JOURNAL OF BACTERIOLOGY, July 1995, p. 3721–3727 Vol. 177, No. 13
0021-9193/95/$04.00+0

Identification and Characterization of a *Bacteroides* Gene, *csuF*, Which Encodes an Outer Membrane Protein That Is Essential for Growth on Chondroitin Sulfate

QIONG CHENG, MICHAEL C. YU, ANDREW R. REEVES, AND ABIGAIL A. SALYERS*

Department of Microbiology, University of Illinois, Urbana, Illinois 61801

Received 25 August 1994/Accepted 27 April 1995

Bacteroides thetaiotaomicron **can utilize a variety of polysaccharides, including charged mucopolysaccharides such as chondroitin sulfate (CS) and hyaluronic acid (HA). Since the enzymes (chondroitin lyases I and II) that catalyze the first step in breakdown of CS and HA are located in the periplasm, we had proposed that the first step in utilization of these polysaccharides was binding to one or more outer membrane proteins followed by translocation into the periplasm, but no such outer membrane proteins had been shown to play a role in CS or HA utilization. Previously we had isolated a transposon-generated mutant, CS4, which was unable to grow on CS or HA but retained the ability to grow on disaccharide components of CS. This phenotype suggested that the mutation in CS4 either blocked the transport of the mucopolysaccharides into the periplasmic space or blocked the depolymerization of the mucopolysaccharides into disaccharides. We have mapped the CS4 mutation to a single gene, *csuF*, which is capable of encoding a protein of 1,065 amino acids and contains a consensus signal sequence. Although CsuF had a predicted molecular weight and pI similar to those of chondroitin lyases, it did not show significant sequence similarity to the *Bacteroides* chondroitin lyase II, a *Proteus* chondroitin ABC lyase, or two hyaluronidases from *Clostridium perfringens* and *Streptococcus pyogenes*, nor was any CS-degrading enzyme activity associated with *csuF* expression in *Bacteroides* species or *Escherichia coli*. The deduced amino acid sequence of CsuF exhibited features suggestive of an outer membrane protein. We obtained antibodies to CsuF and demonstrated that the protein is located in the outer membrane. This is the first evidence that a nonenzymatic outer membrane protein is essential for utilization of CS and HA.**

Bacteroides is a genus of gram-negative obligate anaerobes that accounts for about 25% of the bacteria in the human colonic microflora (23). Some *Bacteroides* spp. are able to utilize a variety of polysaccharides (35), including host-derived as well as dietary polysaccharides, and this capability may contribute to their survival in the colon (27). We have been investigating the mechanism of breakdown of two negatively charged host mucopolysaccharides, chondroitin sulfate (CS) and hyaluronic acid (HA), by *Bacteroides thetaiotaomicron*. CS is composed of repeating dimers of *N*-acetylgalactosamine and glucuronic acid, with a sulfate group on the C-4 or C-6 position of the *N*-acetylgalactosamine residue, and has a molecular weight of at least 10,000. HA is similar in structure to CS except that it contains *N*-acetylglucosamine instead of *N*-acetylgalactosamine and has no sulfate groups. The average molecular weight of HA is greater than 100,000. Chondroitin lyases I and II, the enzymes which degrade CS and HA to disaccharides, are located in the periplasmic space (28). The enzymes that desulfate the sulfated disaccharides (chondro-4-sulfatase and chondro-6-sulfatase) and the β-glucuronidase that cleaves the disaccharide into monosaccharides are located in the cytoplasm (28).

Since the degradative enzymes are not exposed on the cell surface and since the large sizes of CS and HA make it unlikely that these polymers could diffuse through pores in the outer membrane, we suggested that CS and HA were first bound to a protein receptor on the outer membrane and then translocated into the periplasm, where they would be brought into contact with the chondroitin lyases. The enzymes involved in breakdown of CS and HA are produced only when *B. thetaiotaomicron* is grown on one of these polysaccharides or on their disaccharide constituents (26). Consistent with our hypothesis that an outer membrane receptor is involved in CS utilization, we had found previously that several outer membrane proteins appeared to be coregulated with the degradative enzymes (17); i.e., they were produced only when bacteria were grown on CS, HA, or disaccharide components. There was no evidence, however, that any of these proteins was essential for CS or HA utilization. Also, since even the most highly purified preparations of commercially available CS and HA contain small amounts of contaminating proteins, polysaccharides, and ions, we could not rule out the possibility that some of these outer membrane proteins were induced in response to other molecules unrelated to CS and HA.

If outer membrane proteins of *B. thetaiotaomicron* are essential for utilization of CS and HA, it should be possible to find mutants lacking these proteins that are unable to grow on CS and HA. These mutants should still be able to grow on the disaccharides of CS, which are small enough to diffuse through outer membrane porins. We had found previously that wild-type *B. thetaiotaomicron* could grow on the sulfated disaccharides of CS but grew much more slowly on these disaccharides than on intact CS or HA. Presumably this slow growth was due to the fact that the disaccharides had to cross the outer membrane by diffusion through porins rather than by interacting with the hypothetical polysaccharide translocation complex. In this paper, we report the identification and characterization of a gene that encodes an outer membrane protein, which is essential for growth on CS and HA.

* Corresponding author. Mailing address: Department of Microbiology, 131 Burrill Hall, 407 S. Goodwin Ave., University of Illinois, Urbana, IL 61801. Phone: (217) 333-7378. Fax: (217) 244-6697.

TABLE 1. Bacterial strains and plasmids used[a]

Strain or plasmid	Relevant characteristic(s)	Reference, source, or description
Strains		
E. coli		
DH5αMCR	$RecA^-$ Gn^s	13
S17-1	$RecA^-$ (Tp^r) (Sm^r) Gn^s, RP4 inserted in chromosome	33
B. thetaiotaomicron		
5482	Wild type, Gn^r	Anaerobe Laboratory, Virginia Polytechnic Institute, Blacksburg
BT4001	Rif^r Gn^r Em^s	Spontaneous Rif^r mutant of *B. thetaiotaomicron* 5482
ΔCSaseII	Rif^r Gn^r Em^s CSaseII$^-$	Chromosomal deletion in the CSaseII gene in BT4001 (25)
CS4	Rif^r Gn^r Em^r CSaseII$^-$ CS^- HA^- HP^+	Tn*4351* mutagenesis of ΔCSaseII (2)
Plasmids		
pCQW1	(Ap^r) Em^r	GUS fusion suicide vector used to make chromosomal GUS fusions (6)
R751	(Tp^r Tra^+)	IncP plasmid used to mobilize rectors from *E. coli* to *Bacteroides* recipients (32)
pNJR12	(Kn^r) Tc^r	RSF1010-based shuttle vector used for complementation experiments (2)
pCS4A	(Kn^r Tc^r)	pJRD215 (5) containing the 3.4-kbp *tetX* portion of Tn*4351* and a 2.2-kbp flanking chromosomal fragment from CS4 (this study)
pHR4	(Ap^r)	pBR328 containing a 4.1-kbp *Hin*dIII-*Eco*RI chromosomal fragment from BT4001 which spanned the transposon insertion site in CS4 (this study)
pNCS4	(Kn^r) Tc^r	pNJR12 containing the 4.1-kbp *Hin*dIII-*Eco*RI fragment from pHR4 (this study)
pNCS5	(Kn^r) Tc^r	pNJR12 containing the 2.2-kbp *Eco*RV fragment from pHR4 (this study)
pNCS6	(Kn^r) Tc^r	pNJR12 containing the 2.0-kbp *Ssp*I-*Pvu*II fragment from pHR4 (this study)
pNCS7	(Kn^r) Tc^r	pNJR12 containing the 3.3-kbp *Nde*I-*Eco*RI fragment from pHR4 (this study)
pCQW15	(Ap^r) Em^r	pCQW1 containing the 1.1-kbp *Eco*RV fragment from pHR4 fused to the GUS gene in the orientation that the *Pvu*II site on this fragment is closer to the GUS gene (this study)
pCQW16	(Ap^r) Em^r	pCQW1 containing the same 1.1-kbp *Eco*RV fragment fused to the GUS gene in the opposite orientation as in pCQW15 (this study)
pCQW17	(Ap^r) Em^r	pCQW1 containing the 0.5-kbp *Pvu*II fragment from pHR4 fragment fused to the GUS gene (this study)
pCQW18	(Ap^r) Em^r	pCQW1 containing the same 0.5-kbp *Pvu*II fragment fused to the GUS gene in the opposite orientation as in pCQW17 (this study)

[a] Abbreviations: Gn, gentamicin; Tp, trimethoprim; Sm, streptomycin; Rif, rifampin; Em, erythromycin; Ap, ampicillin; Tc, tetracycline; Kn, kanamycin; GUS, β-glucuronidase encoded by *E. coli* gene *uidA*; CSaseII, chondroitin lyase II; CS, chondroitin sulfate; HP, heparin; HA, hyaluronic acid. Genes in parentheses are expressed only in *E. coli*; genes not in parentheses are expressed in *B. thetaiotaomicron*.

MATERIALS AND METHODS

Bacterial strains and plasmids. The bacterial strains and plasmids used in this study are listed in Table 1. The isolation of mutant CS4 has been described previously (2). *B. thetaiotaomicron* strains were grown either in prereduced Trypticase-yeast extract-glucose medium or in a defined medium (15) that contained one of the following as the sole carbon source: CS A (5 mg/ml), HA (3 mg/ml), heparin (4 mg/ml), or glucose (5 mg/ml). *Escherichia coli* strains were grown in Luria-Bertani medium. Concentrations of antibiotics used for selection were as follows: erythromycin, 10 μg/ml; gentamicin sulfate, 200 μg/ml; trimethoprim, 200 μg/ml; tetracycline, 10 μg/ml; ampicillin, 100 μg/ml; and kanamycin monosulfate, 100 μg/ml.

DNA manipulation, transformation, and conjugation. Isolation of plasmids and chromosomal DNA, gel electrophoresis, and Southern hybridization were done as described by Maniatis et al. (22). Transformation of *E. coli* with plasmid DNA and conjugation with *Bacteroides* recipients were performed as described previously (32). pCQW1 derivatives were transformed into *E. coli* S17-1 (33) and were then mobilized from S17-1 into *B. thetaiotaomicron* recipients by the Tra^+ plasmid RP4 inserted in the S17-1 chromosome (10). Other plasmids and their derivatives were introduced into *E. coli* DH5αMCR and were then mobilized from *E. coli* to *Bacteroides* species by the broad-host-range IncP plasmid R751 (32).

Cloning of the wild-type chromosomal DNA segment that contained the transposon insertion site in CS4. Southern hybridization experiments using plasmid

R751 and Tn*4351* as the hybridization probes indicated that CS4 contained a single Tn*4351* insertion and that R751 had not formed a cointegrant (31) with Tn*4351* (data not shown). The junction chromosomal DNA from the insertion site in CS4 was cloned as previously described (2) by taking advantage of a *tetX* gene carried on Tn*4351* which is expressed in *E. coli* (9). The resulting plasmid, pCS4A, contained a 5.6-kbp fragment that included a 3.4-kbp part of Tn*4351* and a 2.2-kbp adjacent chromosomal DNA. pCS4A was used as a hybridization probe to obtain the wild-type locus. ^{32}P-pCS4A was first hybridized to restriction endonuclease-digested chromosomal DNA from wild-type *B. thetaiotaomicron*. A 4.1-kbp *Hin*dIII-*Eco*RI fragment cross-hybridized with the probe. The region containing the 4.1-kbp segments from the *Hin*dIII-*Eco*RI double digests was then extracted from a low-melting-point agarose gel and ligated to the *Hin*dIII-*Eco*RI-digested pBR328. The resulting plasmids were transformed into *E. coli* DH5αMCR, with selection for ampicillin-resistant transformants. The transformants were screened by colony hybridization. Six of the 200 colonies screened cross-reacted with the probe. Plasmids from these isolates were digested with *Hin*dIII plus *Eco*RI. All of them had the same restriction pattern with a 4.1-kbp *Hin*dIII-*Eco*RI fragment that hybridized to the probe on Southern blot. One was designated pHR4. To determine the smallest fragment that would complement the mutation in CS4, the entire 4.1-kbp *Hin*dIII-*Eco*RI fragment or subclones of it were cloned on a low-copy-number shuttle vector pNJR12 and introduced into CS4, with selection for tetracycline resistance. Eighty transconjugants from each mating were patched on minimal medium agar plates to test for the ability to grow on CS, HA, or glucose.

DNA sequence analysis. Nested deletion subclones were generated by using exonuclease III in the Erase-a-Base kit (Promega, Madison, Wis.). Overlapping clones of the progressive unidirectional deletions were sequenced by the M13 dideoxynucleotide chain termination method (29). Sequencing reactions were performed by using a Sequenase 2.0 kit purchased from United States Biochemicals, Inc. (Cleveland, Ohio). The sequencing reaction products were resolved on 6% polyacrylamide–8 M urea denaturing gels. Gels were run at 40 mA (50 to 55°C), fixed, dried, and exposed to X-ray films.

Construction of GUS fusions and targeted insertional disruptions. The chromosomal β-glucuronidase (GUS) fusions were constructed by cloning DNA segments into multiple cloning sites upstream of the GUS gene on the suicide GUS fusion vector pCQW1 (6) and integrating the resulting plasmids into the *B. thetaiotaomicron* chromosome. Four fusions were constructed in this study. ΩCQW15 and ΩCQW16 were constructed by integration of plasmids pCQW15 and pCQW16, which contained a 1.1-kbp *Eco*RV fragment cloned into the *Sma*I site of pCQW1 in both orientations. ΩCQW17 and ΩCQW18 were constructed by integration of plasmids pCQW17 and pCQW18, which contained a 450-bp *Pvu*II fragment cloned into the *Sma*I site of pCQW1 in both orientations. The chromosomal GUS fusions were assayed for GUS activity (6) as well as for the ability to grow on CS, HA, and glucose. Two of the Erase-a-Base clones generated for sequencing were also used for making targeted insertional disruptions. 19Δ11 and 20Δ28 were generated by nested deletions of a 2.2-kbp *Eco*RV fragment cloned into suicide vector pCQW1 in different orientations. Deletion clone 19Δ11 contained an insert of 750 bp. Deletion clone 20Δ28 contained an insert of 1.1 kbp. Both clones were transformed into *E. coli* S17-1 and mobilized from S17-1 into *Bacteroides* recipients, with selection for erythromycin resistance. Transconjugants were tested for the ability to grow on CS, HA, and glucose.

Assay of enzyme activities. Cell extracts were prepared as described previously (28). Protein concentrations in the cell extracts were determined by the method of Lowry et al. (20). GUS activity was determined by measuring the increase in A_{415} when the extracts were incubated with *p*-nitrophenyl glucuronide (Sigma Chemical Co., St. Louis, Mo.) at 37°C (6). Chondroitin lyase activity was assayed by incubating the extracts with CS A (2 mg/ml) and measuring the increase in A_{235} due to the production of unsaturated disaccharides (28). Chondroitin lyase and hyaluronidase activities were also detected by incubating cell extracts with CS or HA and separating the end products by paper chromatography (7). Chondro-4-sulfatase, chondro-6-sulfatase, and β-glucuronidase activities were measured as described previously (28). The substrates were ΔDi-4S (chondro-4-sulfatase assay), ΔDi-6S (chondro-6-sulfatase assay), or ΔDi-OS (β-glucuronidase assay); they were purchased from ICN Biomedical, Inc. (Irvine, Calif.).

Localization of CsuF. Antiserum against CsuF was obtained by immunizing mice with a synthetic peptide linked to keyhole limpet hemacyanin (Sigma). The peptide consisted of amino acids 466 through 491: 5′CKLKVSAGYTYDSRK QDQFNNSNTRY3′. The peptide was synthesized, covalently linked to hemacyanin, and injected into a mouse to generate polyclonal antiserum. These manipulations were done by the University of Illinois Biotechnology Center, using standard procedures. Cytoplasmic and outer membranes were obtained from *B. thetaitoaomicron* grown on CS, as described previously (17). Total membranes and soluble protein fractions were obtained from *B. thetaiotaomicron* grown on glucose or CS as described previously (28). Western blot (immunoblot) analysis of the membrane and soluble fractions was done with the anti-CsuF antiserum as described by Kotarski et al. (17) except that biotinylated immunoglobulin G goat anti-mouse antibody and streptavidin–β-galactosidase conjugate (Bethesda Research Laboratories, Bethesda, Md.) were used to detect bound antibody.

Nucleotide sequence accession number. The nucleotide sequences of the CS4 locus and part of the chondroitin lyase II gene reported in this paper have been deposited in GenBank under accession number L42367.

RESULTS AND DISCUSSION

Characteristics of mutant CS4. *B. thetaiotaomicron* CS4 was isolated originally by screening transposon-generated mutants of a derivative of *B. thetaiotaomicron* 5482 that had a deletion in the *csuB* gene, which encodes chondroitin lyase II (25). The rationale for mutagenizing the deletion mutant rather than the wild type was that we hoped to obtain transposon insertions in the gene encoding chondroitin lyase I, as well as other insertions that rendered *B. thetaiotaomicron* unable to grow on CS. Previously, we had shown that although disrupting the *csuB* gene reduced the chondroitin lyase activity in cell extracts by 70%, it did not affect the ability of *B. thetaiotaomicron* 5482 to grow on CS (11). This might also be true of the gene for chondroitin lyase I, but loss of both genes would be expected to eliminate growth on CS and HA. CS4 was the only mutant found in our screen (2) that was unable to grow on CS or HA but was still able to grow on sulfated disaccharides. This is the phenotype expected either for a mutation that disrupted an essential outer membrane protein or for a mutation that disrupted chondroitin lyase I. Enzyme assays done on cell extracts from CS4 grown on CS disaccharides (data not shown) indicated that chondro-4-sulfatase, chondro-6-sulfatase, and β-glucuronidase activities were still detected at the wild-type level, and chondroitin lyase activity was the same as that of the parent chondroitin lyase II deletion strain (0.4 U/mg of total cell protein). Thus, the transposon insertion in CS4 had not disrupted the chondroitin lyase I, although we could not rule out the possibility that the transposon insertion had disrupted a gene encoding another low-activity lyase or a hydrolyase which was not detected by our assay system. Moreover, the fact that enzymes involved in breakdown of CS were still produced and were still regulated indicated that the transposon insertion had not disrupted a regulatory locus.

Cloning of the wild-type CS4 locus. We cloned a 4.1-kbp *Hin*dIII-*Eco*RI fragment that contained the region of the wild-type chromosome in which the Tn*4351* insertion in CS4 had occurred (pHR4). Restriction digest analysis showed that the Tn*4351* insertion in CS4 had occurred about 2.2 kbp from the *Hin*dIII site. The CS4 strain containing this 4.1-kbp fragment in *trans* was able to grow on CS or HA, and its growth rate was similar to that of the wild type. Several subclones of the 4.1-kbp fragment were also tested, and none of these clones complemented the CS4 mutation. Results of this analysis suggested that at least 3.3 kbp of the cloned region was essential for growth on CS or HA. To determine if the locus disrupted by Tn*4351* in CS4 was essential in the wild-type background and to confirm the size of the region essential for growth on CS or HA, four insertional disruptions were made in the CS4 locus (Fig. 1). Two of the insertions abolished growth on CS and HA, indicating that the region was essential in the wild type. The fact that insertion of the 1.1-kbp *Eco*RV fragment did not abolish growth on CS or HA suggested that one end of the locus was within this fragment.

The 4.1-kbp DNA fragment containing the CS4 locus encodes a single ORF. The complete nucleotide sequence of the 4,114-bp *Hin*dIII-*Eco*RI fragment was determined. A single large open reading frame (ORF) was found (Fig. 1). This ORF had the capacity to encode 1,065 amino acids, the first possible ATG initiation codon of which was 881 bp from the *Hin*dIII site. A TGA termination codon occurred at bp 4078. The gene encoded by this ORF was designated *csuF* (for chondroitin

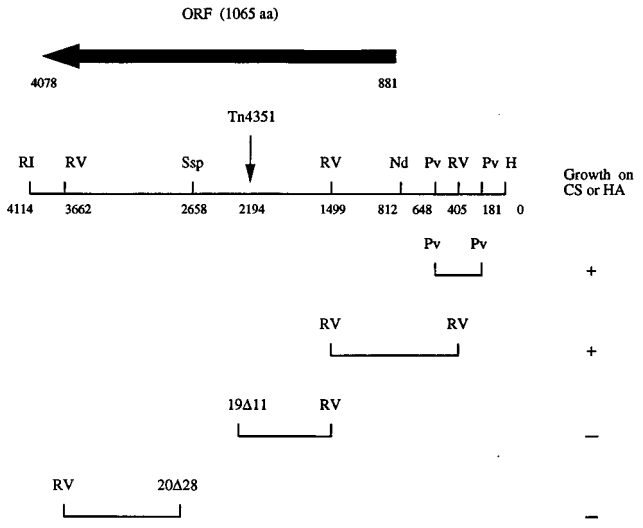

FIG. 1. Growth phenotype of the insertional mutations generated in wild-type *B. thetaiotaomicron* chromosome. The segments indicate the locations of the fragments that were cloned on a suicide vector and were used to make insertional mutations by homologous recombination into the *B. thetaiotaomicron* chromosome. 19Δ11 and 20Δ28 were generated by nested deletions of the 2.2-kbp *Eco*RV fragment in different orientations. Deletion clone 19Δ11 contained an insert of 750 bp and spanned the transposon insertion site. Deletion clone 20Δ28 contained an insert of 1.1 kbp and spanned the *Ssp*I site. The positions (base pairs) of the ORF, the transposon insertion site, and the relevant restriction sites are also given (relative to the right-hand *Hin*dIII site). Abbreviations: aa, amino acids; H, *Hin*dIII; RI, *Eco*RI; RV, *Eco*RV; Pv, *Pvu*II; Nd, *Nde*I; Ssp, *Ssp*I.

sulfate utilization). Genes encoding enzymes involved in CS breakdown had previously been designated *csuA* to *csuE*. By sequencing the junction between IS*4351* and chromosomal DNA in pCS4A, we found that the transposon in CS4 had inserted 2,194 bp from the *Hin*dIII site. There was no evidence of a second ORF immediately downstream of *csuF*. The fact that the 4.1-kbp segment was sufficient to complement the mutation in CS4 suggests that the inability of CS4 to grow on CS or HA was due to disruption of *csuF* itself, not to a polar effect of the transposon insertion on some downstream gene.

Regulation of *csuF* expression. To ascertain that *csuF* was expressed in *B. thetaiotaomicron* and to determine if its expression was regulated, we constructed four chromosomal GUS fusions (Fig. 2). Fusions ΩCQW15 and ΩCQW16 carried the same insertion with GUS fused in different directions. GUS activity was detectable in ΩCQW16 but not in ΩCQW15, indicating that the direction of *csuF* transcription was that predicted from the DNA sequence. There was no detectable GUS activity in the extract made of strain ΩCQW16 grown on glucose, but extracts from ΩCQW16 grown on CS or HA had a GUS specific activity of 2.5 U/mg of total cell protein. Thus, expression of *csuF* was regulated at the transcriptional level, as had been previously found for genes encoding chondroitin lyase II (*csuB*) and chondro-4-sulfatase (*csuC* [6]). However, expression of *csuF* was four- to fivefold lower than expression of these other CS-regulated genes.

So far, only one regulatory gene that controls the expression of genes involved in CS breakdown has been found: *chuR*. ChuR is a putative activator that controls the expression of a subset of genes essential for CS and heparin utilization (2). Only one of the genes encoding CS-degrading enzymes (*csuD* [chondro-6-sulfatase]) is controlled by ChuR (2). To determine if *csuF* was regulated by ChuR, an insertional disruption

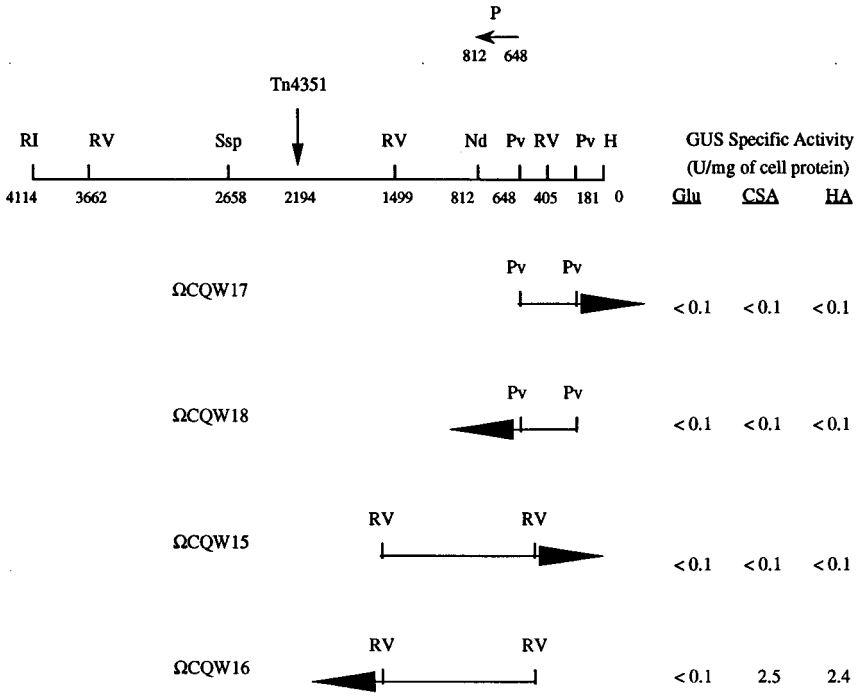

FIG. 2. Transcriptional GUS fusions and their GUS specific activities. The large arrowheads indicate the positions and directions of transcription of the GUS gene in each construct, and the contiguous line shows the region of DNA cloned upstream of GUS. The GUS specific activities were assayed in extracts from strains carrying the indicated chromosomal GUS fusions grown on glucose, CS, or HA. One unit of GUS activity is defined as an increase of 0.01 A_{415} unit/min at 37°C. The horizontal arrow above the restriction map at the top shows the direction of transcription and the region where the promoter was localized. Abbreviations: H, *Hin*dIII; RI, *Eco*RI; RV, *Eco*RV; Pv, *Pvu*II; Nd, *Nde*I; Ssp, *Ssp*I.

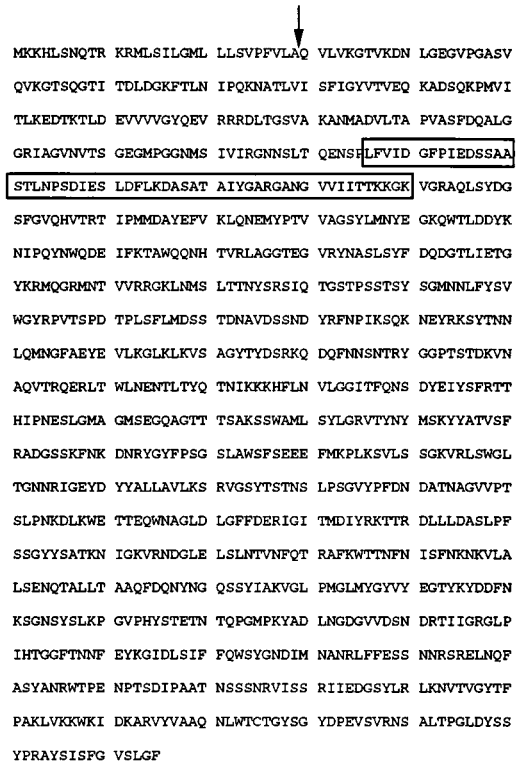

FIG. 3. Deduced amino acid sequence of the CsuF precursor. The CsuF precursor protein is predicted to consist of 1,065 amino acids with an estimated molecular weight of 118,269. Its putative signal peptide is predicted to be 29 residues long. The arrow indicates the putative cleavage site of the signal peptide. The boxed region indicates the location of a domain found in CsuF which is conserved in many known outer membrane proteins.

of *chuR* was generated in ΩCQW16 with a 600-bp internal fragment of *chuR* as previously described (2). Disruption of ChuR had no effect on *csuF* expression, as indicated by the GUS activity of extracts from ΩCQW16/*chuR* (data not shown). We conclude that *csuF* is not regulated by ChuR.

One of the subclones of the 4.1-kbp *Hin*dIII-*Eco*RI region that failed to complement the mutation (pNCS7) contained the entire ORF plus 70 bp of upstream DNA but still failed to complement the CS4 mutation. This finding suggested that the 70 bp of upstream DNA (bp 812 to 881) did not contain the functional promoter for *csuF* and that the promoter was located further upstream. Fusion ΩCQW16 had CS-regulated GUS activity, whereas fusion ΩCQW18 had no detectable GUS activity (Fig. 2). This finding suggested that the *Pvu*II fragment used to make ΩCQW18 was outside the transcriptional unit of *csuF*, and the promoter on the *Eco*RV fragment used to make ΩCQW16 should be downstream of the internal *Pvu*II site on the *Eco*RV fragment. Therefore, the promoter of *csuF* appeared to be located in the region between *Pvu*II and *Nde*I (bp 648 to 812; Fig. 2).

Comparison of *csuF* with known hyaluronidase and chondroitin lyase genes. The deduced amino acid sequence of CsuF is shown in Fig. 3. Some properties of CsuF predicted from the sequence data were similar to those of chondroitin lyases I and II, as determined from previous biochemical studies (19). *csuF* was predicted to encode a precursor protein with a deduced molecular weight of 118,269. The molecular masses chondroitin lyases I and II estimated from sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis analysis were 108 and 104 kDa, respectively. Chondroitin lyases I and II are periplasmic proteins. The deduced amino acid sequence of the N-terminal region of *csuF* displayed features of prokaryotic signal peptides such as a basic N-terminal end, a central hydrophobic core, and a polar C-terminal region (4). A consensus cleavage site based on the −3, −1 rule (4) was also found 29 residues from the N-terminal end of CsuF. The genus *Bacteroides* is phylogenetically distant from the *E. coli* group, and the *E. coli* consensus promoters do not work in *Bacteroides* strains. *csuF* is the first *Bacteroides* gene sequenced that appears to be secreted through the cytoplasmic membrane. Finding a consensus signal sequence in CsuF suggests that the *E. coli* group signal sequence motif, unlike the *E. coli* group promoter sequences, may be found in secreted proteins of the *Bacteroides* group. The deduced isoelectric point of CsuF was strikingly high (pI 9.8). Under physiological conditions (pH 7.0), the protein would be positively charged, as expected for a protein that bound a negatively charged substrate such as CS or HA. Chondroitin lyases I and II are also basic proteins, both with pI values of 7.9 to 8.0 on isoelectric focusing gels. Although the disruption in *csuF* did not affect the chondroitin lyase activity detected by our assay procedure, we could not rule out the possibility that this gene encoded a low-activity lyase or a CS-degrading hydrolase (hyaluronidase). The standard assay for chondroitin lyase activity, i.e., increase in A_{235} due to production of unsaturated disaccharides, does not detect hydrolytic cleavage of CS.

To determine whether CsuF had CS-degrading activity, we first assayed chondroitin lyase activity in extracts of *Bacteroides* strain CS4 carrying the entire *csuF* gene on a multicopy plasmid (complementing clone pNCS4, copy number of 10 to 20). Chondroitin lyase specific activity in extracts from this strain was the same as that in extracts from CS4 (0.4 U/mg of protein). A second approach was to determine if any CS-degrading enzyme activity could be detected in *E. coli*. A strong *E. coli* promoter, P_{tac}, was fused to the *Nde*I site, 70 bp upstream of the putative start codon of the *csu*F ORF. A prominent protein of 90 to 100 kDa was seen on a Coomassie blue-stained SDS-gel of cell extracts from this strain (data not shown). No CS-degrading activity was detected in the cell extract either by the lyase assay or by paper chromatography of incubation mixtures containing cell extracts and CS (data not shown). The paper chromatography assay detects hydrolase as well as lyase activity. Thus, we found no direct evidence that CsuF is a chondroitin lyase or a CS-degrading hyaluronidase.

To determine whether the amino acid sequence of *csuF* resembled those of known chondroitin lyases or hyaluronidases, the EMBL and SWISS databases were searched for proteins with similarity to CsuF protein. CsuF did not show significant overall amino acid similarity with any proteins in the databases. The complete sequences of two hyaluronidase were available in the databases. One, a hyaluronidase gene, *hylP*, with a complete coding sequence of 371 amino acids (16), is encoded on bacteriophage H4489A (from *Streptococcus pyogenes*). The other is a *Clostridium perfringens* hyaluronidase gene, *nagH*, with a complete coding sequence of 1,042 amino acids (1). CsuF had only 21% overall amino acid identity to HylP and 17% identity to NagH. Although the phage hyaluronidase and the clostridial hyaluronidase were also quite different from each other in size and amino acid sequence (21% identity between HylP and NagH), none of the short sequence similarities seen when these two proteins were compared were seen in CsuF. Thus, CsuF appears to be unrelated to these hyaluronidases.

Since no chondroitin lyase sequences had been deposited in the databases, we decided to sequence a portion of the *Bacte-*

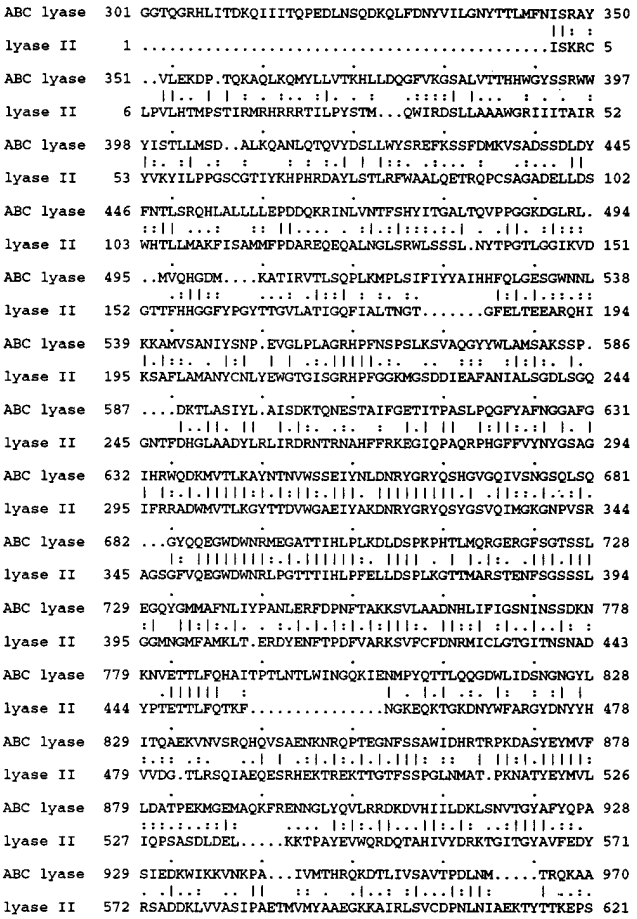

FIG. 4. Comparison of amino acid sequences of chondroitin lyase II from *B. thetaiotaomicron* and chondroitin ABC lyase from *P. vulgaris*. Two-thirds of the lyase II close to the carboxyl terminus was sequenced. Identical amino acids are indicated by vertical lines. Amino acids whose comparison value (scoring based on a simplification scheme for amino acid differences involving a University of Wisconsin Genetics Computer Group computer analysis program) is greater than or equal to 0.50 are indicated by colons. Amino acids whose comparison value is greater than or equal to 0.10 are indicated by dots. The overall amino acid identity between the two sequences was 32%, and the amino acid similarity between the two sequences (comparison value greater or equal to 0.50) was 54%. C-terminal portions of the proteins, where the active site was proposed to be located, had higher similarity (about 50% amino acid identity).

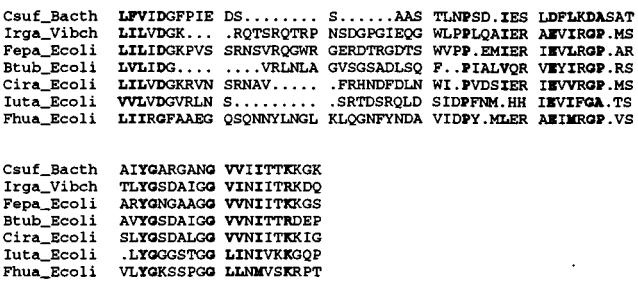

FIG. 5. Alignment of bacterial outer membrane N termini with that of CsuF. Boldface type indicates conserved amino acid residues (scoring greater or equal 0.5; see the legend to Fig. 4 for details) among the seven proteins. CsuF has 1,065 amino acids; the location of the domain starts at residue 186. IrgA has 652 amino acids; the location of the domain starts at residue 101. FepA has 745 amino acids; the location of the domain starts at residue 105. BtuB has 614 amino acids; the location of the domain starts at residue 97. CirA has 663 amino acids; the location of the domain starts at residue 100. IutA has 723 amino acids; the location of the domain starts at residue 100. FhuA has 747 amino acids; the location of the domain starts at residue 123. Bacth, *B. thetaiotaomicron*; Vibch, *V. cholerae*; Ecoli, *E. coli*.

roides gene *csuB*, which encodes chondroitin lyase II. *csuB* was previously cloned from *B. thetaiotaomicron* (12) and is about 3 kbp in size. An internal portion of the *csuB* gene, a 2-kbp *Hin*dIII-*Eco*RV fragment containing two-thirds of the gene close to the carboxyl terminus, was subcloned and sequenced. The *Hin*dIII-*Eco*RV fragment contained an ORF of 1,936 bp, without any stop codons. The deduced amino acid sequence of the sequenced portion of *csuB* had only 19% identity with that of the *csuF* gene. During the preparation of this report, the sequence of another chondroitin lyase, chondroitin ABC lyase from *Proteus vulgaris*, was published by Sato et al. (30). The predicted amino acid sequence of the chondroitin ABC lyase is composed of 1,021 residues, including the 24 residues of the signal peptide. Chondroitin ABC lyase of *P. vulgaris* had 32% overall amino acid identity with the portion of the *B. thetaiotaomicron* chondroitin lyase II that we sequenced (Fig. 4). Although the level of overall identity was not very high, the percent identity between the C-terminal portions of the two lyases, where Sato et al. (30) proposed that the active site of chondroitin ABC lyase might be located, was about 50%. The fact that CsuF not only had only 19% overall amino acid identity with CsuB but also lacked the cluster of high conservation shared by both lyases in the C-terminal portion suggested that CsuF was not a chondroitin lyase.

CsuF is an outer membrane protein. Besides the putative signal peptide sequence present at the amino-terminal end of CsuF, a region near the amino-terminal end of the protein exhibited amino acid sequence similarity to amino termini of known outer membrane proteins. This region included about 60 amino acid residues (amino acids 186 to 240; Fig. 3). In Fig. 5, CsuF is compared with IrgA, the iron-regulated outer membrane protein of *Vibrio cholerae* (8); FepA, the ferric-enterochelin receptor of *E. coli* (21); BtuB, the vitamin B_{12} receptor of *E. coli* (14); CirA, the colicin I receptor of *E. coli* (24); IutA, the ferric-aerobactin receptor of *E. coli* (18); and FhuA, the ferrichrome receptor of *E. coli* (3). The localized similarity at the outer membrane targeting region between CsuF and any of the proteins in the *tonB*-dependent outer membrane protein family was about 40%, whereas these proteins had 49 to 65% identity with each other in the 60-amino-acid region. Given that *Bacteroides* spp. are so distant genetically from *E. coli* and related species, this level of similarity is fairly high. Moreover, CsuF had a C-terminal phenylalanine residue as its last amino acid, a feature shared by many known outer membrane proteins (34). These features of CsuF suggested that it might be an outer membrane protein.

To determine the cellular location of CsuF, we generated polyclonal antibodies against a synthetic peptide, which was designed from the deduced amino acid sequence of CsuF, and used this antibody preparation to localize CsuF. Results of the Western blot analysis are shown in Fig. 6. The antiserum detected a protein of the size expected for CsuF. This protein was seen only in wild-type *B. thetaiotaomicron*, which had been grown on CS, and not in bacteria grown on glucose. This finding was consistent with the GUS fusion results, which showed that expression of *csuF* was regulated by CS. The protein detected by the antiserum was not produced in the mutant, CS4, but was produced by CS4 carrying the complementing clone, pHR4. The protein was located in the mem-

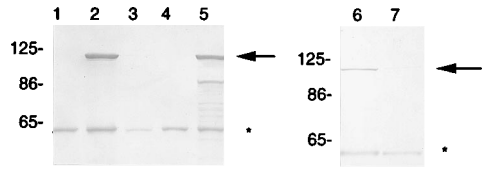

FIG. 6. Cellular location of CsuF. Antiserum raised against a synthetic peptide was used to detect CsuF on Western blots. Approximately 25 μg of protein was loaded in each lane. The arrow at the right indicates the size predicted for CsuF. Sizes of standards are given on the left in kilodaltons. Lanes 1 and 2 contain membranes from *B. thetaiotaomicron* 4001, which had been grown on glucose (lane 1) or CS (lane 2). Lane 3 contains soluble proteins from *B. thetaiotaomicron* 4001 grown on CS. Lane 4 contains membrane proteins from *B. thetaiotaomicron* CS4 grown on glucose. Lane 5 contains membranes from a derivative of *B. thetaiotaomicron* CS4, which carries the complementing clone, pHR4, and was grown on CS. Lane 6 and 7 contain outer membrane and cytoplasmic membrane fractions, respectively. The asterisk at the right marks a streptavidin-binding protein that is present in extracts of *B. thetaiotaomicron* 4001. This protein does not cross-react with the antiserum used but is stained by the reagents used to detect bound antibody.

brane fraction and was not seen in the soluble fraction. Finally, CsuF fractionated with outer membranes and was not seen at all in the cytoplasmic membrane fraction (Fig. 6). Previously, outer membrane proteins had been shown to be coregulated with the CS-degrading enzymes (17), but it was not clear whether these outer membrane proteins were essential for CS utilization. Our findings provide the first evidence that an outer membrane protein is essential for CS utilization. The predicted pI of CsuF is high, as expected for a protein that binds a negatively charged substrate such as CS or HA. If CsuF is a CS-binding protein, however, it appears not to have the same CS-binding motif as the chondroitin lyases. Since radioactively labeled CS or HA is not available commercially, it is not possible to test CS binding to CsuF directly. An alternative possibility is that CsuF is a scaffolding protein that helps to stabilize the actual receptor complex rather than one of the receptor proteins. Further work is needed to determine the precise function of CsuF in CS utilization.

ACKNOWLEDGMENTS

This work was supported by grant AI 17876 from the National Institutes of Health.

We thank Vivian Hwa for the initial cloning of pCS4A.

REFERENCES

1. **Canard, B., T. Garnier, B. Saint-Joanis, and S. T. Cole.** Unpublished data.
2. **Cheng, Q., V. Hwa, and A. A. Salyers.** 1992. A locus that contributes to colonization of the intestinal tract by *Bacteroides thetaiotaomicron* contains a single gene (*chuR*) that links two polysaccharide utilization pathways. J. Bacteriol. **174:**7185–7193.
3. **Coulton, J. W., P. Mason, D. R. Cameron, G. Carmel, R. Jean, and H. N. Rode.** 1986. Protein fusions of β-galactosidase to the ferrichrome-iron receptor of *Escherichia coli* K-12. J. Bacteriol. **165:**181–192.
4. **Dalbey, R. E., and G. von Heijine.** 1992. Signal peptidases in prokaryotes and eukaryotes—a new protease family. Trends Biochem. Sci. **17:**474–478.
5. **Davidson, J., M. Henstersprеute, N. Chevalier, V. Ha-Thi, and F. Brunel.** 1987. Vectors with restriction site banks. V. pJRD215, a wide-host-range cosmid vector with multiple cloning sites. Gene **51:**275–280.
6. **Feldhaus, M. J., V. Hwa, Q. Cheng, and A. A. Salyers.** 1991. Use of an *Escherichia coli* β-glucuronidase gene as a reporter gene for investigation of *Bacteroides* promoters. J. Bacteriol. **173:**4540–4543.
7. **Glaser, J. H., and H. E. Conrad.** 1979. Chondroitin SO_4 catabolism in chick embryo chondrocytes. J. Biol. Chem. **254:**2316–2325.
8. **Goldberg, M. B., S. A. Boyko, J. R. Butterton, J. A. Stoebner, S. M. Payne, and S. B. Calderwood.** 1992. Characterization of a *Vibrio cholerae* virulence factor homologous to the family of TonB-dependent proteins. Mol. Microbiol. **6:**2407–2418.
9. **Guiney, D. G., P. Hasegawa, and C. E. Davis.** 1984. Expression in *E. coli* of cryptic tetracyclin resistance gene from *Bacteroides* R plasmids. Plasmids **11:**248–252.
10. **Guiney, D. G., and E. Yakobson.** 1983. Location and nucleotide sequence of the transfer origin of the broad host range plasmid RK2. Proc. Natl. Acad. Sci. USA **80:**3593–3598.
11. **Guthrie, E. P., and A. A. Salyers.** 1986. Use of targeted insertional mutagenesis to determine whether chondroitin lyase II is essential for chondroitin sulfate utilization by *Bacteroides thetaiotaomicron*. J. Bacteriol. **166:**966–971.
12. **Guthrie, E. P., N. B. Shoemaker, and A. A. Salyers.** 1985. Cloning and expression in *Escherichia coli* of a gene coding for a chondroitin lyase from *Bacteroides thetaiotaomicron*. J. Bacteriol. **164:**510–515.
13. **Hanahan, D.** 1983. Studies on the transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166:**557–580.
14. **Heller, K., and R. J. Kadner.** 1985. Nucleotide sequence of the gene for the vitamin B_{12} receptor protein in the outer membrane of *Escherichia coli*. J. Bacteriol. **161:**904–908.
15. **Holdeman, L. V., and W. E. C. Moore (ed.).** 1975. Anaerobe laboratory manual, 4th ed. Virginia Polytechnic Institute and State University, Blacksburg.
16. **Hynes, W. L., and J. J. Ferretti.** 1989. Sequence analysis and expression in *Escherichia coli* of the hyaluronidase gene of *Streptococcus pyogenes* bacteriophage H4489A. Infect. Immun. **57:**533–539.
17. **Kotarski, S. F., J. Linz, D. M. Braun, and A. A. Salyers.** 1985. Analysis of outer membrane polypeptides that are associated with the growth of *Bacteroides thetaiotaomicron* on chondroitin sulfate. J. Bacteriol. **163:**1080–1086.
18. **Krone, J. A., F. Stegehuis, G. Koningstein, C. van Doorn, B. Roosendaal, F. K. de Graaf, and B. Oudega.** 1985. Characterization of the pColV-K30 encoded cloacin DF13/aerobactin outer membrane receptor protein of *Escherichia coli*: isolation and purification of the protein and analysis of its nucleotide sequence and primary structure. FEMS Microbiol. Lett. **26:**153–161.
19. **Linn, S., T. Chan, L. Lipeski, and A. A. Salyers.** 1983. Isolation and characterization of two chondroitin lyases from *Bacteroides thetaiotaomicron*. J. Bacteriol. **156:**859–866.
20. **Lowry, O. H., N. J. Rosebrough, A. L. Farr, and R. J. Randall.** 1951. Protein measurement with Folin phenol reagent. J. Biol. Chem. **193:**265–275.
21. **Lundrigan, M. D., and R. J. Kadner.** 1986. Nucleotide sequence of the gene for the ferrienterochelin receptor FepA in *Escherichia coli*: homology among outer membrane receptors that interact with TonB. J. Biol. Chem. **261:**10797–10801.
22. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
23. **Moore, W. E. C., and L. V. Holdeman.** 1974. Human fecal flora. Appl. Microbiol. **27:**961–979.
24. **Nau, C. D., and J. Konisky.** 1989. Evolutionary relationship between the TonB-dependent outer membrane transport proteins: nucleotide and amino acid sequences of the *Escherichia coli* colicin I receptor gene. J. Bacteriol. **171:**1041–1047.
25. **Salyers, A. A., and E. P. Guthrie.** 1988. A deletion in the chromosome of *Bacteroides thetaiotaomicron* that abolishes production of chondroitinase II does not affect survival of the organism in gastrointestinal tracts of exgermfree mice. Appl. Environ. Microbiol. **54:**1964–1969.
26. **Salyers, A. A., and S. F. Kotarski.** 1980. Induction of chondroitin sulfate lyase activity in *Bacteroides thetaiotaomicron*. J. Bacteriol. **143:**781–788.
27. **Salyers, A. A., M. Pajeau, and R. E. McCarthy.** 1988. Importance of mucopolysaccharides as substrates for *Bacteroides thetaiotaomicron* growing in intestinal tracts of exgermfree mice. Appl. Environ. Microbiol. **54:**1970–1976.
28. **Salyers, A. A., and M. O'Brien.** 1980. Cellular location of enzymes involved in chondroitin sulfate breakdown by *Bacteroides thetaiotaomicron*. J. Bacteriol. **143:**772–780.
29. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
30. **Sato, N., M. Shimada, H. Nakajima, H. Oda, and S. Kimura.** 1994. Cloning and expression in *Escherichia coli* of the gene encoding the *Proteus vulgaris* chondroitin ABC lyase. Appl. Microbiol. Biotechnol. **41:**39–46.
31. **Shoemaker, N. B., C. E. Getty, J. F. Gardner, and A. A. Salyers.** 1986. Tn*4351* transposes in *Bacteroides* and mediates the integration of R751 into the *Bacteroides* chromosome. J. Bacteriol. **165:**929–936.
32. **Shoemaker, N. B., C. E. Getty, E. P. Guthrie, and A. A. Salyers.** 1986. Regions in *Bacteroides* plasmids pBFTM10 and pB8-51 that allow *Escherichia coli-Bacteroides* shuttle vector to be mobilized by IncP plasmids and by a conjugative *Bacteroides* tetracycline resistance element. J. Bacteriol. **166:**959–965.
33. **Simon, R., U. Priefer, and A. Puhler.** 1983. A broad host range mobilization system for *in vivo* genetic engineering: Transposon mutagenesis in gram negative bacteria. Bio/Technology **1:**784–791.
34. **Struyve, M., M. Moons, and J. Tommassen.** 1991. Carboxy-terminal phenylalanine is essential for the correct assembly of a bacterial outer membrane protein. J. Mol. Biol. **218:**141–148.
35. **Vercellotti, J. R., A. A. Salyers, W. S. Bullard, and T. D. Wilkins.** 1977. Breakdown of mucin and plant polysaccharides in the human colon. Can. J. Biochem. **55:**1190–1196.