INFECTION AND IMMUNITY, July 2001, p. 4342–4350 Vol. 69, No. 7
0019-9567/01/$04.00+0 DOI: 10.1128/IAI.69.7.4342–4350.2001

Polysaccharide Biosynthesis Locus Required for Virulence of *Bacteroides fragilis*

MICHAEL J. COYNE,[1] ARTHUR O. TZIANABOS,[1] BENJAMIN C. MALLORY,[1] VINCENT J. CAREY,[1] DENNIS L. KASPER,[1,2] AND LAURIE E. COMSTOCK[1]*

Channing Laboratory, Department of Medicine, Brigham and Women's Hospital,[1] and Department of Microbiology and Molecular Genetics,[2] Harvard Medical School, Boston, Massachusetts 02115

Received 25 January 2001/Returned for modification 10 April 2001/Accepted 19 April 2001

***Bacteroides fragilis,* though only a minor component of the human intestinal commensal flora, is the anaerobe most frequently isolated from intra-abdominal abscesses. *B. fragilis* 9343 expresses at least three capsular polysaccharides—polysaccharide A (PS A), PS B, and PS C. Purified PS A and PS B have been tested in animal models and are both able to induce the formation of intra-abdominal abscesses. Mutants unable to synthesize PS B or PS C still facilitate abscess formation at levels comparable to those of wild-type 9343. To determine the contribution of PS A to abscess formation in the context of the intact organism, the PS A biosynthesis region was cloned, sequenced, and deleted from 9343 to produce a PS A-negative mutant. Animal experiments demonstrate that the abscess-inducing capability of 9343 is severely attenuated when the organism cannot synthesize PS A, despite continued synthesis of the other capsular polysaccharides. The PS A of 9343 contains an unusual free amino sugar that is essential for abscess formation by this polymer. PCR analysis of the PS A biosynthesis loci of 50 *B. fragilis* isolates indicates that regions flanking each side of this locus are conserved in all strains. The downstream conserved region includes two terminal PS A biosynthesis genes that homology-based analyses predict are involved in the synthesis and transfer of the free amino sugar of PS A. Conservation of these genes suggests that this sugar is present in the PS A of all serotypes and may explain the abscessogenic nature of *B. fragilis*.**

Bacteroides fragilis is an important cause of intra-abdominal abscesses arising from introduction of this organism into the peritoneal cavity via surgical, traumatic, or disease-induced bowel perforation. The capsular polysaccharide complex of the *B. fragilis* strain used for the study of abscess formation, NCTC 9343, is a potent inducer of abscesses. It is composed of at least three distinct capsular polysaccharides: polysaccharide A (PS A), PS B, and PS C. Purified PS A and PS B are each capable of inducing intra-abdominal abscesses in a rodent abscess model, although PS A is more potent: the dose of purified PS A causing abscesses in 50% of animals (AD_{50}) is 0.67 μg, whereas the AD_{50} of PS B is 25 μg (30). PS C has not yet been purified to homogeneity.

The structures of the PS A and PS B capsular polysaccharides of strain 9343 have been determined; each of these polysaccharides contains both positively and negatively charged groups. The repeating unit of PS A has one free amino and one carboxyl group, whereas the repeating unit of PS B has one free amino and two negatively charged (carboxyl and phosphonate) groups (1). The genetic composition of the PS C biosynthesis locus suggests that the PS C repeating unit has one carboxyl group (5).

The presence of both a free amino and at least one negatively charged group is essential for these polysaccharides to induce abscesses (30). Chemical conversion of a free amino group to a positively charged tertiary amine or to a neutral *N*-acetyl group results in loss of abscessogenic activity (29). We have shown that polysaccharides with repeating unit structures that differ from those of PS A or PS B but that possess both free amino and negatively charged groups are also capable of inducing abscesses in animals. Two surface polysaccharides from *Streptococcus pneumoniae*, the type 1 capsular polysaccharide and teichoic acid (C substance), each contain the same free amino sugar component as PS A (10, 16) and have AD_{50}s of 31 and 5 μg, respectively (30). In addition, repeating units with one negatively charged group acquire the ability to induce abscesses in the animal model when chemically converted to possess free amino groups (30).

Although both PS A and PS B in purified form have been shown to induce abscesses in animal models, there has been no demonstration that these are the only, or most potent, abscess-inducing molecules produced by *B. fragilis*. The contribution of PS B to abscess formation was analyzed using a mutant in which 5.7 kb of the PS B biosynthesis region was deleted. This PS B-negative mutant was fully virulent in the rodent abscess model (9). The contribution of PS C to the abscess-inducing ability of 9343 was similarly analyzed using a mutant with a chromosomal deletion in the PS C biosynthesis locus. This mutant was also not attenuated in its ability to induce abscesses (6).

In this study, we identified and characterized the PS A biosynthesis locus of 9343. Through the creation of a PS A-negative mutant strain, the singular importance of this polysaccharide for abscess induction by *B. fragilis* was revealed.

MATERIALS AND METHODS

Bacteria, plasmids, and media. Bacterial strains and plasmids are described in Table 1. *Escherichia coli* strains were grown in L broth or on L agar plates. *B. fragilis* strains were grown anaerobically in basal medium (23) or on brain heart infusion plates supplemented with hemin (50 μg/ml) and menadione (0.5 μg/ml)

* Corresponding author. Mailing address: Channing Laboratory, 181 Longwood Ave., Boston, MA 02115. Phone: (617) 525-2679. Fax: (617) 731-1541. E-mail: lcomstock@channing.harvard.edu.

TABLE 1. Bacterial strains and plasmids

Strain or plasmid	Description	Reference or source
Strains		
E. coli DH5α	F^- ϕ80d*lacZ*ΔM15 Δ(*lacZYA-argF*) *U169 deoR recA1 endA1 hsdR17*($r_K^-m_K^+$) *phoA supE44* λ^- *thi-1 gyrA96 relA1*	3
Bacteroides distasonis 8503	Type strain	14
B. fragilis B16	Clinical isolate	A. B. Onderdonk
B. fragilis B16 mutant 65	Strain B16 with pNJR6Ω4 insertion in PS A region	This study
B. fragilis NCTC 9343	Type strain, appendix abscess	14; NCTC [a]
B. fragilis 9343ΔPSA	9343 mutant with a chromosomal deletion removing *wzx-wcfS* of the PS A region	This study
Plasmids		
pHC79	Cosmid vector; Ap^r	12
pBluescript II SK	Phagemid cloning vector; Ap^r	Stratagene
R751	Mobilizable plasmid used to move plasmids from *E. coli* to *B. fragilis*; IncPβ Tra^+ Tp^r (Tn*402*)	20
pNJR6	*Bacteroides* suicide vector; *mob*$^+$ Tra^- Km^r	28
pNJR6Ω4	Plasmid pNJR6 containing duplicate copies of Tn*4351*; Em^r (*Bacteroides*) Km^r (*E. coli*) *mob*$^+$ Tra^-	28
pLEC12	Junctional clone of *B. fragilis* B16 CE3-negative mutant containing a portion of pNJR6Ω4 and several kb of *B. fragilis* B16 DNA; Km^r	This study
pLEC15	*B. fragilis* 9343 gene bank clone in pHC79; Ap^r	This study
pLEC15.1	4.3-kb *Eco*RI subclone of pLEC15 in pBluescript; Ap^r	This study
pLEC21	*B. fragilis* 9343 gene bank clone in pHC79; Ap^r	This study
pLEC21.1–pLEC21.4	Subclones of pLEC21 in pBluescript; Ap^r	This study
pMJC15	Plasmid pNJR6 containing 4.4 and 4.3 kb of the left and right flanks of the 9343 PS A locus; respectively; *mob*$^+$ Tra^-Km^r	This study

[a] National Collection of Type Cultures.

(BHIS). The following concentrations of antibiotics were added when appropriate: ampicillin, 100 μg/ml; kanamycin, 20 μg/ml; erythromycin, 5 μg/ml; trimethoprim, 100 μg/ml; gentamicin, 200 μg/ml.

Transposon mutagenesis to locate the PS A biosynthesis locus. As 9343 is somewhat resistant to the introduction of foreign DNA, *B. fragilis* strain B16 (also reactive with a monoclonal antibody [MAb] to the PS A of 9343, MAb CE3 [23]) was used for transposon mutagenesis. Transposon suicide vector pNJR6Ω4 was conjugally transferred, as previously described (6), from *E. coli* DH5α into *B. fragilis* B16 using helper plasmid R751 (27). *B. fragilis* transconjugants acquiring resistance to erythromycin encoded by pNJR6Ω4 were screened for loss of MAb CE3 reactivity. One such CE3-negative B16 clone, mutant 65, was further analyzed.

Isolation of 9343 cosmid clones containing the PS A biosynthesis locus and generation of subclones. The chromosome of *B. fragilis* B16 mutant 65 (MAb CE3 negative) yielded a *Hin*dIII fragment containing the origin of replication and Km^r gene of pNJR6Ω4, as well as several kilobases of *B. fragilis* B16 DNA. This fragment was cloned by ligation of *Hin*dIII-digested mutant 65 chromosomal DNA, transformation of DH5α, and selection with kanamycin. The resulting plasmid, pLEC12, has a very low copy number; therefore, a PCR product designated P65 containing a portion of the B16 junctional DNA was synthesized. Using the P65 PCR product as a probe, two pHC79-based cosmid clones, pLEC15 and pLEC21 (Fig. 1A), were selected from a *B. fragilis* 9343 gene bank. Restriction fragments of pLEC15 and pLEC21 were subcloned into pBluescript II SK (Stratagene, La Jolla, Calif.) and used as sequencing templates (Fig. 1A).

Creation of deletion mutant 9343ΔPSA. To create mutant 9343ΔPSA, DNA upstream and downstream of the region to be deleted was ligated into the *Bacteroides* conjugal suicide vector pNJR6 (28). The upstream flank consists of a 4.3-kb *Eco*RI fragment of pLEC21.2 (the upstream *Eco*RI site is vector-based), and the downstream flank consists of the 4.2-kb *Eco*RI insert of pLEC15.1 (Fig. 1A). These *Eco*RI fragments were cloned into the *Eco*RI site of pNJR6 in a three-way ligation, and *E. coli* DH5α was transformed with this ligation mixture. Kanamycin-resistant colonies were screened by PCR for proper orientation of the left and right flanking DNA. Plasmid pMJC15 thus contains both flanks of the 9343 PS A locus ligated in the correct orientation (representing bp 1 through 4350 and 12536 through 16792 of the sequence reported here, for the left and right flanks, respectively).

pMJC15 was conjugally transferred into 9343, and cointegrates were selected by Em^r encoded by pNJR6. The cointegrate strain was passaged three times in basal medium to allow the cointegrate state to resolve by recombination and plated onto BHIS. The resulting colonies were replica plated to BHIS containing erythromycin, and erythromycin-sensitive colonies were screened by PCR to distinguish wild-type revertants from strains acquiring the desired mutant genotype. Clones demonstrating the mutant genotype by PCR were confirmed to be lacking the PS A locus by Southern blot analysis. The resulting mutant, *B. fragilis* 9343 ΔPSA, has 8,185 bp of the PS A region deleted, which impacts all nine biosynthesis genes of the locus (Fig. 1C).

PCR. The locations of the PCR primers and the DNA amplified by each reaction in relation to the 9343 PS A locus are diagrammed in Fig. 1B. The sequences of the primers and the parameters used for each PCR are shown in Table 2. The PCR mixtures contained 20 μl of PCR Supermix (Life Technologies, Gaithersburg, Md.), an additional 0.5 U of *Taq* polymerase (Life Technologies), 3.2 pmol of each primer, and ~20 ng of chromosomal DNA. Each PCR was performed at least twice. Extended PCRs (EPCRs) were carried out as previously described (7).

DNA sequencing and analysis. DNA sequencing and analysis of the *B. fragilis* PS A region were performed as previously described (6). Comparison of WcfR and WcfS with the *S. pneumoniae* unfinished microbial genome was performed using TBLASTN with sequence obtained from The Institute for Genomic Research through the website at http://www.tigr.org.

SDS-polyacrylamide gel electrophoresis and immunoblotting. Bacterial cell lysates were run on discontinuous sodium dodecyl sulfate (SDS)-polyacrylamide gels (4 to 20% gradient; ESA, Inc. Chelmsford, Mass.) and transferred to polyvinylidene difluoride Immobilon membranes (Millipore Corp., Bedford, Mass.) by standard techniques. Immunoblotting was performed as described previously (6). Culture supernatants of MAbs G9 (9343 PS B specific [22]) and QUBF7 (*B. fragilis* 9343 PS C specific [18]) were used at 1:50; ascites fluid containing the PSA-specific MAb CE3 was diluted 1:2,000 before use.

Mouse abscess experiments. A previously described animal model of intra-abdominal infection (26) was modified for these studies. Briefly, male C57BL/6 mice (4 to 6 weeks old; Charles River Laboratories, Wilmington, Mass.) were challenged via the intraperitoneal route with 0.1 ml of inoculum containing 10-fold serial dilutions of 9343 or 9343ΔPSA mixed 1:1 (vol/vol) with sterilized rat fecal contents. Rat fecal contents are used as an adjuvant for abscess formation and do not induce abscesses when implanted alone into the peritonea of mice. Six days after challenge, animals were necropsied and examined for intra-abdominal abscesses. The presence of one or more abscesses in an animal was scored as a positive result. To compare the abscess-inducing potentials of these strains, the AD_{50} was estimated for each strain using linear logistic regression.

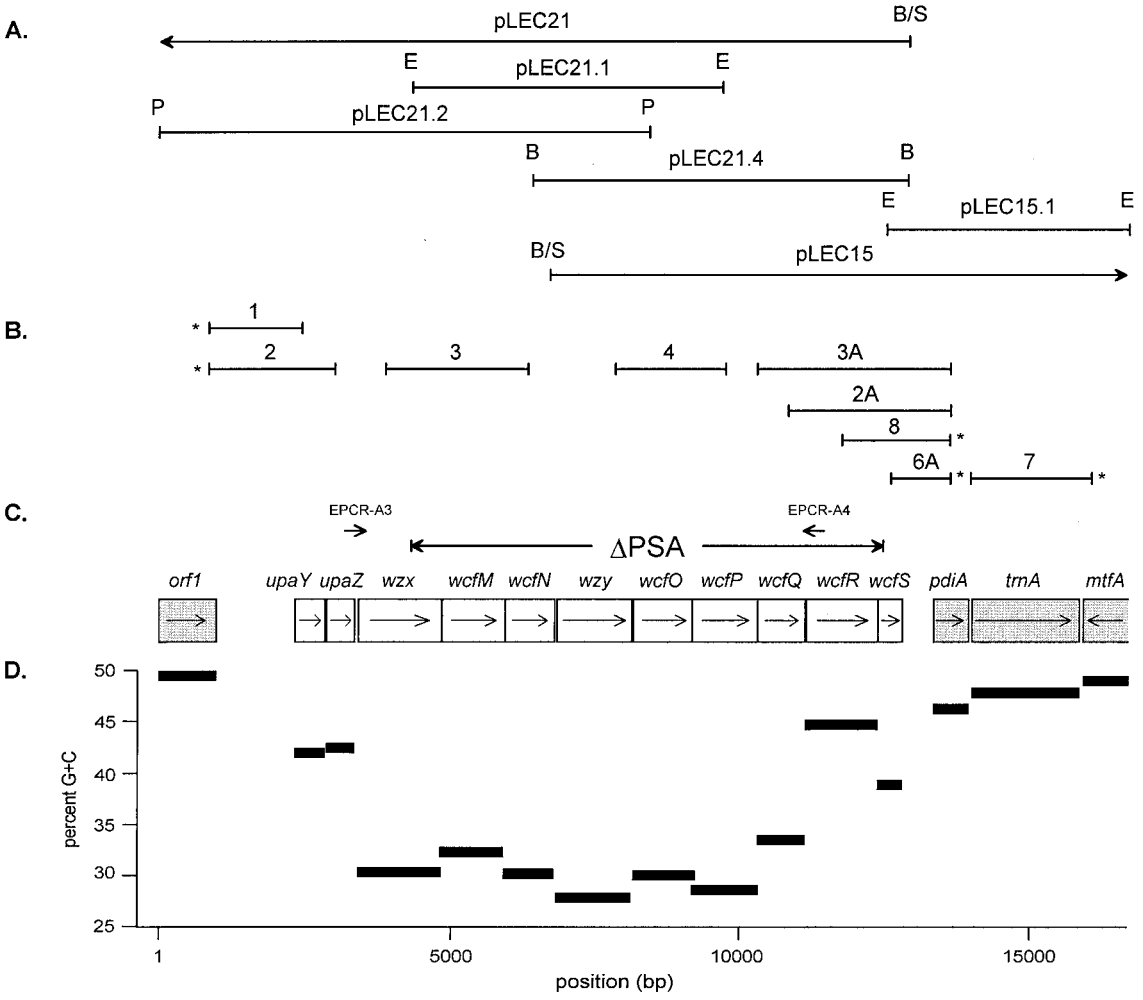

FIG. 1. (A) Cosmids and subclones used in sequencing the 9343 PS A locus. B, *Bam*HI; E, *Eco*RI; P, *Pst*I; B/S, compound site resulting from cloning *Sau*3AI-digested DNA into the vector *Bam*HI site. (B) Diagrammatic representation of the PCRs showing the positions of the primers and the span of the product. PCRs marked with an asterisk were successful in more than 98% of the strains tested. (C) ORF map of the *B. fragilis* 9343 PS A locus. The region deleted in mutant strain 9343ΔPSA and the positions of the primers used for EPCR are indicated above the ORFs. Flanking ORFs not contributing to the production of PS A are shaded. Arrows within the ORFs indicate direction of transcription. (D) G+C content of each of the 9343 PS A locus ORFs.

The likelihood ratio test was used to test for distinct AD_{50}s for the two strains (8).

Nucleotide sequence accession number. The nucleotide sequence described in this work has been submitted to GenBank and assigned accession number AF189282.

RESULTS

Cloning, characterization, and mutation of the PS A locus. As a first step in determining the contribution of PS A to the abscess-inducing capabilities of *B. fragilis* 9343, the PS A locus was cloned, sequenced, and mutated. In all, a 16,792-bp region of the 9343 chromosome was sequenced in both directions, using as templates two cosmid clones isolated from the *B. fragilis* 9343 gene bank (pLEC15 and pLEC21) and subclones derived from them (Fig. 1A).

The PS A locus contains nine genes whose products are similar to other proteins involved in polysaccharide biosynthesis (Table 3). In addition, upstream of the biosynthesis genes are two open reading frames (ORFs) (called *upaY* and *upaZ*, for upstream polysaccharide A) whose products are very similar to those of two small ORFs found just upstream of both the PS B locus (*upbY* and *upbZ*) and the PS C locus (*upcY* and *upcZ*, previously known as *orf3* and *orf4*). The 11 genes of the PS A locus, including *upaY* and *upaZ*, are all transcribed in the same direction and are tightly clustered (Fig. 1C). These genetic characteristics are similar to those of both the PS B and PS C biosynthesis loci and suggest that the PS A locus is an operon. The PS A locus as a whole is AT rich compared with the *B. fragilis* chromosome (41 to 44% G+C), with an average overall G+C content of 33.3% and the content in individual genes ranging from 27.7% (*wzy*) to 44.7% (*wcfR*) (Fig. 1D).

A defined PS A deletion mutant, 9343ΔPSA, was created by removal of an 8,185-bp region of the PS A locus, affecting all nine biosynthesis genes (Fig. 1C). Immunoblot analysis demonstrated that 9343ΔPSA fails to react with MAb CE3 (Fig. 2A), thus confirming that this mutant is unable to synthesize PS A. The 9343ΔPSA mutant expresses both PS B and PS C, as

TABLE 2. Primers and thermal cycler programs used for the PCRs in this study

PCR	Primer	Primer sequence (5′ → 3′)	Primer 5′ position[a]	Gene targeted	Program[b]
1	C464	ACACATATCACTTCCGATGCC	848	*orf1*	94°C, 30 s; 59°C, 30 s; 72°C, 90 s
	C473	AGAAAACTCCTGGTCCTTCTTTG	2484	*upaY*	
2	C464	ACACATATCACTTCCGATGCC	848	*orf1*	94°C, 30 s, 59°C, 30 s, 72°C, 90 s
	C444	GTTGACGGAAATGATCGGTATAG	3056	*upaZ*	
2A	C396	TGCTGATGATATTTCCATGCC	10686	*wcfQ*	94°C, 30 s, 59°C, 30 s, 72°C, 120 s
	C507	CAGAGTAATGGTAAAATCGGGAG	13648	*pdiA*	
3	C432	TAACACGATAGGAGTTGCATGG	3881	*wzx*	94°C, 30 s, 59°C, 30 s, 72°C, 90 s
	C393	ACATTGAGAAATACTCGTCCACC	6366	*wcfN*	
3A	C379	GGCTGAAGTTGTTTGGAAGTTG	10270	*wcfP*	94°C, 30 s, 59°C, 30 s, 72°C, 120 s
	C507	CAGAGTAATGGTAAAATCGGGAG	13648	*pdiA*	
4	C202	GATATTTGGTCACATTTGGGG	7814	*wzy*	94°C, 30 s, 59°C, 30 s, 72°C, 90 s
	C349	GAATTCAGCAGCTTTGTATCCAC	9757	*wcfP*	
6A	C510	AGATCCGAGAATAACTCGTACC	12716	*wcfS*	94°C, 30 s, 58°C, 30 s, 72°C, 60 s
	C507	CAGAGTAATGGTAAAATCGGGAG	13648	*pdiA*	
7	C505	AATGTCCTGATAGACCGGGAG	13940	*pdiA*	94°C, 30 s, 59°C, 30 s, 72°C, 90 s
	C500	CAAGGCCGAAATGATAAAAGAG	16081	*mtfA*	
8	C429	CACTCTTTGGGAGCACGTTAC	11753	*wcfR*	94°C, 30 s, 58°C, 30 s, 72°C, 90 s
	C507	CAGAGTAATGGTAAAATCGGGAG	13648	*pdiA*	
EPCR	EPCR-A3	CTTTTGCCGTATTGGACCATCTTCCCGCTTCTC	3219	*upaZ*	94°C, 30 s, 68°C, 20 min
	EPCR-A4	GGAACAATCACTTCATCGCCCTCCTTCACTCCC	11472	*wcfR*	

[a] In base pairs.
[b] All programs had an initial segment at 94°C for 2 min and were run for 30 cycles.

indicated by its continued reactivity with MAbs G9 and QUBF7 (Fig. 2B and C).

Virulence studies. The contribution of PS A to the ability of *B. fragilis* 9343 to cause abscesses was studied using a mouse abscess model. As outlined in Table 4, the PS A mutant was severely attenuated in its ability to induce abscesses. The wild-type strain consistently induces abscesses in at least 80% of animals at a dose of 10^8 organisms, with a titrating effect as the dose is lowered. In contrast, the PS A mutant, even at the highest dose, was unable to induce the formation of abscesses in 50% of animals. The AD_{50} of the wild-type was $10^{5.2}$ organisms, whereas the AD_{50} of 9343ΔPSA was extrapolated to be $10^{14.9}$ organisms ($P < 0.0001$).

Genetic complement of the PS A locus. PS A is a polymer of the repeating unit {→3)α-*d*-AATGal*p*(1→4)[β-d-Gal*f*(1→3)]α-*d*-Gal*p*NAc(1→3)β-*d*-Gal*p*(1→}, where AATGal is acetamido-amino-2,4,6-trideoxygalactose, and the galactopyranosyl residue is modified by a pyruvate substituent spanning O-4 and O-6 (1). One gene of the PS A locus, *wcfR*, encodes a product that is highly similar to a series of proteins involved in amino sugar biosynthesis (Table 3). This gene product is likely responsible for transfer of the amino group on the AATGal residue of the PS A repeating unit. As this free amino group is essential for the abscess-inducing capabilities of PS A, *wcfR* is expected to be crucial for virulence.

On the basis of the PS A repeating unit structure and of the similarity to other enzyme sequences in the GenBank database, putative functions can be ascribed to many of the other gene products encoded by the PS A locus (Table 3). WcfM is similar to UDP-galactopyranose mutases from several different bacteria. These enzymes catalyze the interconversion of the furanose and pyranose forms of galactose (21). The presence of this gene in the PS A locus was expected, as one of the four sugars of the repeating unit is a galactofuranose.

The locus also encodes four putative transferases (WcfN, WcfP, WcfQ, and WcfS), the number expected for a polysaccharide composed of a four-sugar repeating unit. Homology-based analyses predict that WcfS is responsible for the transfer of the AATGal nucleotide precursor to an undecaprenyl-phosphate lipid carrier as the first step in the synthesis of the PS A repeating unit.

The PS A locus also encodes a putative flippase (*wzx*), responsible for transporting the completed subunits across the inner membrane (17), and a putative polymerase (*wzy*) that links the repeating units together (19). Each of these enzymes are recognizable based on similarities to entries in the GenBank databases (Table 3) and by established characteristics of these proteins.

Conservation of the PS A region among various *B. fragilis* strains. MAb analyses revealed that not only does *B. fragilis* produce multiple capsular polysaccharides, but there is inter-strain heterogeneity in serotypes (22). Because of the importance of the PS A molecule for virulence of *B. fragilis* 9343, the level to which this genetic region is conserved among *B. fragilis* isolates was investigated. A diverse collection of 50 *B. fragilis* strains was used, which included strains isolated from the

TABLE 3. Comparison of products encoded by the *B. fragilis* 9343 PS A locus to sequences in the GenBank database

ORF	Size (aa[a])	Organism and enzyme	Enzyme description	% Identity/similarity[b]	Accession no.
orf1	334	*Mus musculus* Jnkbp1	JNK-binding protein 1	21/39 (95–226)	AB029482
		Pseudomonas aeruginosa OrfC	Serine/threonine protein kinase homolog	34/50 (261–330)	AAD32693
upaY	172	*Bacteroides fragilis* UpbY	Unknown	67/83 (1–170)	AF285774
		Bacteroides fragilis UpcY	Unknown	25/52 (2–168)	AF048749
		Myxococcus xanthus TaA	NusG-like protein	24/41 (8–165)	AJ012097
upaZ	157	*Bacteroides fragilis* UpbZ	Unknown	64/79 (2–157)	AF285774
		Bacteroides fragilis UpcZ	Unknown	40/60 (6–119)	AF048749
wzx	482	*Streptococcus pneumoniae* Cps19BJ	Putative oligosaccharide repeat unit transporter	32/56 (3–476)	AF004325
		Yersinia enterocolitica Wzx	Probable O-antigen transport protein	27/45 (5–416)	Z47767
wcfM	364	*Streptococcus pneumoniae* Cap33fN	UDP-galactopyranose mutase	66/82 (5–360)	AJ006986
		Escherichia coli Glf	UDP-galactopyranose mutase	59/79 (5–360)	P37747
wcfN	291	*Enterococcus faecalis* Orfde 10	Putative glycosyl transferase	31/47 (1–101)	AF071085
		Archaeoglobus fulgidus RfbQ	Putative rhamnosyl transferase	26/46 (3–231)	AE001082
wzy	434	*Vibrio cholerae* WbfQ	Rfc-like protein	21/40 (1–269)	Y07786
		Streptococcus pneumoniae Cps2J	Unknown	22/38 (107–433)	AF026471
wcfO	357	*Bacillus subtilis* YwrD	Gamma-glutamyltransferase homolog	26/47 (27–106)	Z93767
wcfP	378	*Escherichia coli* WbnE	Putative glycosyl transferase	32/51 (81–308)	AF172324
		Bacillus subtilis YveP	Capsular polysaccharide biosynthesis homolog	29/47 (9–372)	Z71928
wcfQ	268	*Erwinia amylovora* AmsE	Glycosyl transferase	33/51 (4–268)	Q46635
		Haemophilus influenzae HI1695	Putative lipopolysaccharide biosynthesis protein	30/51 (1–266)	C64175
wcfR	407	*Streptococcus suis* Cps7G	Putative role in biosynthesis of amino sugar	44/65 (3–393)	AF164515
		Streptococcus pneumoniae type 4	Unknown	43/63 (15–412)	tigr
		Shigella sonnei ORF7S	O-antigen gene	35/56 (3–399)	AB028134
		Bordetella bronchiseptica WlbF	Putative amino sugar biosynthesis protein	32/52 (3–396)	AJ007747
		Pyrococcus abyssi	Aspartate aminotransferase	28/47 (3–396)	D75096
wcfS	195	*Shigella sonnei* ORF8S	O-antigen gene	55/73 (1–194)	AB028134
		Streptococcus suis Cps7F	Putative glycosyltransferase	49/68 (4–195)	AF164515
		Streptococcus pneumoniae type 4	Unknown	44/63 (4–197)	tigr
		Bordetella bronchiseptica WlbG	O-antigen initiating transferase	52/66 (4–197)	AJ007747
pdiA	198	*Deinococcus radiodurans* DR0189	Thiol:disulfide interchange protein	36/51 (55–173)	F75549
		Bacillus subtilis ResA	Thioredoxin related protein	34/55 (59–176)	P35160
tmA	621	*Escherichia coli* YfbS	Putative transport protein	36/59 (7–621)	AE000318
		Deinococcus radiodurans DR1411	Transporter, sodium/sulfate symporter family	31/52 (1–621)	F75398
mtfA	259	*Thermotoga maritima* TM1389	Methyltransferase-related protein	33/55 (82–185)	AE001791
		Deinococcus radiodurans DR0026	Probable methyltransferase	28/48 (70–148)	C75569

[a] Amino acids.
[b] Identities and similarities are based on the segment (in parentheses, in base pairs) of the *B. fragilis* product.

United States, England, and Italy and from a wide spectrum of clinical infections and fecal samples. PCR primer pairs were selected according to the sequence of the 9343 PS A locus such that each primer pair would amplify various portions of this locus or its flanking DNA (Fig. 1B).

As was found to be the case when the diversity of the PS C locus was examined (7), the region upstream of the PS A biosynthesis genes, including *upaY* and *upaZ*, is conserved in all strains (PCR 1) or 49 of 50 strains (PCR 2). Similarly, the region downstream of the PS A locus is also conserved in all strains, as a product was generated by PCR 7 from all 50 strains. Whereas the DNA flanking the 9343 PS A locus was found to be well conserved among the strains tested, the central portion of the locus was not: PCRs 3 and 4 were successful in only 16 of the 50 strains (Fig. 1B and Table 5).

Due to the importance of the free amino group of PS A for virulence of 9343, the distribution of *wcfR*, encoding the aminotransferase homolog, was investigated further. PCRs 6A, 8, 2A, and 3A all utilized the same downstream primer targeted to *pdiA*, a gene downstream of the PS A locus that is conserved in all strains. The upstream primers extended by various degrees into the PS A locus. Two unexpected findings were that PCR 6A and PCR 8 amplified products from all 50 strains tested (Fig. 1B and Table 5). These data indicate that the genes encoding the aminotransferase homolog, thought to be responsible for the synthesis of the critically important free amino

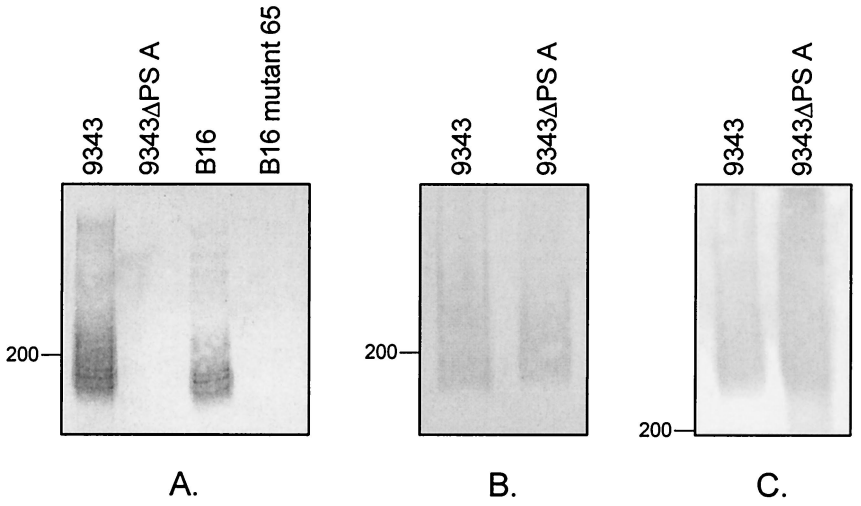

FIG. 2. Western immunoblot of *B. fragilis* wild-type and mutant strains. Bacterial cultures were grown to an optical density at 600 nm of 0.8, pelleted by centrifugation, and lysed by boiling in 1× loading buffer. Bacteria equivalent to 40 μl of the original culture were added to each well of the 4-to-20% gradient SDS-polyacrylamide gel. After separation, the samples were transferred to polyvinylidene difluoride membranes and probed with PS A-specific MAb CE3 (A), PS B-specific MAb G9 (B), and PS C-specific MAb QUBF7 (C). Lanes (all panels): 1. *B. fragilis* 9343; 2. 9343ΔPSA; 3. *B. fragilis* B16; 4. *B. fragilis* B16 mutant 65.

group on AATGal, and the glycosyltransferase that likely initiates PS A synthesis by transfer of AATGal to a lipid carrier are the terminal genes in the PS A locus of all *B. fragilis* strains tested, regardless of the PS A serotype.

The results of the nine PCRs differentiated the PS A locus of these 50 strains into three major genetic types (Table 5). To determine whether there was even greater heterogeneity that could not be differentiated by these nine PCRs, extended PCR (EPCR) was performed to amplify the entire PS A locus from each strain of the collection. An upstream primer (EPCR-A3) that hybridizes to *upaZ*, the last conserved gene of the 5′ end of the locus, was synthesized, and a downstream primer (EPCR-A4) that hybridizes to *wcfR*, the first conserved gene of the 3′ end of the locus (Fig. 1C), was also synthesized. The EPCR primer pair successfully amplified a PCR product from 49 of the 50 strains. The EPCR product sizes helped to differentiate the PS A loci of the 50 strains into 15 genetic types (Table 5). The EPCR products ranged in size from ~8.2 to ~22 kb. The 22-kb product was amplified exclusively from *B. fragilis* 638R, indicating that—at least with respect to the PS A locus—this frequently used laboratory strain is not a particularly representative isolate with regard to its PS A type.

Many studies, using a variety of genetic methods such as DNA-DNA hybridization, restriction fragment length polymorphism analysis, multilocus enzyme electrophoresis, and ribosomal DNA sequencing, have demonstrated that *B. fragilis* comprises two genotypically distinct groups (11, 14, 24, 25). Among other characteristics, these two groups are differentiated by the presence or absence of the *cfiA* gene, encoding a metallo-β-lactamase. Each of the four *cfiA*$^+$ *B. fragilis* strains used in this study (3636, 127746, 1281550II, and TAL2480), while demonstrating conservation of the regions flanking the PS A locus, nevertheless segregated into distinct PS A genetic types: EPCR failed to amplify a product from 3636, and EPCR products of 12.5, 13.0, and 14.8 kb were produced from 1277476, 1281550II, and TAL2480, respectively (Table 5).

The 9343 EPCR product was ~8.2 kb as expected, and a similar-sized product was amplified from 14 additional isolates. In fact, all nine PCRs produced products from these 14 strains. These results suggest that these 14 strains have the same PS A locus as 9343. Thirteen of these 14 strains react with CE3, a MAb reactive to the PS A of 9343. None of the other 35 strains reacted with CE3. These data suggest that approximately one-quarter of *B. fragilis* strains synthesize a PS A that is, by the methods used in this study, genetically and immunologically indistinguishable from the PS A of 9343.

TABLE 4. Abscessogenic potential of *B. fragilis* 9343 and 9343ΔPSA

Strain	No. of mice with abscesses/no. tested at a dose of:					AD_{50} (log_{10})	P^a
	10^8	10^7	10^6	10^5	10^4		
9343	17/20	13/15	12/15	7/15	5/15	5.2	
9343ΔPSA	3/20	5/15	4/15	2/15	1/15	14.9	<0.0001

[a] Compared with the value for *B. fragilis* 9343.

DISCUSSION

The mechanism by which displacement of *B. fragilis* into the peritoneal cavity results in the formation of abscesses is a subject of intense scrutiny in our laboratory. Though much work has been directed at identifying the biochemical, structural, and immunogenic features of the capsular polysaccharides that account for the unusual propensity of *B. fragilis* infection to culminate in the formation of abscesses, comparatively few studies have addressed the genetic features underlying this tendency.

The PS A locus is the third capsular polysaccharide biosynthesis region sequenced from strain 9343. Genes involved in conferring charges to the capsules are present in all three of

TABLE 5. PCR amplification of regions homologous to the 9343 PS A region from various *B. fragilis* strains

Strain	Result of PCR[a]										
	1 (*orf1-upaY*)	2 (*orf1-upaZ*)	3 (*wzx-wcfN*)	4 (*wzy-wcfP*)	3A (*wcfP-pdiA*)	2A (*wcfQ-pdiA*)	8 (*wcfR-pdiA*)	6A (*wcfS-pdiA*)	7 (*pdiA-mtfA*)	EPCR product size (kb)	CE3 reactivity
9343	+	+	+	+	+	+	+	+	+	8.2	+
B16	+	+	+	+	+	+	+	+	+	8.2	+
US324	+	+	+	+	+	+	+	+	+	8.2	+
US1206	+	+	+	+	+	+	+	+	+	8.2	+
US390	+	+	+	+	+	+	+	+	+	8.2	+
US2244	+	+	+	+	+	+	+	+	+	8.2	+
B35	+	+	+	+	+	+	+	+	+	8.2	+
B110	+	+	+	+	+	+	+	+	+	8.2	+
B124	+	+	+	+	+	+	+	+	+	8.2	+
CMR2896	+	+	+	+	+	+	+	+	+	8.2	+
US398	+	+	+	+	+	+	+	+	+	8.2	+
B272	+	+	+	+	+	+	+	+	+	8.2	+
1284-2	+	+	+	+	+	+	+	+	+	8.2	+
IL89375II	+	+	+	+	+	+	+	+	+	8.2	+
1287245I	+	+	+	+	+	+	+	+	+	8.2	
13141	+	+	+	+	+	+	+	+	+	13.0	
CM12	+	+				+	+	+	+	13.0	
1281550II	+	+				+	+	+	+	13.0	
379	+	+				+	+	+	+	14.8	
38310	+	+				+	+	+	+	14.8	
B68	+	+				+	+	+	+	14.8	
CM3	+	+				+	+	+	+	14.8	
1283931I	+	+				+	+	+	+	14.8	
12775LII	+	+				+	+	+	+	14.8	
1285531I	+	+				+	+	+	+	14.8	
B117	+	+				+	+	+	+	14.8	
B356772I	+	+				+	+	+	+	14.8	
TAL2480	+	+				+	+	+	+	14.8	
3636	+	+				+	+	+	+	F[b]	
1291662III	+	+					+	+	+	9.5	
1277476	+	+					+	+	+	12.5	
12905-23V	+	+					+	+	+	13.0	
2429	+	+					+	+	+	13.0	
US365	+	+					+	+	+	15.2	
US357	+	+					+	+	+	15.2	
17905	+	+					+	+	+	15.2	
44355	+	+					+	+	+	15.8	
45703	+	+					+	+	+	15.8	
CM11	+	+					+	+	+	15.8	
US326	+	+					+	+	+	15.8	
26877	+	+					+	+	+	15.8	
1281262I	+						+	+	+	17.0	
1279-2	+	+					+	+	+	17.0	
1279155I	+	+					+	+	+	20.0	
US388	+	+					+	+	+	20.0	
US52540	+	+					+	+	+	20.0	
PA5	+	+					+	+	+	20.0	
US332	+	+					+	+	+	20.0	
1277810I	+	+					+	+	+	20.0	
638R	+	+					+	+	+	22.0	

[a] A plus sign indicates that a product was successfully amplified.
[b] EPCR failed to produce a product from this strain.

these loci (6, 9). Recent data demonstrated that homologs of genes whose products confer these charge groups are present in all *B. fragilis* strains tested and suggest that the synthesis of capsular polysaccharides with charged groups is a characteristic of the species *B. fragilis* (7).

The synthesis of at least three capsular polysaccharides by *B. fragilis* and the demonstration that purified PS A and PS B each induce abscesses in the animal model suggested that more than one of the capsular polysaccharides were contributing to the organism's ability to induce abscesses. If so, a mutant devoid of multiple capsular polysaccharides might be necessary to produce an attenuated strain. The first mutants created, 9343Δ*wcfF* and 9343ΔPSB, which are unable to produce PS C and PS B, respectively, supported this contention. These mutants were each found to be fully virulent, with AD_{50}s statistically indistinguishable from that of wild-type 9343 (6, 9). The data presented here demonstrate that the capsular polysaccharides of 9343 do not contribute equally to abscess induction;

rather, the synthesis of PS A is essential for *B. fragilis* 9343 to facilitate abscess formation (Table 4). The 9343ΔPSA mutant strain had an AD_{50} that was extrapolated to be approximately 10 log units higher than the wild-type strain. These data definitively demonstrate that PS A is the major virulence determinant in the formation of abscesses by *B. fragilis* 9343.

PCR studies and hybridization analysis of the PS C locus of 50 *B. fragilis* strains demonstrated that no polysaccharide biosynthesis gene was conserved in the PS C locus of all strains, although the flanking DNA was conserved (7). Therefore, the presence of *wcfR* and *wcfS* in the PS A locus of all strains tested was an unexpected finding. Our conclusion that WcfR and WcfS are involved in the synthesis and transfer of the acetamido-amino-2,4,6-trideoxygalactose (AATGal) of PS A is strongly supported by homology data. Close homologs of both of these proteins are encoded by the region responsible for the synthesis of the O17 O antigen of *Plesiomonas shigelloides* and by the *Shigella sonnei* form I antigen locus (4) (Table 3). Both of these polysaccharides contain the rare monosaccharide AATGal (15). Homologs of WcfR and WcfS were also detected in several species for which the polysaccharide structures have not been determined. With few exceptions, the WcfR and WcfS homologs are encoded by adjacent genes, suggesting that they participate in a common pathway, as predicted for WcfR and WcfS. The teichoic acid polymer (C substance) of *S. pneumoniae* also contains AATGal (2, 10). Based on the putative functions of WcfR and WcfS in AATGal biosynthesis and transfer, homologs of these genes should be present in the genome of *S. pneumoniae*. To test our hypothesis, a search of the unfinished *S. pneumoniae* type 4 genome (available from www.tigr.org) using the TBLASTN program with WcfR and WcfS as query sequences was performed. This search revealed two adjacent ORFs encoding putative products with similarity to WcfR and WcfS, as predicted (Table 3). These genes are not flanked by other polysaccharide biosynthesis genes and therefore do not appear to be contained in a capsular polysaccharide biosynthesis locus. Since the *S. pneumoniae* type 4 capsule repeating unit does not contain AATGal (13), we predict that these genes are involved in the synthesis and transfer of the AATGal of the *S. pneumoniae* teichoic acid. As this polymer is common to this species, these genes are likely contained in the genome of all *S. pneumoniae* strains, regardless of capsular serotype.

The importance of the free amino group of AATGal to the ability of PS A to cause abscesses is well documented (29–31). The conservation of *wcfR* and *wcfS* in the PS A locus of *B. fragilis* indicates that this unusual monosaccharide is likely present in the PS A repeating unit of this species as a whole. The only other *B. fragilis* strain for which the structure of the PS A repeating unit has been determined is 638R. We have shown that the PS A locus of 638R is ~24 kb, compared with the ~10.7-kb PS A locus of 9343 (Table 5). As predicted by this large locus, the 638R PS A repeating unit (termed PS A2) is more complex than the PS A of 9343. PS A2 contains five monosaccharides (compared with the four residues in the repeating unit of 9343 PS A [1]), and several of these are deoxymonosaccharides (32). Despite the great structural differences between PS A and PS A2, PS A2 also contains AATGal, as the results presented here predict it should. The possibility that this sugar is conserved throughout the species, as the genetic analyses suggest, has profound biological and pathogenic implications.

ACKNOWLEDGMENTS

We are grateful to Jessica Kadis for help with mutant screening and to Jaylyn Olivo for editorial services.

This work was supported by Public Health Service grants AI44193 and AI39576 from the National Institute of Allergy and Infectious Diseases and by the Edward and Amalie Kass Fellowship.

REFERENCES

1. **Baumann, H., A. O. Tzianabos, J. R. Brisson, D. L. Kasper, and H. J. Jennings.** 1992. Structural elucidation of two capsular polysaccharides from one strain of *Bacteroides fragilis* using resolution NMR spectroscopy. Biochemistry **31:**4081–4089.
2. **Behr, T., W. Fischer, J. Peter-Katalinic, and H. Egge.** 1992. The structure of pneumococcal lipoteichoic acid. Improved preparation, chemical and mass spectrometric studies. Eur. J. Biochem. **207:**1063–1075.
3. **Bethesda Research Laboratories.** 1986. BRL pUC host: *E. coli* DH5α competent cells. Focus **8:**9.
4. **Chida, T., N. Okamura, K. Ohtani, Y. Yoshida, E. Arakawa, and H. Watanabe.** 2000. The complete DNA sequence of the O antigen gene region of *Plesiomonas shigelloides* serotype O17 which is identical to *Shigella sonnei* form I antigen. Microbiol. Immunol. **44:**161–172.
5. **Comstock, L. E., M. J. Coyne, A. O. Tzianabos, and D. L. Kasper.** 1999. Interstrain variation of the polysaccharide B biosynthesis locus of *Bacteroides fragilis*: characterization of the region from strain 638R. J. Bacteriol. **181:** 6192–6196.
6. **Comstock, L. E., M. J. Coyne, A. O. Tzianabos, A. Pantosti, A. B. Onderdonk, and D. L. Kasper.** 1999. Analysis of a capsular polysaccharide biosynthesis locus of *Bacteroides fragilis.* Infect. Immun. **67:**3525–3532.
7. **Comstock, L. E., A. Pantosti, and D. L. Kasper.** 2000. Genetic diversity of the capsular polysaccharide C biosynthesis region of *Bacteroides fragilis*. Infect. Immun. **68:**6182–6188.
8. **Cox, D., and E. Snell.** 1989. Analysis of binary data, 2nd ed. Chapman and Hall, London, United Kingdom.
9. **Coyne, M. J., W. Kalka-Moll, A. O. Tzianabos, D. L. Kasper, and L. E. Comstock.** 2000. *Bacteroides fragilis* NCTC9343 produces at least three distinct capsular polysaccharides: cloning, characterization, and reassignment of the PS B and PS C biosynthesis loci. Infect. Immun. **68:**6176–6181.
10. **Fischer, W., T. Behr, R. Hartmann, J. Peter-Katalinic, and H. Egge.** 1993. Teichoic acid and lipoteichoic acid of *Streptococcus pneumoniae* possess identical chain structures. A reinvestigation of teichoid acid (C polysaccharide). Eur. J. Biochem. **215:**851–857.
11. **Gutacker, M., C. Valsangiacomo, and J. Piffaretti.** 2000. Identification of two genetic groups in *Bacteroides fragilis* by multilocus enzyme electrophoresis: distribution of antibiotic resistance (*cfiA, cepA*) and enterotoxin (*bft*) encoding genes. Microbiology **146:**1241–1254.
12. **Hohn, B., and J. Collins.** 1980. A small cosmid for efficient cloning of large DNA fragments. Gene **11:**291–298.
13. **Jansson, P. E., B. Lindberg, and V. Lindquist.** 1981. Structural studies of the capsular polysaccharide from *Streptococcus pneumoniae* type 4. Carbohydr. Res. **95:**73–80.
14. **Johnson, J. L.** 1978. Taxonomy of the bacteroides. I. Deoxyribonucleic acid homologies among *Bacteroides fragilis* and other saccharolytic *Bacteroides* species. Int. J. Syst. Bacteriol. **28:**245–268.
15. **Kenne, L., B. Lindberg, K. Peterson, E. Katzenellenbogen, and E. Romanowska.** 1980. Structural studies of the O-specific side-chains of the *Shigella sonnei* phase I lipopolysaccharide. Carbohydr. Res. **78:**119–126.
16. **Lindberg, B., B. Lindqvist, J. Lonngren, and D. A. Powell.** 1980. Structural studies of the capsular polysaccharide from *Streptococcus pneumoniae* type 1. Carbohydr. Res. **78:**111–117.
17. **Liu, D. L., R. A. Cole, and P. R. Reeves.** 1996. An O-antigen processing function of Wzx (RfbX): a promising candidate for O-unit flippase. J. Bacteriol. **178:**2102–2107.
18. **Lutton, D. A., S. Patrick, A. D. Crockard, L. D. Stewart, M. J. Larkin, E. Dermott, and T. A. McNeill.** 1991. Flow cytometric analysis of within-strain variation in polysaccharide expression by *Bacteroides fragilis* by use of murine monoclonal antibodies. J. Med. Microbiol. **35:**229–237.
19. **Morona, R. M., M. Mavris, A. Fallarino, and P. A. Manning.** 1994. Characterization of the *rfc* region of *Shigella flexneri*. J. Bacteriol. **176:**733–747.
20. **Myer, R. J., and J. A. Shapiro.** 1980. Genetic organization of the broad-host-range IncP-1 plasmid R751. J. Bacteriol. **143:**1362–1373.
21. **Nassau, P. M., S. L. Martin, R. E. Brown, A. Weston, D. Monsey, M. R. McNeil, and K. Duncan.** 1996. Galactofuranose biosynthesis in *Escherichia coli* K12: identification and cloning of the UDP-galactopyranose mutase. J. Bacteriol. **178:**1047–1052.
22. **Pantosti, A., R. Colangeli, A. O. Tzianabos, and D. L. Kasper.** 1995. Monoclonal antibodies to detect capsular diversity among *Bacteroides fragilis* iso-

lates. J. Clin. Microbiol. **33:**2647–2652.
23. **Pantosti, A., A. O. Tzianabos, A. B. Onderdonk, and D. L. Kasper.** 1991. Immunochemical characterization of two surface polysaccharides of *Bacteroides fragilis*. Infect. Immun. **59:**2075–2082.
24. **Podglajen, I., J. Breuil, I. Casin, and E. Collatz.** 1995. Genotypic identification of two groups within the species *Bacteroides fragilis* by ribotyping and by analysis of PCR-generated fragment patterns and insertion sequence content. J. Bacteriol. **177:**5270–5275.
25. **Ruimy, R., I. Podglajen, J. Breuil, R. Christen, and E. Collatz.** 1996. A recent fixation of *cfiA* genes in a monophyletic cluster of *Bacteroides fragilis* is correlated with the presence of multiple insertion elements. J. Bacteriol. **178:**1914–1918.
26. **Shapiro, M. E., D. L. Kasper, D. F. Zaleznik, S. Spriggs, A. B. Onderdonk, and R. W. Finberg.** 1986. Cellular control of abscess formation: role of T cells in the regulation of abscesses formed in response to *Bacteroides fragilis*. J. Immunol. **137:**341–346.
27. **Shoemaker, N. B., C. Getty, E. P. Guthrie, and A. A. Salyers.** 1986. Two *Bacteroides* plasmids, pBFTM10 and pB8–51, contain transfer regions that are recognized by broad-host-range IncP plasmids and by a conjugative *Bacteroides* tetracycline resistance element. J. Bacteriol. **166:**959–965.
28. **Stevens, A. M., N. B. Shoemaker, and A. A. Salyers.** 1990. The region of a *Bacteroides* conjugal chromosomal tetracycline resistance element which is responsible for production of plasmidlike forms from unlinked chromosomal DNA might also be involved in transfer of the element. J. Bacteriol. **172:** 4271–4279.
29. **Tzianabos, A. O., R. W. Finberg, Y. Wang, M. Chan, A. B. Onderdonk, H. J. Jennings, and D. L. Kasper.** 2000. T cells activated by zwitterionic molecules prevent abscesses induced by pathogenic bacteria. J. Biol. Chem. **275:**6733–6740.
30. **Tzianabos, A. O., A. B. Onderdonk, B. Rosner, R. L. Cisneros, and D. L. Kasper.** 1993. Structural features of polysaccharides that induce intra-abdominal abscesses. Science **262:**416–419.
31. **Tzianabos, A. O., A. B. Onderdonk, R. S. Smith, and D. L. Kasper.** 1994. Structure-function relationships for polysaccharide-induced intra-abdominal abscesses. Infect. Immun. **62:**3590–3593.
32. **Wang, Y., W. Kalka-Moll, M. H. Roehrl, and D. L. Kasper.** 2000. Structural basis of the abscess-modulating polysaccharide A2 from *Bacteroides fragilis*. Proc. Natl. Acad. Sci. USA **97:**13478–13483.

Editor: R. N. Moore