stringent control (Josaitis et al., 1999). Based on that sequence, Shine–Dalgarno sequences of E. coli (Hawley & McClure, 1983) and B. fragilis (Bayley et al., 2000). (c) Sequences of the −33 and −7 regions of the P1 and P2 promoters from BT4001 as well as the substitutions made in the P1 region of the 16S rRNA promoter. The −33 and −7 regions and the transcription initiation sites (+1) of the P1 and P2 promoters are underlined. The spacing between the −33 and −7 regions is 21 and 20 bases for P1 and P2, respectively. Substitutions made in these two regions are boxed. The P1 promoter with no upstream sequence starts 1 bp upstream of the first base of the −33 region.

**Fig. 1.** Comparison of E. coli and B. fragilis promoter consensus sequences (a), Shine–Dalgarno sequences (b) and promoter substitutions (c). (a) Promoter consensus sequences of E. coli (Hawley & McClure, 1983) and B. fragilis (Bayley et al., 2000). (b) Shine–Dalgarno sequences of E. coli (Weaver, 1999) and B. fragilis (Tribble et al., 1999). (c) Sequences of the −33 and −7 regions of the P1 and P2 promoters from BT4001 as well as the substitutions made in the P1 region of the 16S rRNA promoter. The −33 and −7 regions and the transcription initiation sites (+1) of the P1 and P2 promoters are underlined. The spacing between the −33 and −7 regions is 21 and 20 bases for P1 and P2, respectively. Substitutions made in these two regions are boxed. The P1 promoter with no upstream sequence starts 1 bp upstream of the first base of the −33 region.

region (Travers, 1984) between the −10 and the transcription start site that is required for stringent control (Josiatis et al., 1995; Murray et al., 2003).

Our in vivo study complements the in vitro work done by others (Vingadassalom et al., 2005, 2007) and further demonstrates that gene expression capabilities between Bacteroides and E. coli depend on unique promoter structure. In our study the role of the RBSs of E. coli and Bacteroides has also been analysed. We developed two luminescence-based reporter systems that utilize the lux genes from the symbiotic luminescent soil bacterium Photorhabdus luminescens for use in Bacteroides species to comparatively examine the role of promoter elements and the RBS in controlling gene expression. These reporters will be useful tools for the Bacteroides–Cytophaga–Flavobacterium group.

**Isolation of a BT4001 16S rRNA promoter.** As this work was initiated prior to publication of the genome sequence (Xu et al., 2003), an inverse PCR method was used to obtain the sequence of the region upstream of the 16S rRNA promoter. Primers 89R and 599F (Supplementary Table S1) were then designed to amplify the putative BT4001 16S rRNA promoter from chromosomal DNA. The resulting PCR fragment was cloned into pGEM-T (Promega; pMLT2). Primers MLT2B and MLT2A (Supplementary Table S1) were used to add flanking PstI and BamHI sites to the end of the cloned PCR product (pMLT7).

**Development of a bioluminescent gene reporter.** Two reporters were constructed, pMMD and pMMT series vectors (Fig. 2), containing the entire lux operon (luxCDABE) or just containing...
replication is located. The transcription into the downstream region, where the origin of the second transcription terminator was used to prevent read-through from the phage t0 transcription terminator was inserted between the 3' end of the promoter. A second series of vectors was created, for easy one-step cloning of promoters into the lux reporter. pMMT series vectors were constructed by cloning transcription terminators and the luxAB cassette into pGWA48.3 (Whittle et al., 2002). The E. coli rrrB T1T2 transcription terminators were amplified from the plasmid pK223-3 (Aman et al., 1983) using primers T1T2F3 and T1T2R3 (Supplementary Table S1), and cloned into pGEM-T. Both of these fragments were cloned into pMDM7A using KpnI and EcoRI, creating pMLT15 (Bacteroides RBSs) and pMLT16 (E. coli RBSs). A second transcription terminator was used to prevent read-through transcription into the downstream region, where the origin of replication is located. The phage t0 transcription terminator was amplified from the plasmid pZA3p16S (Lutz & Bujard, 1997) using primers T0F and T0R (Supplementary Table S1), which amplified the region with flanking EcoRI sites. It was cloned into pMLT15 and pMLT16, creating pMMT1 and pMMT2, with both P1 and P2 with no upstream sequence of P1, were isolated from pMLT7 via PCR, using primers listed in Supplementary Table S1, which introduced a PstI site upstream and a BamHI site downstream. Primers MLT12 and MLT17 were used for P1, MLT3A and MLT37B for P2, and MLT7A and MLT37B for P1 and P2 with no upstream sequences.

Substitutions were made to the P1 promoter region (Fig. 1c) using template pMLT12 and primers listed in Supplementary Table S1. The first round of PCR was done using DeepVent polymerase (New England Biolabs) with the forward primers 33P1, 7P1U and 7P1D, used for the second round of PCR with the primer MLT2B, which adds a BamHI site to the 3' end. The P1 promoter with no upstream sequence was amplified using the primers MLT7A and MLT7C. The products were subsequently gel-extracted using a Qiagen Gel Purification kit (Qiagen) and used for a second round of PCR with the primer MLT12B, which adds a PstI site to the 3' end, to obtain full-length P1.

Bacteroides promoters of interest were cloned into pMMT1 or pMMT2 via PstI and BamHI sites, while PstI and KpnI sites were used for the E. coli rrrB P1. See Table 1 for designs and designations of these promoter–reporter fusions.

**Conjugation procedure.** All of the E. coli–Bacteroides lux reporter/ shuttle vectors were transferred into BT4001 by diparental mating using E. coli strain S17-1 (Simon et al., 1983) using standard filter-mating protocols (Shoemaker et al., 1986).

**In vitro luciferase assays.** E. coli JM109 and BT4001 strains containing lux fusion constructs were grown to OD600 0.5, 1 ml aliquots of each were sedimented by centrifugation and the pellets were stored at −70 °C. Luciferase assays using the harvested crude cell extracts were performed as previously described (Finney et al., 2002). In brief, each cell pellet was resuspended in 1 ml lysis buffer (10 mM EDTA, pH 7.0, 10 mM potassium phosphate, 5 mg lysozyme, 1 mM DTT, 0.1 µg BSA per 100 ml). The samples were incubated at −70 °C for 1 h and thawed at room temperature prior to analysis of the activity of aliquots. Activities were expressed as

---

**Fig. 2.** Maps of two E. coli–Bacteroides shuttle/expression vectors. (a) pMDM1 is a fusion of pVAL1 with pJMP20 using the restriction sites EcoRI and PstI. The cepA promoter was previously cloned into pT205 via PstI and BamHI, creating pJMP20. (b) pMMT1 contains luxAB with flanking KpnI and EcoRI sites. The t0 transcription terminator and the T1T2 transcription terminator are flanked by EcoRI and HindIII restriction sites, respectively. The multiple cloning sequence contains four unique sites, PstI, Smal, BamHI and KpnI, which may be used for cloning additional promoters. See main text for details of construction.
Relative light units (RLU) using a TD-20/20 luminometer (Turner Designs). The fold increase was defined as the RLU of the promoter construct divided by the RLU of the promoterless negative control.

In vivo whole-cell luminescence assays. Cellular luminescence was measured from BT4001 strains carrying the pMDM series as well as the pMMT1 constructs with the cepA promoter and two variant forms (Table 1). Assays were then performed as described by Phillips-Jones (1993). In brief, the cells were grown overnight in 10 ml TYG with the appropriate antibiotics, subcultured into fresh medium and grown to OD₆₀₀ 0.5. A 1 μl volume of 10% sonicated n-decyl aldehyde (decanal; Sigma) was added to 100 μl live cells and the RLU were monitored over a period of 10 min using a TD-20/20 luminometer.

Primer extension. Primers BtPE1 and MLT37B (Supplementary Table S1) were used to examine the transcription start site of BT4001 p1 and p2, respectively. Each primer was 5’ end-labelled using [γ-³²P]dATP (Amersham) and a Primer Extension System kit (Promega). Primer extension was performed using this kit according to the manufacturer’s instructions, except that an ethanol precipitation step was used to remove unincorporated [γ-³²P]dATP from the primer. For ethanol precipitation, 90 μl RNase-free water, 11 μl 3 M RNase-free sodium acetate and 220 μl ethanol were added, mixed by vortexing, and kept at room temperature for 1 h. Sequencing reactions of purified plasmids were performed using an fmol DNA cycle sequencing system kit (Promega) according to the manufacturer’s instructions. Samples were analysed by electrophoresis in a 6% polyacrylamide-urea sequencing gel.

Statistics. Results were analysed using Instat 3 software (Graphpad Software) with an unpaired Student’s t test. A two-tailed P value <0.05 was considered statistically significant.

Table 1. Plasmids used in this study

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Relevance to study*</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pGEM-T</td>
<td>TA cloning vector used for sequencing, Ap&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Promega</td>
</tr>
<tr>
<td>pUC19</td>
<td>Cloning vector used for T&lt;sub&gt;1&lt;/sub&gt;T&lt;sub&gt;2&lt;/sub&gt; transcription terminator, Ap&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Yanisch-Perron et al. (1985)</td>
</tr>
<tr>
<td>pGWA48.3</td>
<td>E. coli–Bacteroides shuttle vector, Kan&lt;sup&gt;R&lt;/sup&gt;, Em&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Whittle et al. (2002)</td>
</tr>
<tr>
<td>pVAL1</td>
<td>E. coli–Bacteroides shuttle vector, Tc&lt;sup&gt;R&lt;/sup&gt;, Ap&lt;sup&gt;R&lt;/sup&gt;, Em&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Valentine et al. (1988)</td>
</tr>
<tr>
<td>pKK223-3</td>
<td>Source of the T&lt;sub&gt;1&lt;/sub&gt;T&lt;sub&gt;2&lt;/sub&gt; transcription terminator, Kan&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Amann et al. (1983)</td>
</tr>
<tr>
<td>pZ3AP&lt;i&gt;lac&lt;/i&gt;</td>
<td>Source of the t&lt;sub&gt;0&lt;/sub&gt; transcription terminator, Cm&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Lutz &amp; Bujard (1997)</td>
</tr>
<tr>
<td>pFD711</td>
<td>Source of wild-type cepA promoter, Spc&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Bayley et al. (2000)</td>
</tr>
<tr>
<td>pFD714</td>
<td>Source of variant cepA promoter (0% activity), Spc&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Bayley et al. (2000)</td>
</tr>
<tr>
<td>pFD729</td>
<td>Source of variant cepA promoter (33% activity), Spc&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Bayley et al. (2000)</td>
</tr>
<tr>
<td>pMLT7</td>
<td>Full-length BT4001 16S rRNA promoter, flanked by PsI and BamHl sites in pGEM-T</td>
<td>This study</td>
</tr>
<tr>
<td>pMLT12</td>
<td>BT4001 16S rRNA P1 flanked by PsI and BamHl sites in pGEM-T</td>
<td>This study</td>
</tr>
<tr>
<td>pMLT15</td>
<td>&lt;i&gt;P. luminescens luxAB&lt;/i&gt; with Bacteroides RBS downstream from T&lt;sub&gt;1&lt;/sub&gt;T&lt;sub&gt;2&lt;/sub&gt; in pMDM7A-7</td>
<td>This study</td>
</tr>
<tr>
<td>pMLT16</td>
<td>&lt;i&gt;P. luminescens luxAB&lt;/i&gt; with &lt;i&gt;E. coli&lt;/i&gt; RBS downstream from T&lt;sub&gt;1&lt;/sub&gt;T&lt;sub&gt;2&lt;/sub&gt; in pMDM7A-7</td>
<td>This study</td>
</tr>
<tr>
<td>pMDM1</td>
<td>Wild-type cepA promoter construct, with entire lux operon in pVAL1&lt;sup&gt;+&lt;/sup&gt;, Tc&lt;sup&gt;R&lt;/sup&gt;, Ap&lt;sup&gt;S&lt;/sup&gt;, Em&lt;sup&gt;R&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pMDM4</td>
<td>Promoterless lux operon reporter control for pMDM series&lt;sup&gt;+&lt;/sup&gt;, Tc&lt;sup&gt;R&lt;/sup&gt;, Ap&lt;sup&gt;S&lt;/sup&gt;, Em&lt;sup&gt;R&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pMDM7A-7</td>
<td>T&lt;sub&gt;1&lt;/sub&gt;T&lt;sub&gt;2&lt;/sub&gt; &lt;i&gt;E. coli&lt;/i&gt; transcription terminators in pGWA48.3</td>
<td>This study</td>
</tr>
<tr>
<td>pMDM8</td>
<td>BT4001 16S rRNA promoter–lux operon reporter construct in pVAL1&lt;sup&gt;+&lt;/sup&gt;, Tc&lt;sup&gt;R&lt;/sup&gt;, Ap&lt;sup&gt;S&lt;/sup&gt;, Em&lt;sup&gt;R&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pMMT1</td>
<td>luxAB reporter system with the &lt;i&gt;Bacteroides&lt;/i&gt; consensus RBS, Kan&lt;sup&gt;R&lt;/sup&gt;, Em&lt;sup&gt;R&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pMMT1A</td>
<td>&lt;i&gt;E. coli&lt;/i&gt; rrnB P1 promoter in pMMT1</td>
<td>This study</td>
</tr>
<tr>
<td>pMMT1B</td>
<td>BT4001 full-length 16S rRNA promoter in pMMT1</td>
<td>This study</td>
</tr>
<tr>
<td>pMMT1C</td>
<td>BT4001 16S rRNA P1 promoter in pMMT1</td>
<td>This study</td>
</tr>
<tr>
<td>pMMT1D</td>
<td>BT4001 16S rRNA P2 promoter in pMMT1</td>
<td>This study</td>
</tr>
<tr>
<td>pMMT1E</td>
<td>BT4001 P1P2 16S rRNA promoter with no upstream sequence in pMMT1</td>
<td>This study</td>
</tr>
<tr>
<td>pMMT1F</td>
<td>Wild-type cepA promoter in pMMT1</td>
<td>This study</td>
</tr>
<tr>
<td>pMMT1G</td>
<td>Variant cepA promoter construct (0% activity) in pMMT1</td>
<td>This study</td>
</tr>
<tr>
<td>pMMT1H</td>
<td>Variant cepA promoter construct (33% activity) in pMMT1</td>
<td>This study</td>
</tr>
<tr>
<td>pMMT1I</td>
<td>BT4001 16S rRNA P1 promoter with substituted −7 downstream region in pMMT1</td>
<td>This study</td>
</tr>
<tr>
<td>pMMT1J</td>
<td>BT4001 16S rRNA P1 promoter with a substituted −7 upstream region in pMMT1</td>
<td>This study</td>
</tr>
<tr>
<td>pMMT1K</td>
<td>BT4001 16S rRNA P1 promoter with a substituted −33 region in pMMT1</td>
<td>This study</td>
</tr>
<tr>
<td>pMMT1L</td>
<td>BT4001 16S rRNA P1 promoter with no upstream sequence in pMMT1</td>
<td>This study</td>
</tr>
<tr>
<td>pMMT2</td>
<td>luxAB reporter system with the &lt;i&gt;E. coli&lt;/i&gt; consensus RBS, Kan&lt;sup&gt;R&lt;/sup&gt;, Em&lt;sup&gt;R&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pMMT2A</td>
<td>&lt;i&gt;E. coli&lt;/i&gt; rrnB P1 promoter in pMMT2</td>
<td>This study</td>
</tr>
<tr>
<td>pMMT2B</td>
<td>BT4001 full-length 16S rRNA promoter in pMMT2</td>
<td>This study</td>
</tr>
<tr>
<td>pMMT2C</td>
<td>BT4001 16S rRNA P1 promoter in pMMT2</td>
<td>This study</td>
</tr>
</tbody>
</table>

*For use in <i>E. coli</i>: Tc<sup>R</sup>, tetracycline resistance; Ap<sup>R</sup>, ampicillin resistance; Tp<sup>R</sup>, trimethoprim resistance; Kan<sup>R</sup>, kanamycin resistance; Cm<sup>R</sup>, chloramphenicol resistance; Spc<sup>R</sup>, spectinomycin resistance; Str<sup>R</sup>, streptomycin resistance. For use in <i>Bacteroides</i>: Rif<sup>R</sup>, rifampicin resistance; Gm<sup>R</sup>, gentamicin resistance; Em<sup>R</sup>, erythromycin resistance.

<sup>†</sup>The ampicillin resistance gene is no longer functional in the pMDM series <i>lux</i> reporter.
Efforts to improve the whole-cell assays. Efforts were made to develop high-throughput whole-cell luminescence assays. First, cells were grown on TYG plates followed by saturation of the plate with 10% sonicated decanal. Luminescence output was observed visually and by exposure of X-ray film. Second, isolated colonies were scraped off TYG plates and resuspended in a luminometer cuvette in 50 μl luciferase assay buffer (10 mM KPO₄, 1 mM DTT, 0.1 g BSA per 100 ml). A 1 μl volume of 10% sonicated decanal was added to the cell suspension and light output was monitored in a luminometer with an integration time of 4 s. Finally, 96-well plate assays were performed with 200 μl TYG agar plugs or broth cultures. Two or five microlitres of 10% sonicated decanal were added to broth and agar cultures, respectively. Light output was evaluated using a SpectraFluor Plus plate reader (Tecan), with a gain of either 100 or 125 and an integration time of 500 ms.

RESULTS

Development of the lux gene reporter system

Preliminary work was done to evaluate the feasibility of using a lux-based reporter system in Bacteroides. The pMDM series reporters (Fig. 2a) were developed with the entire lux operon, encoding both the luciferase and the fatty acid reductase complex, and were used to measure the production of light using both whole-cell luminescence assays and crude cell extract luciferase assays. Expression of the reporter was examined using the cepA promoter (data not shown) or the BT4001 16S rRNA promoter (Fig. 3). BT4001 was able to produce a functional luciferase enzyme and produce light in the presence of oxygen using a lux operon reporter system under the control of either of the two Bacteroides promoters tested, but only when exogenous decanal substrate was added. The output of the negative control pMDM4 was negligible with respect to the background production of light in these two assay systems (Fig. 3). However, there are concerns when using the pMDM series vectors. First, promoters are not easily cloned, because each promoter has to be cloned into pJT205 before the constructs can be fused with pVAL1. Second, the vectors are large, ~17 kb in size. Third, strains carrying these large vectors have a roughly two- to threefold increase in generation time. Since the decanal substrate had to be supplied exogenously to the cells, the luxCDE genes were not necessary for the reporters. Therefore, the pMMT series vector system was developed as an improved reporter for Bacteroides because a shorter one-step cloning method could be used to introduce promoters into the vector using the unique BamHI, KpnI, PstI and SmaI sites. Besides easier manipulation of the vector, it is also smaller (~10 kb in size) and has less of an effect on the growth of strains. In order to evaluate the sensitivity of the new pMMT lux reporter (Fig. 2b), prior to experiments with the as yet uncharacterized rRNA promoter, expression levels were measured from fusions to the well-characterized wild-type cepA promoter and two variants, which have decreased activity (0 and 33% activity), as demonstrated by the chloramphenicol acetyltransferase (CAT) gene reporter system (Bayley et al., 2000; Table 1). The strains carrying pMMT1F (wild-type), pMMT1G (0%) and pMMT1H (33%) were used for both whole-cell luminescence and crude cell extract luciferase assays (data not shown), which yielded similar results to the published CAT assays (Bayley et al., 2000). Hence, the new reporter was deemed to be sensitive enough to be used to analyse expression from 16S rRNA constructs.

Heterologous promoter expression

Luciferase assays performed on cell extracts of E. coli demonstrated the activity of the E. coli rnbP1 promoter,
pMMT1A and pMMT2A, which expressed the lux reporter at 250–160-fold higher levels than the negative controls (Fig. 4a, b). *rrnB P1* allowed expression of the lux reporter in the presence of either the *Bacteroides* or the *E. coli* RBS, but expression was slightly enhanced in the presence of the *Bacteroides* RBS. Differences in the spacing between the RBS and start codon being recognized in *E. coli* might explain the enhanced levels of expression observed from the pMMT1A construct (seven bases) when compared with pMMT2A (eight bases) (Fig. 4a, b). Luciferase assays on crude cell extract from *E. coli* also showed that, despite the presence of the *E. coli* RBS, the BT4001 full-length 16S rRNA promoter (pMMT2B) and 16S rRNA P1 (pMMT2C) were able to permit expression of the lux reporter at levels significantly higher than the negative control, but about an order of magnitude lower than the *E. coli* P1 promoter fusions (Fig. 4b). This demonstrates that differences in gene expression between *E. coli* and *Bacteroides* species occur at the level of promoter recognition.

Luciferase assays performed on cell extracts of BT4001 demonstrated the strength of the BT4001 16S rRNA promoter in that in the presence of its own RBS (pMMT1B), it permitted expression of the lux reporter at ~260-fold higher levels than the promoterless control pMMT1 (Fig. 4c). Likewise, *B. thetaiotaomicron* 16S rRNA P1 (pMMT1C) permitted expression of the reporter at 230-fold higher levels than the negative control (Fig. 4c). *E. coli* *rrnB P1* was unable to drive expression of the lux reporter in BT4001, even in the presence of the *Bacteroides* RBS.

**Ribosomal binding site**

![Diagram](Image)

**Fig. 4.** Luciferase assays using both *E. coli* and *B. thetaiotaomicron* whole-cell lysates. (a, b) Transcription levels from *E. coli* and BT4001 16S rRNA promoters using *E. coli* whole-cell lysates with the *Bacteroides* RBS (a) and the *E. coli* RBS (b). (c, d) Transcription initiation from *E. coli* and BT4001 16S rRNA promoters using BT4001 whole-cell lysates with the *Bacteroides* RBS (c) and the *E. coli* RBS (d). From left to right, the graphs show: pMMT1 or pMMT2, promoterless control with the *Bacteroides* RBS or *E. coli* RBS, respectively; pMMT1A or pMMT2A, *E. coli* *rrnB P1*; pMMT1B or pMMT2B, BT4001 full-length 16S rRNA promoter; pMMT1C or pMMT2C, BT4001 16S rRNA P1. Results shown are the mean (±SD) of two independent trials, performed in triplicate. The mean RLU values for the promoterless controls, for (a) and (b), were 0.71 and 0.13, respectively. The mean RLU values for the promoterless controls, for (c) and (d), were 0.062 and 0.081, respectively. Asterisks show statistically significant differences as compared with the promoterless controls, determined using Student’s unpaired *t* test (two-tailed); *P* < 0.001.
(pMMT1A; Fig. 4c). The *Bacteroides* promoters in combination with the *E. coli* RBS did permit initiation of luciferase expression at levels significantly higher than those for the negative control. In the presence of the *E. coli* RBS, BT4001 16S rRNA P1 allowed expression of the *lux* reporter at 20-fold higher levels than the promoterless control (pMMT2C; Fig. 4d). However, when compared with the strength of the BT4001 16S rRNA P1 promoter in the presence of the *Bacteroides* RBS (pMMT1C; Fig. 4c), this level of increase was about an order of magnitude lower. These results suggest that in *Bacteroides* species, translation initiation also plays a role in the disparity of gene expression between the two species.

**Analysis of a *B. thetaiotaomicron* 16S rRNA promoter region**

The *lux* reporter also allowed us to analyse the different regions of the BT4001 16S rRNA promoter to evaluate the role they play in initiating transcription. The BT4001 full-length 16S rRNA promoter, P1, P2 and full-length promoter with no upstream sequences of P1 (pMMT1B, pMMT1C, pMMT1D and pMMT1E, respectively) were also examined (Fig. 5). *In vitro* luciferase assays demonstrated that the regions believed to contain BT4001 16S rRNA P1 and P2 contained a functional promoter, and that each was capable of permitting initiation of gene expression in the absence of the other (pMMT1C and pMMT1D, respectively; Fig. 5). This suggests that, as in *E. coli*, the two promoters work together in order to maintain the appropriate amount of rRNA in a cell at a given growth condition. Also, BT4001 full-length 16S rRNA promoter without any upstream sequences of P1 was capable of being expressed at levels close to that of the same promoter with upstream sequences (pMMT1E versus pMMT1B, respectively; Fig. 5).

Primer extension was done in order to examine the start site of transcription in *Bacteroides*. The results of these experiments showed that the transcription start site from the P1 promoter is a cytosine (Fig. 6b). The dominant start site of transcription for the P2 promoter is guanine with a secondary adjacent start site at a cytosine (Fig. 6a). These results indicate that the spacing between the centre of the −7 region and the transcription start site is 8 bp.

**Examination of the putative P1 −33, −7 and upstream sequences**

Substitutions were made in the *Bacteroides* P1 promoter to examine their effect on promoter strength. *In vitro* luciferase assays of crude cell extracts from cells that contained P1 (pMMT1C) detected ~550-fold higher RLU than the promoterless control (pMMT1; data not shown). The activity levels of the other promoters as compared with the wild-type were: 11.2 % for the −33 substituted (TTTG→AAAC), 4.0 % for P1 with no upstream sequence, 1.8 % for the −7 upstream substituted region (TA→AT), and 0.44 % for the −7 downstream substituted region (TTTG→AAAC) (pMMT1K, pMMT1L, pMMT1J and pMMT1I, respectively; data not shown). These results indicate that the −7 upstream and downstream sequences are equally important to the function of this promoter, more so than the −33 region, and that the region upstream of P1 influences promoter strength.

**Efforts to improve whole-cell luminescence assays**

Colonies containing pMDM1 on TYG plates flooded with decanal and exposed to X-ray film showed a very faint light emittance that could be seen by the naked eye, but lasted for less than a minute. Unfortunately, images on X-ray film were too diffuse to visualize individual colonies. When examining colonies resuspended in assay buffer, some light output above background was observed for up to 30 min from cells containing pMDM1 and for up to 60 min for cells containing pMMT1 and pMMT1H. These time limitations were presumably due to the decreased energy/
Reduced coenzyme levels of the cells upon exposure to oxygen.

The amount of light output from the broth cultures grown in 96-well plates containing strains with pMMT1F and pMMT1H was 2–15 times above the background levels for ~30 min. The background levels from the negative control cultures were 2–5 RLU. However, with broth cultures containing pMDM1, the maximum light output was only four to six times above background for 5 min only, and the light output from these strains was very weak. When the cultures containing pMMT1F and pMMT1H were grown in the agar plugs, the RLU from these strains were above background for only about 10 min. Therefore, high-throughput qualitative measurements to score for the presence of light output above background are possible for a short amount of time following the addition of the substrate. The best results screening for light output were achieved using the pMMT series vectors grown anaerobically in broth cultures on microtitre plates. However, downstream in vitro luciferase assays are necessary to accurately quantify promoter strength.

**DISCUSSION**

The barriers that impede successful expression of foreign bacterial genes in *Bacteroides* are still not fully understood. To date, most of the studies in the area of gene expression of *Bacteroides* have focused on a limited number of metabolic reporter systems. The four reporter systems that have previously been used in *Bacteroides* make use of β-glucuronidase (GUS) (Feldhaus et al., 1991), xylanase/arabinosidase (Whitehead, 1997), CAT (Bayley et al., 2000; Smith et al., 1992) and catechol 2,3-dioxygenase (XylE) (Chatzidaki-Livanis et al., 2008; Krinos et al., 2001). Endogenous enzyme production can be an issue with the first two systems, the substrate for the CAT assay can be cost-prohibitive, and only the XylE system is amenable to high-throughput differential screening. Thus, there is a need for the development of additional reporters with low background levels and rapid screening capabilities.

The new approach that we have used here involves the luminescence or *lux* operon as a tool for evaluating levels of gene expression in *Bacteroides*. To our knowledge, there is only one other study that has focused on using bioluminescence as a reporter system for an anaerobic bacterium. This study was performed on another human pathogen, *Clostridium perfringens*, using the *lux* operon from the marine bacterium *Vibrio fisheri* (Phillips-Jones, 1993). The problem with using the *V. fisheri* lux operon for studies in *Bacteroides* is that the Lux proteins are not fully active at temperatures above 30 °C, while *Bacteroides* has an optimal growth temperature of 37 °C. To overcome this problem, the Lux system from *P. luminescens*, which functions at up to 45 °C (Szittner & Meighen, 1990), was used for this study. Light output may be measured quantitatively over several orders of magnitude, making it an ideal reporter, but only under aerobic conditions.

Although *Bacteroides* are anaerobes, they are aerotolerant, capable of surviving in an oxygenated environment for a period of time, but are incapable of maintaining the activities of key enzymes necessary for energy production via central metabolism (Pan & Imlay, 2001). Therefore, whole-cell luminescence assays (Fig. 3a) were found to be feasible on a limited time frame, where sufficient energy/reduced coenzymes were available to produce the substrates FMNH₂ and decanal necessary for luciferase activity. This problem was eliminated by adding the substrates exogenously during *in vitro* luciferase assays performed on crude cell extracts as described above. These reporters are useful tools for examining promoter activity and gene expression in BT4001 and could be applied to other species of *Bacteroides*, *Cytophaga* and *Flavobacterium*.
The *Bacteroides* 16S rRNA promoter region examined here has the same overall rRNA operon promoter structure as in *E. coli*. There are two BT4001 16S rRNA promoters, both containing the −7 and −33 *Bacteroides* consensus sequences, −150 bp apart. Overall, there are five rRNA operons in BT4001 (Xu et al., 2003), four of which are 94–98 % identical in sequence. In these four, both P1 and P2 promoters are 100 % conserved. The fifth rRNA operon is 88 % identical to this sequence and only contains putative P2.

The results from the luciferase assays on cell extracts showed that BT4001 16S rRNA P1 and P2 promoters were both capable of initiating gene expression in the absence of the other and that the full-length 16S rRNA promoter was able to initiate gene expression in the absence of any sequences upstream of the P1 promoter (Fig. 4). Estrem et al. (1998) identified a consensus UP element sequence in *E. coli* by using *in vitro* selection for upstream sequences that promote rapid RNA polymerase binding to the *E. coli* rrrB P1 promoter, followed by *in vivo* screening for high promoter activity. UP element effects on promoter activity differ, correlating generally with the degree of similarity to the consensus sequence (Ross et al., 1998). A + T-rich upstream sequences have also been identified in compilations of *Bacillus subtilis* (Helmann, 1995) and *Clostridium* promoters (Graves & Rabinowitz, 1986). Even though there is no sequence upstream from BT4001 16S rRNA P1 and P2 similar to the *E. coli* consensus UP element sequence, there appear to be A + T-rich regions. It has been demonstrated that sequences out to −54 of certain regulated *B. fragilis* promoters are required for optimal activity (Bayley et al., 2000), but this may reflect a requirement for activator binding. The examination of the region upstream of the BT4001 P1 promoter indicates that it is directly involved in its activity (data not shown), but this region does not seem to affect the activity of the full-length 16S rRNA promoter (Fig. 5). Similar results have been observed in another member of the phylum *Bacteroidetes*, *Flavobacterium*. Deletion of the sequences upstream of the −33 region between −50 and −42 of the *Flavobacterium johnsoniae* ompA promoter leads to a dramatic reduction in promoter activity (Chen et al., 2007a). BT4001 has an A-rich region, 7 bp in length, which is 12 bp upstream of the −33 region of both P1 and P2. *F. johnsoniae*, however, has a T-rich region in the same area upstream of the *ompA* promoter.

The Fis protein can bind DNA in a non-specific manner as well as at specific sites, making determination of a clear consensus sequence difficult (Bettermier et al., 1994). Fis-binding sites could be present upstream of the BT4001 rRNA operon, but due to a weak consensus, further studies will need to be done to determine if they are present. However, sequence analysis of the BT4001 genome sequence (Xu et al., 2003) reveals no obvious Fis homologues.

A comparison of the P1 16S rRNA promoter variants with variants of the *B. fragilis* cepA promoter in the study by Bayley et al. (2000) confirmed the role of the −7 and −35 consensus sequences. With respect to the −7 downstream region (TTTG), both the cepA (Bayley et al., 2000) and the P1 substitutions show a lack of expression when this region is changed (Fig. 6). The P1 promoter with no upstream sequence and −33 substituted sequence helps confirm that the −33 region and the bases upstream of it are both important to the overall expression of the promoter. However, it seems that the upstream sequence has a larger effect on the activity of the promoter than the −33 region. These results also agree with earlier studies using the cepA promoter (Bayley et al., 2000). One discrepancy was seen in comparing the results from the P1 −7 upstream region (TA) with those for the cepA promoter. We observed that the substitution of these two bases resulted in 1.8 % activity as compared with the wild-type, unlike the cepA promoter, where a 38 % activity was observed (Bayley et al., 2000).

Perhaps the −7 upstream region (TA) is critical to the activity of BT4001 P1, but not to other promoters examined thus far, since those examined by Bayley et al. (2000) do not all share the consensus −7 upstream region. Our findings are similar to those of studies of *Flavobacterium*, where deletion of the −33 region of the ompA promoter (Chen et al., 2007b) or changes in either of the −7 upstream bases (TA) almost completely abolish gene expression (Chen et al., 2007a). Indeed, most substitutions at bases in both the −33 and the −7 region of the *Flavobacterium* promoter lead to decreased GFP reporter expression in comparison with the wild-type promoter levels (Chen et al., 2007a).

Overall, the results of this study expand on the work of others (Bayley et al., 2000; Vingadassalom et al., 2005, 2007) to provide additional insights into the requirements for transcription initiation in *Bacteroides*. Characterization of the P1 and P2 promoters and P1 promoter variants confirms the importance of the −33 and −7 sites for promoter activity *in vivo*. In *E. coli*, promoter recognition is the limiting factor in expressing *Bacteroides* genes (Fig. 4a, b). However, in *Bacteroides*, it seems that the *E. coli* promoters and RBS are not efficiently utilized by the *Bacteroides* transcription and translation machineries (Fig. 4c, d). Furthermore, it appears that the RBS sequence and spacing recognition by the ribosome in *Bacteroides* are more selective than those in *E. coli*.

**ACKNOWLEDGEMENTS**

We thank Jennifer Meador-Parton for her technical assistance, Reinhardt Rosson (Bio-Technical Resources, Manitowoc, WI) for providing p7T205, and Richard Gourse, Timothy Larson, Abigail Salyers, Nadja Shoemaker and C. Jeff Smith for the materials that they provided for this study. This work was supported by NSF CAREER award MCB-9875479, the Virginia Tech Optical Sciences and Engineering Research (OSER) Center and Carilion Biomedical Institute.

**REFERENCES**


Edited by: H. J. Flint