JOURNAL OF BACTERIOLOGY, Apr. 2005, p. 2858–2869 Vol. 187, No. 8
0021-9193/05/$08.00+0 doi:10.1128/JB.187.8.2858–2869.2005

Genetic and Structural Analysis of the *Bacteroides* Conjugative Transposon CTn341

M. Bacic, A. C. Parker, J. Stagg, H. P. Whitley, W. G. Wells, L. A. Jacob, and C. J. Smith*

Department of Microbiology and Immunology, Brody School of Medicine, East Carolina University, Greenville, North Carolina

Received 13 October 2004/Accepted 7 January 2005

The genetic structure and functional organization of a *Bacteroides* conjugative transposon (CTn), CTn341, were determined. CTn341 was originally isolated from a tetracycline-resistant clinical isolate of *Bacteroides vulgatus*. The element was 51,993 bp long, which included a 5-bp coupling sequence that linked the transposon ends in the circular form. There were 46 genes, and the corresponding gene products fell into three major functional groups: DNA metabolism, regulation and antibiotic resistance, and conjugation. The G+C content and codon usage observed in the functional groups suggested that the groups belong to different genetic lineages, indicating that CTn341 is a composite, modular element. Mutational analysis of genes representing the different functional groups provided evidence for the gene assignments and showed that the basic conjugation and excision genes are conserved among *Bacteroides* spp. A group IIA1 intron, designated B.f.I1, was found to be inserted into the *bmhA* methylase gene. Reverse transcriptase PCR analysis of CTn341 RNA showed that B.fr.I1 was functional and was spliced out of the *bmhA* gene. Six related CTn-like elements were found in the genome sequences of *Bacteroides fragilis* NCTC9343 and *Bacteroides thetaiotaomicron* VPI5482. The putative elements were similar to CTn341 primarily in the *tra* and *mob* regions and in the *exc* gene, and several appeared to contain intron elements. Our data provide the first reported sequence for a complete *Bacteroides* CTn, and they should be of considerable benefit to further functional and genetic analyses of antibiotic resistance elements and genome evolution in *Bacteroides*.

Conjugation is one of the most important mechanisms of horizontal gene transfer in prokaryotes, leading to genetic variation within a species and the acquisition of new traits, such as antibiotic resistance or the ability to produce toxins (10). Conjugation was originally thought to be exclusively a plasmid-encoded function, but now it also is known to be mediated by a diverse group of specialized genetic elements, collectively referred to as either conjugative transposons (CTns) or integrative and conjugative elements (9, 10, 33). These elements can be broadly defined as self-transmissible, integrating elements that cannot be maintained extrachromosomally and they have now been identified in a wide array of gram-negative and gram-positive species from clinical or environmental sources (9, 33). CTns were first identified for their role in the dissemination of antibiotic resistance. For example, in gram-positive organisms Tn*916* encodes tetracycline resistance and Tn*1546* carries vancomycin resistance determinants (9). In gram-negative bacteria there is the IncJ family of integrating elements and the related element SXT that carry resistance to multiple antibiotics in enteric organisms and *Vibrio cholerae* (3, 6).

The *Bacteroides* CTns were among the first CTns discovered, and they have since been shown to play a central role in the dissemination of antibiotic resistance genes in this genus and related genera (26, 31, 33, 37). This role stems from two novel features of the *Bacteroides* CTns. The first novel feature is the ability to mobilize unlinked plasmids and other specialized transposons (mobilizable transposons [MTns]). The elements being mobilized must encode a specific Mob protein (in the relaxase superfamily) and a *cis oriT* which can make use of the CTn conjugation apparatus for transfer. A wide range of antibiotic resistance plasmids (e.g., pBFTM10, *ermF*) and MTns (e.g., Tn*4555*, *cfxA*) have been shown to transfer in this manner (33). The second novel feature is the ability of most *Bacteroides* CTns to induce conjugation 1,000- to 10,000-fold in the presence of tetracycline. The transfer of both the CTn and unlinked elements is enhanced by growth in the presence tetracycline at a concentration of ≤ 1 μg/ml, and this induction is mediated by a complex signal transduction pathway that includes a two-component regulatory system (RteA and RteB) and at least one other regulatory protein, RteC (25, 43, 44). Together, the ability to mobilize other genetic elements and the antibiotic-stimulated transfer may in part explain the 80% tetracycline resistance and the high frequency of other antibiotic resistance genes in *Bacteroides* (37).

There is a paucity of information on the mechanisms of *Bacteroides* CTn integration, excision, and conjugation, but a general framework is emerging. Prior to conjugation the elements excise from the chromosome and circularize by using a process that requires multiple gene products and appears to be unlike any known excision mechanism (11, 13). The mechanism of conjugation is even less well understood, and only 10% of the genes in the transfer regions of CTnDOT appear to match anything in the public databases. One exception, TraG, has similarity to VirB4, which is a superfamily of proteins involved in DNA transfer in all type IV secretion systems (7). It is assumed that conjugation involves the transfer of a single

* Corresponding author. Mailing address: Department of Microbiology and Immunology, 600 Moye Blvd., East Carolina University, Greenville, NC 27834. Phone: (252) 744-2700. Fax: (252) 744-3104. E-mail: smithcha@mail.ecu.edu.

TABLE 1. Bacterial strains and plasmids

Strain or plasmid	Relevant properties[a]	Reference or source
Strains		
B. thetaiotaomicron BT5482	Rf^r, no plasmids or known transmissible elements	4
B. thetaiotaomicron IB394	Rf^r Tp^r, Δ*thyA* derivative of BT5482	This study
B. thetaiotaomicron BTΩpFD699	Rf^r Tc^r Em^r, independent transconjugant from EC100(pFD699) × BT5482 mating	41
B. thetaiotaomicron IB395	Rf^r Tc^r Δ*thyA*, independent transconjugant from *B. vulgatus* CLA341 × IB394 mating	This study
B. thetaiotaomicron IB399	Rf^r Tc^r Em^r Δ*thyA*, independent transconjugant from EC100(pFD699) × IB394 mating	This study
B. fragilis 638R	Rf^r, no plasmids or known transmissible elements	31
B. vulgatus CLA341	Tc^r Fx^r, clinical strain containing CTn341 and Tn*4555*	29
E. coli EC100	F^- *mcrA* Δ(*mrr-hsdRMS-mcrBC*) ϕ80d*lacZ* ΔM15 Δ*lacX74 recA1 endA1 araD139* Δ(*ara leu*)*7697 gaIU gaIK* λ^- *rpsLnupG*	Epicentre
E. coli DH10B	F^- *mcrA* Δ(*mrr-hsdRMS-mcrBC*) ϕ80*lacZ* ΔM15 Δ*lacX74 deoR recA1 endA1 araD139* Δ(*ara leu*)*7697 gaIU gaIK* λ^-*rpsL nupG tonA*	Invitrogen
Plasmids		
pFD699	Em^r Tc^r (Ap^r), pFD670::CTn341 hybrid element that replicates as plasmid in *E. coli* and integrates as CTn in *Bacteroides*, 56 kb	41
pYT102	Tc^r (Cm^r), suicide plasmid containing *thyA*, 8.2 kb	2
pYT102:Fx	Fx^r (Cm^r), derivative of pYT102 in which *tetQ* was replaced by *cfxA*, 8 kb	This study

[a] The antibiotic resistance markers in parentheses Ap^r and Cm^r, are expressed only in *E. coli* strains; Rf^r, Tp^r, Tc^r, Em^r, and Fx^r determinants are expressed only in *Bacteroides* strains.

DNA strand which subsequently undergoes replication in the recipient and that then a tyrosine integrase belonging to the phage lambda superfamily mediates integration by a mechanism similar to that described for Tn*916* (12, 13). Currently there is a wide gap in our knowledge of most transposon gene products and their functions in *Bacteroides* CTns.

In this paper we describe for the first time the complete sequence of a *Bacteroides* CTn. The element sequenced, CTn341, is a member of the most common CTn group, the CTnDOT family (33). Sequence analysis revealed 46 genes and one functional group II intron. The genes fell roughly into three major functional groups: DNA metabolism, regulation and antibiotic resistance, and conjugation. Mutational analyses of genes from each group were used to verify the proposed functional assignments. Based on G+C content and codon usage, the functional groups appeared to belong to different genetic lineages, indicating that CTn341 is a composite, modular element. Comparisons with *Bacteroides* genome sequences suggested that the basic conjugation and excision genes are conserved in *Bacteroides* spp. Our data should be of considerable benefit to further functional and genetic analyses of these important antibiotic resistance elements.

MATERIALS AND METHODS

Bacterial strains and growth. The relevant bacterial strains and plasmids used in this study are described in Table 1. *Bacteroides* strains were grown in BHIS medium in an anaerobic chamber with an atmosphere containing 80% N_2, 10% CO_2, and 10% H_2 (38). *Escherichia coli* strains were grown aerobically in L-broth (1% tryptone, 0.5% NaCl, 0.5% yeast extract). The following antibiotics were used for selection: rifampin (20 μg/ml), tetracycline (5 μg/ml), ampicillin (50 μg/ml), erythromycin (10 μg/ml), chloramphenicol (10 μg/ml), cefoxitin (20 μg/ml), and gentamicin (50 μg/ml).

Routine DNA manipulations. Standard procedures were used for agarose gel electrophoresis of DNA, restriction endonuclease digestion, and DNA ligation reactions (34). Large-scale plasmid samples for sequencing and in vitro mutagenesis were prepared by CsCl-ethidium bromide equilibrium ultracentrifugation of lysates prepared by a gentle high-salt lysis procedure or by purification on QIAGEN Maxi columns by using protocols for low-copy-number plasmids (QIAGEN, Valencia, Calif.) (38). Chromosomal DNA from *Bacteroides* species was isolated from total cell lysates by a previously described method (39). Southern blot analysis of chromosomal DNA was carried out as described previously by using nylon membranes probed with biotinylated probes (41). Following overnight hybridization at 65°C, the membranes were washed under stringent conditions and then exposed to X-ray film for autoradiography. The probe (217 bp) encompassed both ends of CTn341 and was PCR amplified from pFD699 with primers 5′GGATTGTAATGGCTTCACGGA3′ and 5′CGCACAGAAGAATA ATGCAGC3′.

Bacterial mating. Standard filter mating protocols were used for plasmid transfer experiments (23, 30). Aerobic mating was used for transfer from *E. coli* to *Bacteroides*, and anaerobic mating was used for all mating with *Bacteroides* donors. Mating with *Bacteroides* donors was performed by using cultures that were either induced with 1 μg of tetracycline per ml or not induced. For mating with IB395 isogenic donors we used *Bacteroides fragilis* IB220 (40) as the recipient, and for mating with IB399 isogenic donors (including Jo strains) we used *E. coli* DH10B as the recipient.

RNA extraction and cDNA preparation. RNA was extracted from log-phase cultures by the hot acid-phenol method (32). RNA was treated with DNase I (Invitrogen, Carlsbad, Calif.) and then used in reverse transcriptase PCR with ImProm II reverse transcriptase (Promega, Madison, Wis.). RNAguard (Amersham Biosciences, Piscataway, N.J.), an RNase inhibitor, was added to all reaction mixtures. The primer used to create cDNA of the intron was 5′-CGCAAT CGTAGTCGTTGTTC-′3, which occurred 200 bp downstream of the intron within the methylase gene, *bmhA1*. The control primer used for *tetQ* cDNA was 5′-GGATATTATCAGAATAACCG-′3. The cDNA produced from these reactions then was amplified with primers Intron/For (5′-CGATGTATAACCGCA AGTTC-′3) and Intron/Rev (5′-GCTTTCAGCCCGATAATCAT-′3). Intron/For was 160 bp upstream of the intron start site, and Intron/Rev was 70 bp after the end of the intron.

DNA sequencing. Plasmid pFD699 (containing CTn341) and mutagenized derivatives of this plasmid were purified from *E. coli* and used as templates in automated DNA sequencing reactions. The pFD699 template was randomly mutagenized in vitro with a modified mini-Mu phage containing a chloramphenicol resistance cassette (Template Generation System; Finnzymes, Espoo, Finland). The mutagenized pFD699 was transformed into *E. coli* EC100, and then individual Cm^r and Ap^r colonies were screened for plasmid content. Plasmid DNA from each randomly chosen colony was used as a template in two sequencing reactions performed with primers orientated outward from the Mu transposon termini in opposite directions. The two primer sequences used were 5′AT CAGCGGCCGCGATCC3′ and 5′TTTATTCGGTCGAAAAGGATCC3′, and at least 500 bp of high-quality sequence was generated in each reaction.

Nucleotide sequences were aligned by using the Fragment Assembly program

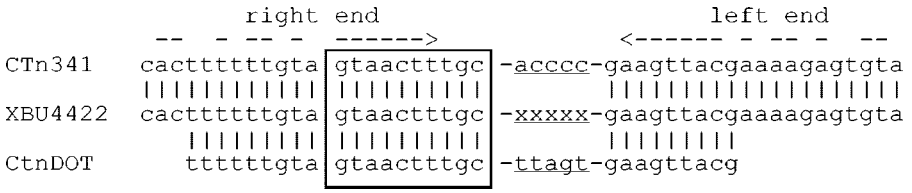

FIG. 1. DNA sequence alignment of the CTn341, XBU4422, and CTnDOT ends, as found in the closed circular intermediate. The 5-bp coupling sequence also is shown where it is known. The conserved 10-bp region of homology between the transposon right end and the insertion target site is enclosed in a box, and the imperfect indirect repeats at the ends are indicated by the dashed arrows above the sequence. xxxxx represents the bases of an unknown coupling sequence. The sequence of XBU4422 was obtained from GenBank accession numbers S75303 and S75304, and the CTnDOT sequence was obtained from reference 13.

from the Wisconsin Package (Accelrys Inc., San Diego, Calif.), and gaps in the sequence were filled in by a primer walking strategy. Additional sequencing reactions were needed to resolve the sequence around the N terminus of the CTn341 *tetQ* gene, which was duplicated during the construction of pFD699 (41). In order to obtain an unaltered sequence from this region, a portion of the *tetQ* gene and upstream region was PCR amplified from genomic DNA of the original CTn341 host strain, *Bacteroides vulgatus* CLA341. This region, which corresponded to bp 34942 to 35859, was cloned into pGEM-T (Promega), and then the nucleotide sequence was determined for both strands. All sequencing was performed at the Molecular Biology Resource Facility, University of Tennessee, Knoxville, by using ABI Prism Dye terminator chemistry and an ABI model 3100 genetic analyzer.

DNA sequence analysis. The Wisconsin Package was used for all routine DNA sequence analyses (18). Identification of protein-encoding genes was done by first using the program Frames to find open reading frames (ORFs) more than 240 bp long. These ORFs were then analyzed by using the program TestCode, which plotted a measure of the nonrandomness of the codon choices along the DNA strand and produced a statistical threshold above which there was a high degree of confidence that the ORF was a gene (21). A second program, Codonpreference, was used to confirm the gene assignments by comparing the codon usage to a codon frequency table that we constructed from known *Bacteroides* genes.

Preliminary results of National Center for Biotechnology Information BLAST searches suggested that there were four regions of the *Bacteroides thetaiotaomicron* VPI5482 genome (accession number NC_004663) and two regions of the *B. fragilis* NCTC9343 genome (http://www.sanger.ac.uk/Projects/B_fragilis/) that had extensive homology with CTn341. This homology was examined more closely at the nucleotide sequence level by using BlastZ, which is a modified Gapped BLAST program that aligns sequences in the megabase range (35, 36). The program generated a percent identity plot, and MultiPip Maker showed the relationships among more than two sequences. The programs were accessed at http://bio.cse.psu.edu/pipmaker/. The output showed aligned regions (55 to 70% identity over >100 bp without a gap) and strongly aligned regions (>70% identity over >100 bp without a gap). For these comparisons the following sequences were used for *B. thetaiotaomicron*: BT-CTn1 (bp 63000 to 133000), BT-CTn2 (bp 2850000 to 2930000), BT-CTn3 (bp 3214000 to 3308000), and BT-CTn4 (bp 6210000 to 6260361 and 1 to 45000). The *B. fragilis* sequences used were BF-CTn1 (bp 1463022 to 1537981) and BF-CTn2 (bp 2015029 to 2080000).

Mutant construction. CTn341 mutant strains containing the mini-Mu insertions are described above, and all of them were derived from pFD699 and given Jo designations with identifying numbers. Other mutations were constructed by using the double-crossover positive-selection technique described by Baughn and Malamy (2). For these mutants, IB395 and IB399 were the *thyA* host strains, and either pYT102 or pYT102:FX was the vector used for allelic exchange. The deletion mutations were designed so that the reading frame was preserved and the deleted sequences were replaced by a restriction site (SalI or PstI) that was used to aid in their construction. The following deletions were created: Δ*traA* (strain 395-TA; bp 18111 to 18710 deleted), Δ*traD* (strain 395-TD; bp 16518 to 16925 deleted), Δ*traJ* (strain 399-TJ; bp 11298 to 12029 deleted), Δ*traO* (strain 399-TO; bp 7160 to 7624 deleted), Δ*traQ* (strain 399-TQ; bp 5772 to 6206 deleted), and Δ*mobA* (strain 395-MA; bp 19558 to 19911 deleted).

Nucleotide sequence accession number. The final CTn341 sequence has been deposited in the National Center for Biotechnology Information database under accession number AY515263.

RESULTS AND DISCUSSION

Determination of the CTn341 DNA sequence. Previously, CTn341 was captured in *E. coli* on a low-copy-number plasmid by using a simple strategy in which an *E. coli* plasmid bearing a portion of the CTn341 *tetQ* gene was inserted by single-crossover, homologous recombination into CTn341 (41). The resulting construct, pFD699, is a stable plasmid in *E. coli*, and by using RK231 it can be mobilized from *E. coli* into *Bacteroides* hosts, where it does not replicate as a plasmid but is a fully functional conjugative transposon. In order to determine the CTn341 nucleotide sequence, pFD699 was mutagenized with a mini-Mu phage, and randomly chosen clones were sequenced by using primers oriented outward from both ends of Mu. Many regions of CTn341 were underrepresented by Mu insertions; thus, a primer walking strategy was used to provide coverage of those regions and to close gaps in the sequence between Mu insertions.

A total of 493 sequencing reactions were used to compile the final sequence, and the coverage was more-than-fivefold redundant. The sequence was 51,993 bp long and had a G+C content of 49.5%. When pFD699 is transferred into *E. coli* from *Bacteroides*, it is locked in the circular form, so in order to define the ends of CTn341, *attL* and *attR* were identified by homology to the ends of *Bacteroides* conjugative transposons XBU4422 and CTnDOT, which have been described previously (Fig. 1) (4, 13). The ends are imperfect, indirect 18-bp repeats which in *attL* include a conserved region that is homologous to the preferred chromosomal target sites of CTnDOT (13). Interestingly, CTn341 has a second copy of this conserved region located at bp 19236, precisely between the divergently transcribed *tra* and *mob* genes. The ends of all three elements were identical in the 18-bp indirect repeat region, and a comparison to XBU4422 revealed 92% identity to CTn341 for at least 184 bp on the left end and 77% identity for 207 bp on the right end. The CTn341 ends were joined by a 5-bp coupling sequence, which is consistent with the current model of transposition suggested for CTnDOT (13).

Another notable feature of the sequence was the lack of sites for the common restriction enzymes. For example, there were just two EcoRI sites and no HindIII sites when the predicted number based on mononucleotide probabilities was 12 each for a molecule that was the size of CTn341 and had the G+C content of CTn341. Also, there were no sites for the *B. fragilis* restriction enzyme BfaI (CTAG) and only 44 Sau3A1 sites when both enzymes were predicted to have 200 sites.

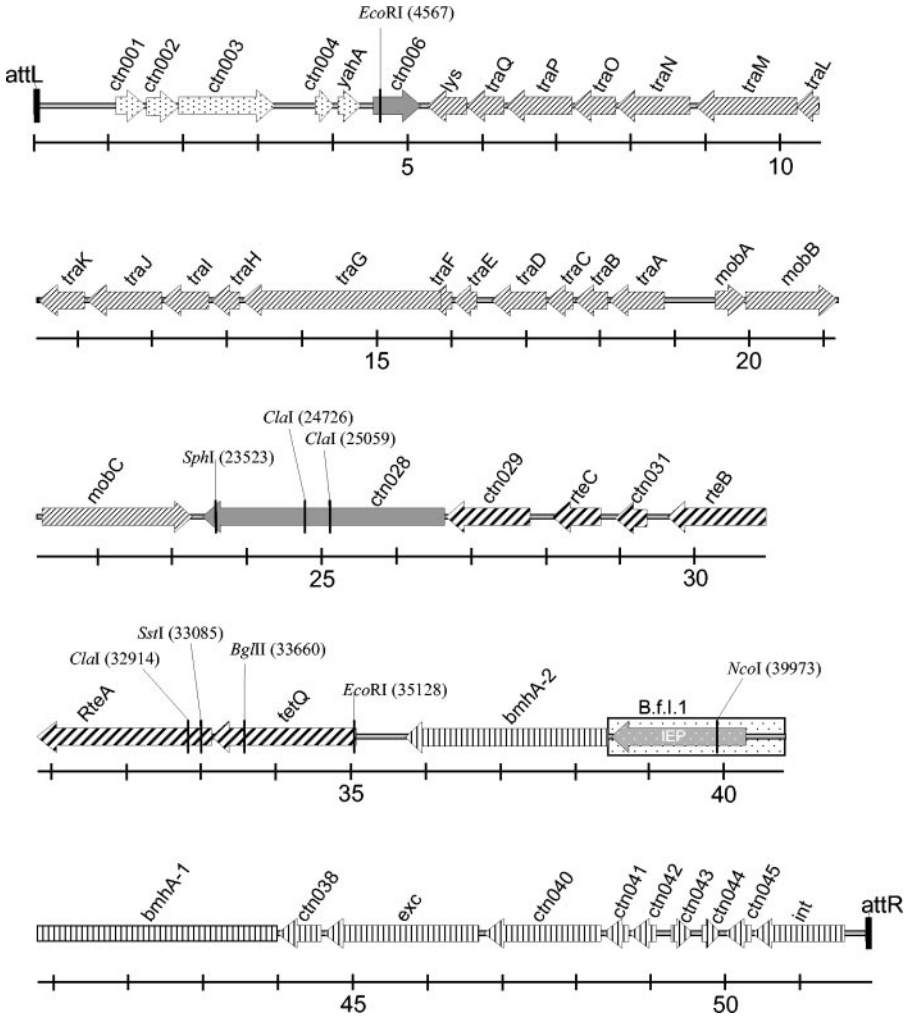

FIG. 2. Genetic organization and ORF map of CTn341. The ORFs are indicated by arrows, which show the orientation of transcription. Arrows with light diagonal lines, transfer region; arrows with dark diagonal lines, regulation and resistance; arrows with vertical lines, DNA metabolism; arrows with stippling, UF-A region with unknown function; dark grey arrows, UF-B region with unknown function. The group II intron B.f.I.1 is indicated by a light stippled box, and the associated intron-encoded protein (maturase) is indicated by the labeled grey arrow.

CTn341 gene assignments. Forty-six ORFs were identified that met the criteria described above (Fig. 2 and Table 2). The initial screening criteria were ORFs that began with an ATG and had ≥80 codons. Putative genes were then examined with the program Testcode, which plots a measure of nonrandomness of the composition at every third base (Fickett's statistic), and final gene assignments were chosen from the genes that had a majority of the ORF predicted at the 95% confidence level. The ORFs then were searched against the GenBank database by using a BLAST program, and we found that the products of 32 of the 46 ORFs did not match other known proteins or only matched proteins encoded by other *Bacteroides* genetic elements. The genes fell into five groups that could be readily identified by abrupt changes in the pattern of the G+C content (Fig. 3). These groups were further supported based on common function, location, and codon usage (Fig. 3 and Table 3).

(i) DNA metabolism. The first group, consisting of 11 genes, was associated with transposon function and DNA metabolism (Table 2). The average G+C content of these genes was 55%, the genes were located just inside *attR*, and most of the genes were oriented so that the direction of transcription was inward from the end (Fig. 2 and 3). Included in this group were genes for the integrase (*int*) and excision (*exc*), which exhibited 97% identity to CTnDOT genes that have been shown to encode these activities (11, 13). Int is a member of the XerD family of site-specific recombinases related to the lambda phage integrases, which are known to cleave DNA substrates by using staggered cuts. Int is required for both insertion and excision of CTnDOT, and recent evidence showed that it makes a 5-base staggered cut adjacent to the conserved sequence of the chromosomal target sites (12). We found that when we used CTn341 (pJo57 and pJo99) *int* mutant *E. coli* strains as donors for mating, no *B. thetaiotaomicron* transconjugants were observed, which was consistent with a role in transposition. Even though the Int proteins are similar, the integration of CTn341 appeared to be more random than that reported for CTnDOT, which has seven preferred insertion sites (4). As shown in Fig.

TABLE 2. Features of CTn341 open reading frames

ORF	Designation	CDS[a]	Protein mol wt	No. of amino acids[b]	Homolog[c]	% Identity (no. of amino acids)[d]	Motif[e]	Putative function	Organism	Accession no.[f]
ORF-1	*ctn001*	1056–1448f	15,749	131	Hypothetical	80 (71)			*B. thetaiotaomicron*	NP_811196
ORF-2	*ctn002*	1465–1890f	16,386	142						
ORF-3	*ctn003*	1890–3170f	49,482	427	Hypothetical	20 (146)			*Enterococcus faecalis*	NP_817052
ORF-4	*ctn004*	3702–3947f	9,444	82						
ORF-5	*yahA*	4004–4309f	11,864	102	YahA	35 (100)			*E. coli*	BAB12595
ORF-6	*ctn006*	4477–5109f	24,422	211						
ORF-7	*lys*	5188–5712r	20,287	175	Lys	29 (153)	COG3772 phage-related lysozyme (muraminidase	Lysis protein	*B. subtilis*, phage B103	NP_690649
ORF-8	*traQ*	5712–6215r	19,064	168	TraQ	86 (164)		Transfer	*B. thetaiotaomicron*	AAG17842
ORF-9	*traP*	6237–7112	32,943	292	TraP	74 (254)	pfam01807, CHC2 zinc finger	Transfer primase	*B. thetaiotaomicron*	AAG17841
					PrmN1	43 (288)	COG0358 DnaG, DNA primase		*Bacteroides uniformis*	AF238307
ORF-10	*traO*	7124–7696r	21,021	191	TraO	94 (191)		Transfer	*B. thetaiotaomicron*	AAG17840
ORF-11	*traN*	7702–8685r	37,114	328	TraN	94 (329)		Transfer	*B. thetaiotaomicron*	AAG17839
ORF-12	*traM*	8757–10103r	48,847	449	TraM	85 (454)		Transfer	*B. thetaiotaomicron*	AAG17838
ORF-13	*traL*	10087–10392r	11,764	102	TraL	69 (102)		Transfer	*B. thetaiotaomicron*	AAG17837
ORF-14	*traK*	10401–11021	24,028	207	TraK	96 (207)		Transfer	*B. thetaiotaomicron*	AAG17836
ORF-15	*traJ*	11056–12057r	37,164	334	TraJ	90 (305)		Transfer	*B. thetaiotaomicron*	AAG17835
ORF-16	*traI*	12064–12690r	24,168	209	TraI	94 (209)		Transfer	*B. thetaiotaomicron*	AAG17834
ORF-17	*traH*	12720–13106r	14,750	129	TraH	85 (128)		Transfer	*B. thetaiotaomicron*	AAG17833
ORF-18	*traG*	13141–15822r	103,339	894	TraG	95 (834)	COG3451, VirB4, type IV secretory pathway	Transfer	*B. thetaiotaomicron*	AAG17832
ORF-19	*traF*	15642–15971r	12,599	110	TraF	96 (108)		Transfer	*B. thetaiotaomicron*	AAG17831
ORF-20	*traE*	15985–16293r	10,654	103	TraE	65 (105)		Transfer	*B. thetaiotaomicron*	AAG17830
ORF-21	*traD*	16497–17234r	27,970	246	TraD	72 (246)		Transfer	*B. thetaiotaomicron*	AAG17828
ORF-22	*traC*	17234–17593r	13,942	120	TraC	69 (120)		Transfer	*B. thetaiotaomicron*	AAG17828
					TraB	29 (65)		Transfer	*B. thetaiotaomicron*	AAG17827
ORF-23	*traB*	17620–18066r	17,003	149	TraB	79 (149)		Transfer	*B. thetaiotaomicron*	AAG17827
					TraC	30 (63)		Transfer	*B. thetaiotaomicron*	AAG17828
ORF-24	*traA*	18075–18827r	28,277	251	traA	94 (251)	COG1192, Soj, ATPases, chromosome partitioning	Transfer	*B. thetaiotaomicron*	AAG17826
ORF-25	*mobA*	19509–19934f	16,822	142	Hypothetical	73 (138)			*B. thetaiotaomicron*	NP_812927
ORF-26	*mobB*	19916–21160f	45,600	415	MocA-like	93 (415)	pfam03432, relaxase	Transfer and mobilization	*B. thetaiotaomicron*	AAG17843
ORF-27	*mobC*	21194–23209f	77,081	672	Putative Mob	74 (667)	pfam02534, TraG; COG3505, VirD4	Transfer and mobilization	*B. thetaiotaomicron*	AAG17844
ORF-28	*ctn028*	23347–26598r	124,219	1,084	Hypothetical	23 (278)			*Shewanella oneidensis*	NP_716343
ORF-29	*ctn029*	26618–27745r	43,078	376	Hypothetical	28 (376)	COG1373, ATPase COG4565, CitB, response regulator		*Coxiella burnetii*	AAO91368
ORF-30	*rteC*	28050–28700r	25,174	217	RteC	96 (203)		Regulation	*B. thetaiotaomicron*	AAA22922.1
ORF-31	*ctn031*	28894–29313r	16,089	140	Hypothetical	38 (55)			*Synechococcus*	AAN46185
ORF-32	*rteB*	29596–30915r	76,229	675	RteB	98 (641)	COG2204, AtoC, response regulator	Regulation	*B. thetaiotaomicron*	AAA22921
ORF-33	*rteA*	30911–33226r	87,098	772	RteA	92 (771)	COG0642, BaeS signal transduction histidine kinase	Regulation	*B. thetaiotaomicron*	AAA22920
ORF-34	*tetQ*	33229–35151r	72,156	641	TetQ	98 (641)	COG0480, FusA, translation elongation factors (GTPases)	Tetracycline resistance	*B. thetaiotaomicron*	Q00937

ORF-35	*bmhA-2* (exon 2)	35812–38574r	99,481	878	BmhA	65 (859)	COG4646, DNA methylase; COG0553, HepA, DNA/RNA helicases	DNA methylase	*B. thetaiotaomicron*	NP_811239
ORF-36	*matR* (intron-encoded protein)	38564–40360r	70,285	599	MatR	42 (599)	COG3344, retron-type reverse transcriptase; pfam01348, type II, intron maturase	Intron splicing	*Lactococcus*	Q57005
ORF-37	*bmhA-1* (exon1)	40885–44073r	120,370	1,060	BmhA	45 (756)	COG4646, DNA methylase; COG0286, HsdM	DNA methylase	*B. thetaiotaomicron*	NP_813660
ORF-38	*ctn038*	44066–44650r	22,264	195	Hypothetical	100 (195)			*B. thetaiotaomicron*	CAD24288
ORF-39	*exc*	44681–46765r	77,600	695	Exc	95 (695)	COG0550, TopA, topoisomerase	Excision	*B. thetaiotaomicron*	CAD24285
ORF-40	*ctn040*	46829–48397r	59,798	523	Hypothetical	92 (505)			*B. thetaiotaomicron*	CAD24284
ORF-41	*ctn041*	48421–48768r	13,923	116	Hypothetical	93 (116)			*B. theaiotaotmicron*	CAD24283
ORF-42	*ctn042*	48775–49134r	14,057	120	Hypothetical	53 (120)			*B. thetaiotaomicron*	CAD24282
ORF-43	*ctn043*	49336–49632f	11,589	99	Hypothetical	98 (99)			*B. thetaiotaomicron*	CAD24281
ORF-44	*ctn044*	49750–49998f	8,432	83	Hypothetical	88 (43)			*B. thetaiotaomicron*	CAC47936
ORF-45	*ctn045*	50053–50412r	13,612	120	Hypothetical	87 (120)	HTH; COG3943, virulence protein		*B. thetaiotaomicron*	CAC47922
					RhuM	28 (87)			*Salmonella*	BAA89634
ORF-46	*int*	50427–51659r	47,885	411	Int	97 (411)	COG4974, XerD, site-specific recombinase	Integration	*B. thetaiotaomicron*	CAC47921

[a] CDS is the coding region of the gene.
[b] Number of amino acids in the predicted protein.
[c] Most similar protein in the databases.
[d] Level of identity of the protein to the homolog for the number of amino acids indicated in parentheses.
[e] Motifs were obtained from the National Center for Biotechnology Information conserved domain database.
[f] Accession number of the most similar protein.

4, chromosomal DNA from independent *Bacteroides* transconjugants was probed with the CTn341 ends to identify the chromosomal junction fragments. For both *B. fragilis* and *B. thetaiotaomicron* the hybridization patterns varied for each of the transconjugants tested, indicating that the insertion mechanism was much more random and suggesting that there are differences in target site selection mechanisms between CTn341 and CTnDOT.

The Exc protein is similar to the TopA family of topoisomerases. These proteins preferentially relax negatively supercoiled DNA, and they have the ability to cleave single- or double-stranded DNA and then rejoin the broken phosphodiester backbone. Although Exc is required for excision of CTnDOT, the topoisomerase (DNA-unwinding) activity of Exc was not required for excision, leaving open the possibility that the cleavage reactions may be more important for excision (46). However, the role of Exc may be even more complex since a recent study showed that the protein is apparently needed for excision in vivo but not in vitro (45). Consistent with this role, CTn341 *exc* mutant Jo97 was able to integrate when it was mobilized from *E. coli* to *Bacteroides* recipients, but it was not able to transfer from *Bacteroides* to *E. coli*, suggesting that excision was impaired (Table 4).

Another DNA metabolism gene, *bmhA* (*Bacteroides* methylase/helicase), was disrupted by a group II intron encoding an intron-encoded protein (IEP) with similarity to reverse transcriptase/maturase proteins found in other group II introns (see below). The complete *bmhA* coding region encoded 1,938 amino acids, and there were three conserved regions: an HsdM-like methylase domain, a prokaryote methylase domain whose function was unknown, and an SNF2 methylase/helicase domain found in many proteins that have activity in transcriptional regulation, recombination, and DNA repair activities. Typical of the SNF2 family, there was a DEXDc motif that is common to ATP-dependent helicases. The remaining seven ORFs in this region did not match any proteins with known functions, although ORF-45 did have a helix-turn-helix motif and was similar to a portion of RhuM found in *Salmonella* pathogenicity island 3. RhuM is thought to be part of a family of virulence proteins, but there is little evidence to support this hypothesis since RhuM is associated with a variable region of the pathogenicity islands (5).

(ii) Conjugative transfer. The second group contained 21 genes that are involved in transfer, and this group likely is phylogenetically related to or coevolved with the DNA metabolism group. The average G+C content of the transfer genes was 53%, and these genes displayed codon usage patterns similar to those of genes in the DNA metabolism group (Table 3). The genes from *traA* to *traQ* were similar to the *tra* region of CTnDOT, and the pairwise levels of identity of the translated products ranged from 69 to 98% (average, 87%). A new gene not previously associated with CTnDOT was *lys*, whose ATG overlapped the *traQ* stop codon. Lys has homology to phage-related lysozyme. Although a mutant is not yet available, it is conceivable that this gene product could participate in the formation of a mating pore which passes through a peptidoglycan layer. Alternatively, this gene could be the remnant of a bacteriophage genome, which leads to intriguing questions regarding the origin of this CTn module.

In general, the *tra* gene coding regions were not similar to

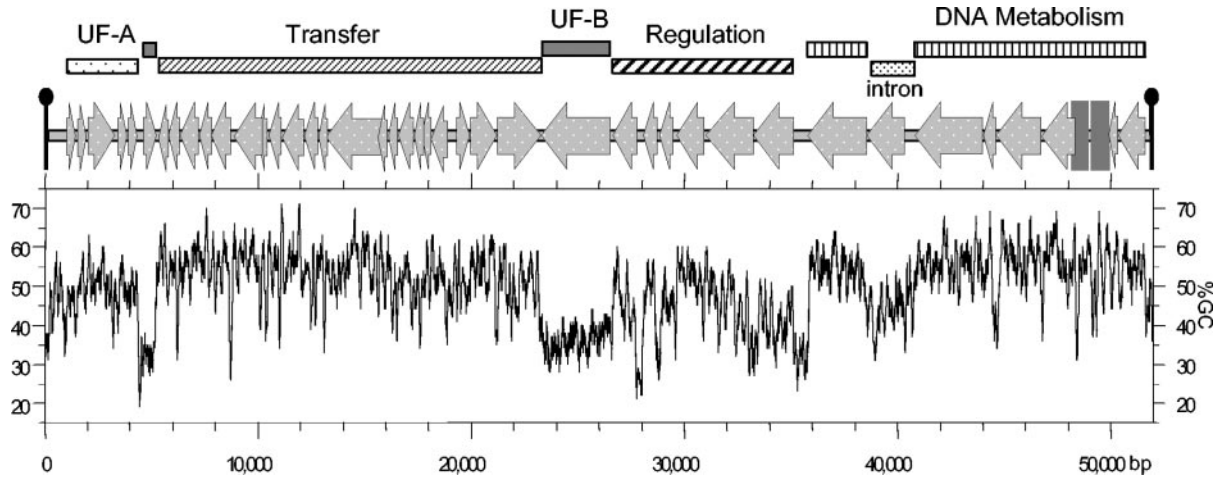

FIG. 3. Plot of G+C content of CTn341 along the length of the molecule. The G+C content was determined by the Window program in the GCG analysis package. *attL* and *attR* are represented at the ends of the map by the stickball symbols. ORFs are represented by arrows except for four small ORFs which are indicated by thick vertical lines. The functional regions of the conjugative transposon are indicated by the bars above the ORF map.

anything in the databases. The three exceptions were TraA, TraG, and TraP. TraA was closely associated with the ParA family of ATPases, which are important for chromosome or plasmid partitioning and often play transcriptional regulatory roles in these processes as well. This family includes MinD and other orthologues, such as ParA-pB171 with the Walker-type ATPase, which are thought to position the plasmid or the chromosome for correct partitioning (17, 20). By analogy, the TraA gene product may be required for recruitment of the CTn to the mating pore. The TraG protein is another ATPase, but it is similar to the family containing VirB4, one of the most conserved proteins found in all type IV secretion pathways (19). In *Bacteroides* another VirB4 orthologue, BctA, was found to be required for transfer of pBF4 (27). VirB4 proteins are membrane associated and are thought to be structural components of the mating pore.

TraP appears to be related to DNA primases in the DnaG family, but only the N-terminal half of the protein is similar. The dominating feature of this motif is the CHCC Zn finger binding domain for single-stranded DNA target recognition. The protein most closely related to TraP in the database is PrmN1, which is required for excision of the mobilizable transposon NBU1. TraP was not required for CTn341 transfer, but it was required for mobilization of an unlinked plasmid (Table 4 and data not shown). It is possible that TraP has some role in DNA processing reactions for unlinked elements that share the mating pore.

TABLE 3. Codon usage comparisons for major CTn341 regions[a]

Region[b]	UF-A	UF-B	Transfer	Regulation	DNA metabolism
UF-A	0.0	3.87	0.71	0.83	0.86
UF-B	3.87	0.0	5.65	2.32	6.00
Tra	0.71	5.65	0.0	1.17	0.10
Reg	0.83	2.23	1.17	0.0	1.24
DNA	0.86	6.00	0.10	1.24	0.0

[a] The values are the D^2 statistic from the GCG Correspond program, which compares codon frequency tables for ORFs. The lower the value, the more similar the codon usage for a given ORF.

[b] UF-A and UF-B are two regions with unknown functions, as shown in Fig. 2 and 3. Tra, transfer; Reg, regulation and antibiotic resistance; DNA, DNA metabolism.

We also included in the conjugation group three genes divergently transcribed from *traA*; *mobA*, *mobB*, and *mobC* appeared to be in an operon but were not previously reported to be part of the CTnDOT *tra* region (Fig. 2). MobB had a relaxase domain similar to that of the RP4 TraI family, and it exhibited homology to other *Bacteroides* mobilization proteins, such as BmgA and MocA, which have been shown to be required for transfer of their cognate mobilizable transposons (1, 28, 42). The *mobC* coding region contained the TraG/VirD4 domain, which is found in coupling proteins involved in DNA processing reactions and is thought to mediate interaction among the relaxasome, the DNA, and the mating pore (19). The TraG/VirD4 family is highly conserved and is found in nearly all type IV systems. The presence of orthologues of both VirB4 and VirD4 suggests that the *Bacteroides* system is related to type IV secretion-conjugation systems found in most other eubacteria but has diverged significantly.

(iii) Antibiotic resistance and regulation. The third module was associated with antibiotic resistance and tetracycline regulation of transposon function. This group was not as well defined with respect to G+C content. The average G+C content was 44%, but the values ranged from 40% for *tetQ* to 51% for *rteB*. The codon usage for this group was distinct from that for all other CTn341 genes, leading to the conclusion that this region may have evolved independently. The *tetQ* gene, which codes for a ribosome protection mechanism of tetracycline resistance, has been described previously and is the primary tetracycline resistance gene found in *Bacteroides* strains (33, 37). There was 97% nucleotide sequence identity to the *tetQ* genes of CTnDOT and pRR1, a tetracycline resistance plasmid found in *Prevotella ruminicola* (37). This homology extended 511 bp upstream into a noncoding region, at which point the CTnDOT sequences diverged. The CTn341 and pRR1 se-

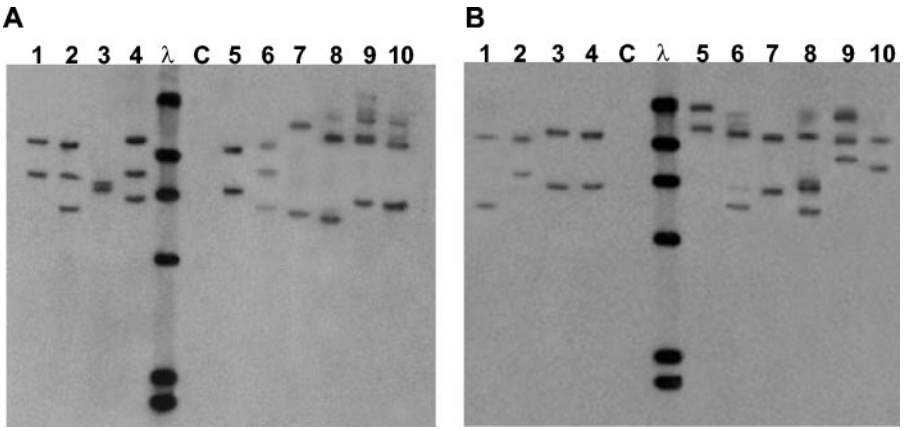

FIG. 4. Analysis of CTn341 insertion sites by Southern hybridization. Chromosomal DNA from independent *B. fragilis* (A) and *B. thetaiotaomicron* (B) transconjugants was digested with AvaII, subjected to Southern blot analysis, and probed with a biotinylated probe containing the CTn341 ends. Lane C contained the parent strain without CTn341; lane λ contained the molecular size standard phage λ DNA digested with HindIII. Lanes 1 through 10 contained independent transconjugants.

quences continued to match for about 250 bp more, extending into the 3′ end of the *bmhA* gene, and then abruptly ended. These data are indicative of extensive gene rearrangements and the modular structure of these elements.

The *rteA*, *rteB*, and *rteC* genes were similar to the genes described previously, and at the amino acid level there was 92 to 98% identity to the CTnDOT sequences (43). RteA and RteB are a two-component regulatory system that works with RteC to control tetracycline-induced excision and transfer, but there is not much additional information concerning the mechanism by which this control works (44). Insertional mutations of *rteA* and *rteB* were both defective for conjugal transfer, as expected, since they are proposed to up-regulate the transfer genes (Table 4). One other gene in this central regulatory region, *ctn029*, had similarity to the AAA general family of ATPases with a conserved P loop. A *ctn029* deletion mutant had no obvious effect on transfer; thus, it is unlikely to play a role in the tetracycline-induced excision and transfer reactions (Table 4).

TABLE 4. Transfer frequencies of CTn341 mutations

Strain	Mutation	Module[a]	Transfer frequency[b]	
			With tetracycline	Without tetracycline
IB399	Wild type		8.8×10^{-4}	1.7×10^{-8}
IB395	Wild type		9.1×10^{-5}	ND[c]
395-TA	Δ*traA*	Con	8.3×10^{-7}	ND
395-TD	Δ*traD*	Con	$<10^{-8}$	ND
Jo37	*traG*	Con	$<10^{-8}$	$<10^{-8}$
Jo90	*traI*	Con	$<10^{-8}$	$<10^{-8}$
399-TJ	Δ*traJ*	Con	1.2×10^{-7}	ND
Jo22	*traK*	Con	$<10^{-8}$	$<10^{-8}$
Jo32	*traM*	Con	$<10^{-8}$	$<10^{-8}$
399-TO	Δ*traO*	Con	2.9×10^{-4}	ND
Jo103	*traP*	Con	1.0×10^{-5}	9.7×10^{-8}
399-TQ	Δ*traQ*	Con	3.1×10^{-3}	ND
395-MA	Δ*mobA*	Con	$<10^{-8}$	ND
Jo36	*mobB*	Con	$<10^{-8}$	$<10^{-8}$
Jo26	*mobC*	Con	6.2×10^{-5}	$<10^{-8}$
Jo68	*rteA*	Reg	$<10^{-8}$	$<10^{-8}$
Jo69	*rteB*	Reg	$<10^{-8}$	$<10^{-8}$
395R29	*ctn029*	Reg	1.1×10^{-4}	ND
Jo97	*exc*	DNA	$<10^{-8}$	$<10^{-8}$
Jo3	*bmhA1*	DNA	8.7×10^{-5}	1.0×10^{-8}
Jo70	*bmhA2*	DNA	6.6×10^{-7}	$<10^{-8}$
Jo49	*iep*	DNA	2.5×10^{-5}	$<10^{-8}$

[a] The modules are the functional regions shown in Fig. 2 and 3. Con, conjugation; Reg, regulation and antibiotic resistance; DNA, DNA metabolism.
[b] Transfer frequencies were determined by determining the number of transconjugants per input donor cell and are means of three replicates. Cultures were either induced with tetracycline as indicated in text or not induced.
[c] ND, not determined.

(iv) Regions with unknown functions. Two regions, *ctn006* and *ctn028*, were notable for their distinctly different G+C contents (32 to 35%) (Fig. 2). In both cases there was a very sharp border between regions, and in fact it appeared that the regions were once one region that was split by insertion of the *tra* genes. The putative genes *ctn006* and *ctn028* did not encode any sequences that matched sequences in the protein databases, and the Testcode predictions for these genes were not convincing, with much of the ORFs falling below the 95% confidence limit.

Another set of unknown genes was located at the left end of the element and was oriented so that transcription was inward. The G+C content of these genes was 48%, and the codon usage did not specifically distinguish these genes as members of any other group of CTn341 genes. None of the putative gene products matched known proteins in the public databases.

Mutational analysis of the *tra* region. As mentioned above, it appears that the transfer region encodes a new type IV secretion system that bears little resemblance to better-studied systems, such as the *Agrobacterium* Vir system or the enteric Rp4 broad-host-range system. In order to confirm the role in transfer, we examined Mu insertional mutations in pFD699 and CTn341 deletion mutations. First we examined the three-gene operon that encoded the Mob proteins which should be important for DNA processing. Conjugation mediated by type IV secretion requires that the relaxase nick a single DNA strand at *oriT*, covalently attach to the 5′ end of the DNA, and then interact with the mating pore. The type IV coupling protein also is thought to be essential for this process and specifically recruits substrate DNA-protein complexes to the mating

pore (19). Consistent with this role, mutants with insertions in *mobB* (relaxase) were unable to transfer CTn341 in mating assays (Table 4). This result was similar to findings with the MTn Tn*4399*, in which the related protein MocA was required for specific *oriT* nicking and subsequent mobilization (28). In contrast to this result, *mobC* mutations were fully functional. Since the coupling protein is absolutely required in other type IV systems, we confirmed the results by testing three independent *mobC* mutations, and none was defective for transfer. There are at least four *mobC* homologues in the *B. thetaiotaomicron* chromosome, and one of these may compensate for the loss. The small gene *mobA* appears to be the first gene in the Mob operon, and thus it seemed likely that it could have a role in transfer. An in-frame deletion mutant was constructed and was found to be deficient in CTn341 transfer. Although MobA did not have any obvious homologues, a position-specific iterated (PSI)-BLAST search did find a convincing match to the Nsp1_C motif associated with nucleoporin p62. This is a component of the nuclear RNA export complex in eukaryotes, and conceivably this motif could be involved in nucleic acid export during conjugation in prokaryotes (16).

As shown in Table 4, we tested 10 of the *tra* genes for a role in transfer. Our results were in general agreement with data reported for the related element CTnDOT, but in addition we identified several new *tra* genes whose functions were required for maximal transfer. Mutations fell into three general classes. The first class resulted in a loss of transfer, and the genes were *traD*, *traG*, *traI*, *traK*, and *traM*. The *traD* and *traK* genes have not been shown previously to be required for transfer, and the results with *traG*, *traI*, and *traM* were similar to the results in previous reports for the homologs in CTnDOT (8). Although little is known about the transcriptional organization of the CTn341 *tra* region, in the case of *traG*, *traI*, *traK*, and *traM* it is possible that the loss of function resulting from the mutations was due to polar effects of the Mu insertions on downstream genes. Deletion mutations in *traA* and *traJ* represented the second class and were the only mutations tested that resulted in a significant decrease in transfer without a complete loss of function. The third class included mutations in *traO*, *traP*, and *traQ*, which resulted in no significant transfer defect or in a slight increase in transfer frequency (Table 4). Previously, Bonheyo et al. suggested that *traQ* and *traP* might encode proteins with some regulatory role since in CTnDOT, insertion mutations in these genes affected the levels of TraN and TraG in the cell and these mutations resulted in increased transfer frequency (8). In our study we found that the *traQ* mutation led to higher transfer frequencies, but the *traO* and *traP* mutations were neutral and did not significantly affect transfer, suggesting that there might be some minor differences in the regulation of conjugation. An alternate explanation for the difference is that the large single-crossover insertions (>5 kb) in CTnDOT *traO* and *traP* may have had a polar effect on the downstream gene, *traQ*,. In contrast, the 1.3-kb insertion of mini-Mu into *traP* of CTn341 may not have had the same polar effect. Overall, the mutational studies firmly established a role for the *tra* genes and provided convincing evidence that the mobilization genes also are part of the conjugation system.

Identification of a functional group II intron. Sequence analysis revealed that ORF-36 encoded an RNA maturase/reverse transcriptase homologue that had similarity to an IEP in *Lactococcus lactis* intron Ll.LtrB. The IEP-like ORF was located in the *bmhA* DNA methylase gene, and closer examination revealed a full-length 2,448-bp intron, designated B.f.I1. This intron sequence started with CUGCG (bp 40893) and ended with GCCTGT (bp 38446). The novel start sequence differed from the consensus GUGYG sequence, but otherwise the intron appeared to be a typical group IIA1 intron in the mitochondrial lineage based on the IEP sequence and RNA structure predictions (15, 47; S. Zimmerly, personal communication). The folded RNA structure contained the basic group IIA1 elements, including the exon binding sequences EBS1 and EBS 2 and the highly conserved domains V and VI with the catalytic bulged A (14, 47). Although IEP genes are actually rare in most group II introns, bacterial group II introns usually encode an IEP which is required for in vivo splicing. The B.f.I1 IEP possessed the standard reverse transcriptase domain and domain X, which is required for splicing activity.

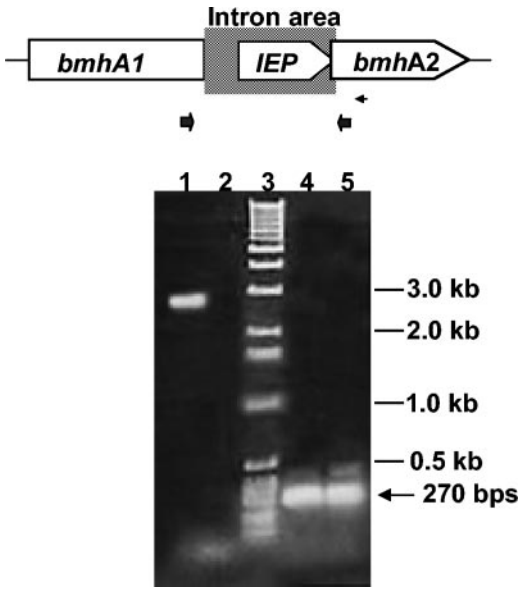

FIG. 5. Schematic diagram of intron area and agarose gel of PCR products from amplified cDNA. The small arrow indicates the priming site used for *bmhA* cDNA synthesis. The large arrows flanking the intron region indicate the primer binding sites used to detect the presence of the intron in *bmhA* mRNA. Lane 1, amplification product produced from chromosomal DNA template; lane 2, control without reverse transcriptase; lane 3, 1-kb ladder molecular size standard; lanes 4 and 5, amplification products from cDNA template.

In order to determine if the putative intron was active, total RNA was reverse transcribed by using a primer located downstream of the intron in *bmhA* and then was PCR amplified by using nested primers that flanked the intron (Fig. 5). The results revealed a 270-bp product which would be possible only if the intron had spliced out of the *bmhA* gene. When the same primers were used to amplify IB399 chromosomal DNA, a 2.7-kb product was observed, and this size corresponded to the size predicted with the intron in place in *bmhA*. Controls in which RNA reverse transcribed with a primer from the *tetQ* gene was used and controls to which reverse transcriptase was not added did not result in any product formation. These data show that B.f.I1 is fully functional and the first functional intron found in *Bacteroides*. We are currently looking to see if

FIG. 6. Summary of percent identity plots showing regions of nucleotide similarity for CTn341 and several *B. fragilis* and *B. thetaiotaomicron* chromosomal segments with homology to the *tra* genes. The plots measured similarity between ungapped, aligned regions consisting of at least 100 bp (green, 55 to 70% identity; red, >70% identity). The contiguous sequence compared is indicated by a genetic map at the top of each panel, and other sequences are represented by open boxes.

there is a role for BmhA and the intron in the transfer of CTn341.

CTn-related sequences in *Bacteroides* chromosomes. A recent description of the *B. thetaiotaomicron* genome revealed that there were multiple homologues of genes similar to those in CTns (48). It was of interest to examine the genetic structure and arrangement of these homologues to better understand their relationship to bone fide active CTns. We examined the two available *Bacteroides* genomes and found four regions in *B. thetaiotaomicron* and two regions in *B. fragilis* with extensive homology to the *tra* region of CTn341. At the amino acid level the homology ranged from 27 to 83% identity, and the similarity extended into the Mob region as well. To determine if these sequences could be part of transmissible elements, we expanded the search around the *tra* genes, looking for other sequences similar to transposon genes. The nucleotide sequences were then aligned and compared by using web-based tools for genome analysis. The results are presented as percent identity plots, which show regions of similarity between one contiguous reference sequence at the top and four secondary sequences that are not necessarily contiguous (Fig. 6). The regions with the greatest identity were the *tra* genes (including divergently transcribed *mobABC*) and *exc*, for which there was between 55 and 70% identity. Three of the regions had *int*

genes related to CTn341, but those that did not have such genes had one or more coding regions for other integrases or transposases. The *tra* genes had very similar structural organizations in all of the genome sequences examined, and they all formed a contiguous module. The only exceptions were *traBC*, which exhibited the lowest levels of identity among all the sequences, and *traD*, which was absent in five of six *tra* regions. The untranslated region between the divergently transcribed *traA* and *mobA* genes also was dissimilar. None of the putative elements had a *tetQ* gene, and only two elements had *rteAB*-related sequences.

Several other genetic features of CTn341 were found in these regions. Genes encoding methylases related to *bmhA* also were found in two of the putative elements, and there was a marked presence of group II intron-like sequences associated with genes in the *tra* regions. Four distinct introns were found based on the amino acid sequences of the IEPs, and in two cases (matR1 and matR4) the introns disrupted *traG* homologues (Fig. 6). Group II introns have been found in a variety of conjugal elements in both gram-negative and gram-positive bacteria, and often they are associated with transfer-related genes (15). The relatively high frequency with which the IEPs were found associated with these CTn-like sequences is consistent with the findings of other workers who found that bacterial introns have an affinity for mobile DNA (14). There appears to be a bias toward intron insertion during lagging-strand synthesis, which would occur in a CTn when it is transferred into the recipient cell (24). It is an obvious advantage for introns to associate with any conjugative element plasmid or transposon in order to facilitate their spread in bacterial populations.

Recently, Franco reported on CTn-like sequences from *B. fragilis* strain NCTC9343 and enterotoxigenic strain 86-5443-2-2 (22). These two elements were closely related to each other, but they did not exhibit significant nucleotide sequence identity to the elements described in this paper. CTn86 included a pathogenicity island that contained the toxin gene *bft*, and CTn9343 had several genes not present in CTn86, but otherwise the sequences were nearly identical to each other. Although there were no data on the ability of these sequences to conjugally transfer, they did encode proteins that had some similarity to MobB, MobC, TraG, TraI, TraM, and TraN (28 to 43% amino acid identity), but the complete *tra* region as found in CTn341 was not present. Potential or proven conjugal elements of many types appear to make up a large portion of the genomes of the two *Bacteroides* species that have been sequenced. These elements arguably promote intraspecies variability, as well as provide a mechanism for interspecies (genus) horizontal gene transfer. The rapid dissemination of antibiotic resistance and the evolution *Bacteroides* pathogenicity islands are potentially driven by these CTns, and further study of their transfer mechanism should increase our understanding of genome plasticity in *Bacteroides*.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grant AI-28884 to C.J.S.

We acknowledge S. Zimmerly for helpful discussions on intron identification and structure. We also thank E. Rocha for critical reading of the manuscript.

REFERENCES

1. **Bass, K. A., and D. W. Hecht.** 2002. Isolation and characterization of cLV25, a *Bacteroides fragilis* chromosomal transfer factor resembling multiple *Bacteroides* sp. mobilizable transposons. J. Bacteriol. **184:**1895–1904.
2. **Baughn, A. D., and M. H. Malamy.** 2002. A mitochondrial-like aconitase in the bacterium *Bacteroides fragilis*: implications for the evolution of the mitochondrial Krebs cycle. Proc. Natl. Acad. Sci. USA **99:**4662–4667.
3. **Beaber, J. W., B. Hochhut, and M. K. Waldor.** 2002. Genomic and functional analyses of SXT, an integrating antibiotic resistance gene transfer element derived from *Vibrio cholerae*. J. Bacteriol. **184:**4259–4269.
4. **Bedzyk, L. A., N. B. Shoemaker, K. E. Young, and A. A. Salyers.** 1992. Insertion and excision of *Bacteroides* conjugative chromosomal elements. J. Bacteriol. **174:**166–172.
5. **Blanc-Potard, A. B., F. Solomon, J. Kayser, and E. A. Groisman.** 1999. The SPI-3 pathogenicity island of *Salmonella enterica*. J. Bacteriol. **181:**998–1004.
6. **Boltner, D., C. MacMahon, J. T. Pembroke, P. Strike, and A. M. Osborn.** 2002. R391: a conjugative integrating mosaic comprised of phage, plasmid, and transposon elements. J. Bacteriol. **184:**5158–5169.
7. **Bonheyo, G., D. Graham, N. B. Shoemaker, and A. A. Salyers.** 2001. Transfer region of a *Bacteroides* conjugative transposon, CTnDOT. Plasmid **45:**41–51.
8. **Bonheyo, G. T., B. D. Hund, N. B. Shoemaker, and A. A. Salyers.** 2001. Transfer region of a *Bacteroides* conjugative transposon contains regulatory as well as structural genes. Plasmid **46:**202–209.
9. **Burrus, V., G. Pavlovic, B. Decaris, and G. Guedon.** 2002. Conjugative transposons: the tip of the iceberg. Mol. Microbiol. **46:**601–610.
10. **Burrus, V., and M. K. Waldor.** 2004. Shaping bacterial genomes with integrative and conjugative elements. Res. Microbiol. **155:**376–386.
11. **Cheng, Q., Y. Sutanto, N. B. Shoemaker, J. F. Gardner, and A. A. Salyers.** 2001. Identification of genes required for excision of CTnDOT, a *Bacteroides* conjugative transposon. Mol. Microbiol. **41:**625–632.
12. **Cheng, Q., N. Wesslund, N. B. Shoemaker, A. A. Salyers, and J. F. Gardner.** 2002. Development of an in vitro integration assay for the *Bacteroides* conjugative transposon CTnDOT. J. Bacteriol. **184:**4829–4837.
13. **Cheng, Q. I., B. J. Paszkiet, N. B. Shoemaker, J. F. Gardner, and A. A. Salyers.** 2000. Integration and excision of a *Bacteroides* conjugative transposon, CTnDOT. J. Bacteriol. **182:**4035–4043.
14. **Dai, L., N. Toor, R. Olson, A. Keeping, and S. Zimmerly.** 2003. Database for mobile group II introns. Nucleic Acids Res. **31:**424–426.
15. **Dai, L., and S. Zimmerly.** 2002. Compilation and analysis of group II intron insertions in bacterial genomes; evidence for retroelement behavior. Nucleic Acids Res. **30:**1091–1102.
16. **Dargemont, C., M. S. Schmidt-Zachmann, and L. C. Kuhn.** 2004. Direct interaction of nucleoporin p62 with mRNA during its export from the nucleus. J. Cell Sci. **108:**257–263.
17. **de Boer, P. A., R. E. Crossley, A. R. Hand, and L. I. Rothfield.** 1991. The MinD protein is a membrane ATPase required for the correct placement of the *Escherichia coli* division site. EMBO J. **10:**4371–4380.
18. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:**387–395.
19. **Ding, Z., K. Atmakuri, and P. J. Christie.** 2003. The outs and ins of bacterial type IV secretion substrates. Trends Microbiol. **11:**527–535.
20. **Ebersbach, G., and K. Gerdes.** 2001. The double *par* locus of virulence factor pB171: DNA segregation is correlated with oscillation of ParA. Proc. Natl. Acad. Sci. USA **98:**15078–15083.
21. **Fickett, J. W.** 1982. Recognition of protein coding regions in DNA sequences. Nucleic Acids Res. **10:**5303–5318.
22. **Franco, A. A.** 2004. The *Bacteroides fragilis* pathogenicity island is contained in a putative novel conjugative transposon. J. Bacteriol. **186:**6077–6092.
23. **Guiney, D. G., P. Hasegawa, and C. E. Davis.** 1984. Plasmid transfer from *Escherichia coli* to *Bacteroides fragilis*: differential expression of antibiotic resistance phenotypes. Proc. Natl. Acad. Sci. USA **81:**7203–7206.
24. **Ichiyanagi, K., A. Beauregard, and M. Belfort.** 2003. A bacterial group II intron favors retrotransposition into plasmid targets. Proc. Natl. Acad. Sci. USA **100:**15742–15747.
25. **Li, L. Y., N. B. Shoemaker, and A. A. Salyers.** 1995. Location and characteristics of the transfer region of a *Bacteroides* conjugative transposon and regulation of transfer genes. J. Bacteriol. **177:**4992–4999.
26. **Mays, T. D., C. J. Smith, R. A. Welch, C. Delfini, and F. L. Macrina.** 1982. Novel antibiotic resistance transfer in *Bacteroides*. Antimicrob. Agents Chemother. **21:**110–118.
27. **Morgan, R. M., and F. L. Macrina.** 1997. *bctA*: a novel pBF4 gene necessary for conjugal transfer in *Bacteroides*. Microbiology **143:**2155–2165.
28. **Murphy, C. G., and M. H. Malamy.** 1993. Characterization of a "mobilization cassette" in transposon Tn4399 from *Bacteroides fragilis*. J. Bacteriol. **175:**5814–5823.
29. **Parker, A. C., and C. J. Smith.** 1993. Genetic and biochemical analysis of a novel ambler class A β-lactamase responsible for cefoxitin resistance in *Bacteroides* species. Antimicrob. Agents Chemother. **37:**1028–1036.
30. **Parker, A. C., and C. J. Smith.** 2004. A multicomponent system is required for tetracycline-induced excision of Tn*4555*. J. Bacteriol. **186:**438–444.
31. **Privitera, G., M. Sebald, and F. Fayolle.** 1979. Common regulatory mecha-

nism of expression and conjugative ability of a tetracycline resistance plasmid in *Bacteroides fragilis*. Nature **278:**657–659.
32. **Rocha, E. R., and C. J. Smith.** 1997. Regulation of *Bacteriodes fragilis katB* mRNA by oxidative stress and carbon limitation. J. Bacteriol. **179:**7033–7039.
33. **Salyers, A. A., N. B. Shoemaker, A. M. Stevens, and L. Y. Li.** 1995. Conjugative transposons: an unusual and diverse set of integrated gene transfer elements. Microbiol. Rev. **59:**579–590.
34. **Sambrook, J., and D. W. Russell.** 2001. Molecular cloning: a laboratory manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
35. **Schwartz, S., L. Elnitski, M. Li, M. Weirauch, C. Riemer, A. Smit, E. D. Green, R. C. Hardison, and W. Miller.** 2003. MultiPipMaker and supporting tools: alignments and analysis of multiple genomic DNA sequences. Nucleic Acids Res. **31:**3518–3524.
36. **Schwartz, S., Z. Zhang, K. A. Frazer, A. Smit, C. Riemer, J. Bouck, R. Gibbs, R. Hardison, and W. Miller.** 2000. PipMaker—a web server for aligning two genomic DNA sequences. Genome Res. **10:**577–586.
37. **Shoemaker, N. B., H. Vlamakis, K. Hayes, and A. A. Salyers.** 2001. Evidence for extensive resistance gene transfer among *Bacteroides* spp. and among *Bacteroides* and other genera in the human colon. Appl. Environ. Microbiol. **67:**561–568.
38. **Smith, C. J.** 1985. Characterization of *Bacteroides ovatus* plasmid pBI136 and structure of its clindamycin resistance region. J. Bacteriol. **161:**1069–1073.
39. **Smith, C. J., and D. R. Callihan.** 1992. Analysis of rRNA restriction fragment length polymorphisms from *Bacteroides* spp. and *Bacteroides fragilis* isolates associated with diarrhea in humans and animals. J. Clin. Microbiol. **30:**806–812.
40. **Smith, C. J., and A. C. Parker.** 1993. Identification of a circular intermediate in the transfer and transposition of Tn*4555*, a mobilizable transposon from *Bacteroides* spp. J. Bacteriol. **175:**2682–2691.
41. **Smith, C. J., A. C. Parker, and M. Bacic.** 2001. Analysis of a *Bacteroides* conjugative transposon using a novel "targeted capture" model system. Plasmid **46:**47–56.
42. **Smith, C. J., G. D. Tribble, and D. P. Bayley.** 1998. Genetic elements of *Bacteroides* species: a moving story. Plasmid **40:**12–29.
43. **Stevens, A. M., J. M. Sanders, N. B. Shoemaker, and A. A. Salyers.** 1992. Genes involved in production of plasmidlike forms by a *Bacteroides* conjugal chromosomal element share amino acid homology with two-component regulatory systems. J. Bacteriol. **174:**2935–2942.
44. **Stevens, A. M., N. B. Shoemaker, L.-Y. Li, and A. A. Salyers.** 1993. Tetracycline regulation of genes on *Bacteroides* conjugative transposons. J. Bacteriol. **175:**6134–6141.
45. **Sutanto, Y., J. M. DiChiara, N. B. Shoemaker, J. F. Gardner, and A. A. Salyers.** 2004. Factors required in vitro for excision of the *Bacteroides* conjugative transposon, CTnDOT. Plasmid **52:**119–130.
46. **Sutanto, Y., N. B. Shoemaker, J. F. Gardner, and A. A. Salyers.** 2002. Characterization of Exc, a novel protein required for the excision of *Bacteroides* conjugative transposon. Mol. Microbiol. **46:**1239–1246.
47. **Toor, N., G. Hausner, and S. Zimmerly.** 2001. Coevolution of group II intron RNA structures with their intron-encoded reverse transcriptases. RNA **7:**1142–1152.
48. **Xu, J., M. K. Bjursell, J. Himrod, S. Deng, L. K. Carmichael, H. C. Chiang, L. V. Hooper, and J. I. Gordon.** 2003. A genomic view of the human-*Bacteroides thetaiotaomicron* symbiosis. Science **299:**2074–2076.