

Analysis of Operons Encoding 23S rRNA of *Clostridium botulinum* Type A

ALISON K. EAST, DAPHNE E. THOMPSON, AND MATTHEW D. COLLINS*

Department of Microbiology, AFRC Institute for Food Research, Reading Laboratory, Earley Gate, Whiteknights Road, Reading RG6 2EF, United Kingdom

Received 6 July 1992/Accepted 13 October 1992

Southern hybridization analysis of *Clostridium botulinum* type A chromosomal DNA indicated the presence of six copies of the 23S rRNA gene. Fragments of DNA encoding 23S rRNA were amplified by polymerase chain reaction and cloned in *Escherichia coli*. Three clones examined by restriction enzyme and sequence analysis were found to be derived from different operons. Sequence determination of the entire insert of two clones revealed nine nucleotide changes in the genes coding for 23S rRNA (99.7% sequence identity) between operons encoded on the same chromosome, showing microheterogeneity in the rRNA operons of this organism.

The species *Clostridium botulinum* comprises a diverse group of gram-positive, anaerobic, spore-forming, rod-shaped bacteria (5). *C. botulinum* type A produces a botulinum neurotoxin (BoNT/A) which, with toxins of type B (BoNT/B) and type E (BoNT/E), is responsible for the majority of cases of foodborne human botulism (17). Strains of *C. botulinum* have been grouped together under a common name because of their ability to produce BoNT. Within toxin-producing *C. botulinum*, four distinct groups have been identified on the basis of phenotypic characteristics (12). In addition, some strains of *Clostridium barati* and *Clostridium butyricum* have been reported to produce BoNT (16), and strains which phenotypically resemble *C. botulinum* but are nontoxicogenic have been described (10). These inconsistencies clearly demonstrate that the classification and nomenclature of the *C. botulinum* group are unsatisfactory. Chromosomal DNA-DNA pairing studies have revealed considerable heterogeneity within the *C. botulinum* group of species (14, 15, 22). It is known that this technique can only be used to measure the relationship of closely related organisms (i.e., strains within a species or closely related species). There is growing recognition that rRNA sequencing is one of the most powerful and precise methods for determining the phylogenetic interrelationships of bacteria (25), but little is known of sequence variation in rRNA operons (*rrn*) of bacteria other than *Escherichia coli* (4) or species of *Bacillus* (20). As the foundation for an investigation of the genealogical relationships within the *C. botulinum* complex of species, we have determined the nucleotide sequence of the large subunit rRNA gene and intergenic spacer regions of more than one *rrn* operon of *C. botulinum* type A. We report microheterogeneity between the operons, including regions encoding 23S rRNA, and comment on the implications of this for determining close genetic interrelationships.

Analysis of cloned *rrn* operons. *C. botulinum* type A (NCTC 7272) was grown anaerobically in reinforced clostridial medium (Oxoid) at 37°C, and chromosomal DNA was prepared by the method of Farrow et al. (7). Polymerase chain reaction (PCR) products were generated in a reaction with primers designed to conserved regions of 16S and 5S rRNA (5'-AATCATCATGCCCTTATGACCTGGCTA-3' [positions 1197 to 1224, *E. coli* numbering] and 5'-TTAAGT

TCCATGTTTCGGTATGGGAA-3' [positions 33 to 57, *E. coli* numbering], respectively) by using 1 U of *Taq* polymerase (Amersham International) under the following conditions: denaturation at 95°C for 5 min followed by 25 cycles of denaturation at 92°C for 1 min, annealing at 50°C for 1 min, and primer extension at 72°C for 5 min. Amplification was completed by a final incubation at 72°C for 10 min. Template DNA was at final concentration 20 ng/μl, and deoxynucleotides (BCL) were at final concentration 200 μM. The fragments were cloned into the *EcoRV* site of pBluescript KS+ (Stratagene) after modification as described previously (6). From a cloning experiment, three recombinant plasmids were isolated (designated pCBA2, pCBA4, and pCBA5) containing DNA inserts (ca. 3.5 kb) encoding 23S rRNA. Sequence analysis using the dideoxynucleotide chain termination method (6, 21) revealed the organization of the cloned *rrn* operons to be a classical 16S rRNA-spacer-23S rRNA-spacer-5S rRNA structure. The restriction maps of the three recombinant plasmids were found to be similar but not identical (Fig. 1). Approximately 400 bases at the 16S rRNA end of the PCR fragment was missing in clone pCBA4. This was revealed by sequencing (Fig. 2) and is probably due to nuclease activity during cloning. This accounts for the loss of the *SmaI* site in pCBA4, but the *BamHI* of pCBA2 is within the coding region for 23S rRNA and is absent from the other two clones. Sequence analysis showed that in pCBA2, the nucleotide at position 100 (Fig. 3) is a G instead of the A of pCBA4 and pCBA5. This base substitution introduces a *BamHI* site into pCBA2. A *ScaI* site is present towards the end of genes coding for 23S rRNA (rDNA) in pCBA2 and pCBA5 but absent from pCBA4. The base substitution which introduces the *ScaI* site into pCBA5 is at position 2182 of 23S rRNA (Fig. 3).

Further evidence that the clones are distinct from each other, and therefore are likely to represent different operons, was obtained from analysis of the 16S-23S rRNA- and 23S-5S rRNA-spacer regions (Fig. 2 and 4). The 16S-23S rRNA-spacer sequences of pCBA2 and pCBA4 are very similar (three mismatches in 187 nucleotides), whereas that of pCBA5 has a deletion of approximately 70 nucleotides (Fig. 2). In contrast, the 23S-5S rRNA-spacer sequences of pCBA2 and pCBA5 are almost identical and contain a deletion of approximately 40 bp compared with that of pCBA4 (Fig. 4). Sequence analysis of these spacer regions has not revealed the coding for tRNA molecules. tRNA

* Corresponding author.

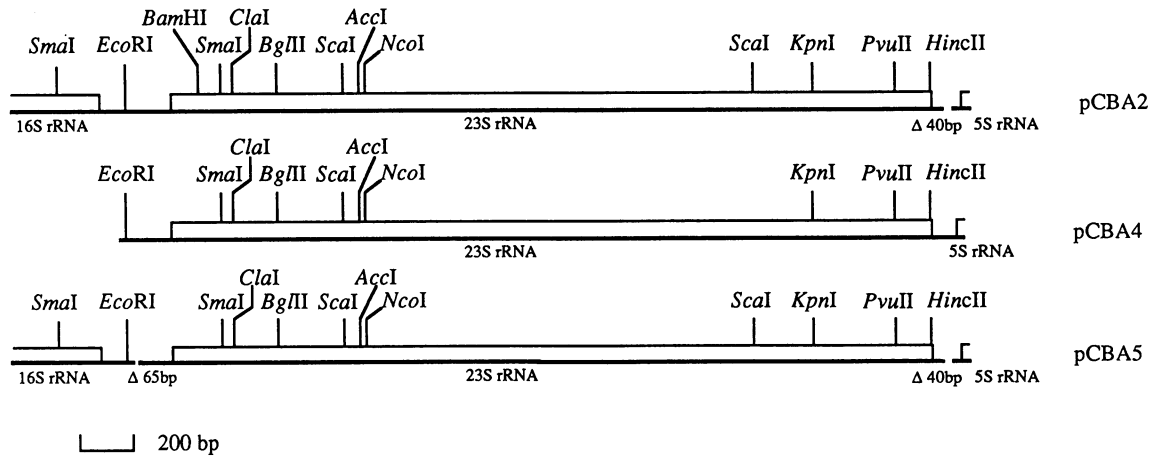


FIG. 1. Restriction map of plasmids pCBA2, pCBA4, and pCBA5. Deletions are shown with respect to pCBA4.

molecules are found in the 16S-23S rRNA intergenic spacers of all *E. coli* *rrm* operons (26) and some of those of *Streptococcus pneumoniae* (2), *Plesiomonas shigelloides* (6), and *Listeria monocytogenes* (23). Studies on *Clostridium perfringens* have revealed the presence of 10 *rrm* operons, and the authors suggest two of these may contain coding for tRNA molecules between the 16S and 23S rDNA (9). However, since preliminary Southern blot analysis has revealed the presence of six copies of the gene on the *C. botulinum* type A chromosome (Fig. 5), the sequence data shown here do not exclude the possibility of tRNA encoded in the intergenic spacer regions of the other *rrm* operons. There is no evidence in the sequence for inverted repeats upstream and downstream of the 23S rDNA, such as make up the RNase III sites used in rRNA processing in gram-negative organisms such as *E. coli* (27) and *Caulobacter crescentus* (8). In gram-positive organisms, a different consensus sequence implicated in rRNA processing has been described (9). This sequence is found in all gram-positive organisms so far sequenced, including mycoplasmas (13, 19), *Streptomyces*

coelicolor (3), and *C. perfringens* (9). These sequences, with some modifications, are found in the 16S-23S and 23S-5S rDNA-spacer regions of *C. botulinum* type A (Fig. 4 and 5).

Analysis of 23S rDNA. The complete nucleotide sequence of the region encoding 23S rRNA of *C. botulinum* type A determined from pCBA4 (2,896 nucleotides) is shown in Fig. 3. Sequence determination of the same region of a second clone (pCBA5) revealed 99.7% identity (corresponding to nine nucleotide differences; Fig. 3). Two of the differences between the clones (positions 150 and 174; Fig. 3) are base-paired. Other nucleotide heterogeneity occurred in open regions or in positions of uncertain secondary structure. Although it is usual for the product of more than one PCR experiment to be sequenced to confirm the data, the sequence deviation shown here is too high to be attributable to *Taq* polymerase (error rate estimated to be approximately 1 in 9,000 nucleotides [24]) and, together with the sequence variation of the intergenic spacer regions, is clearly evidence of microheterogeneity within *rrm* operons of this organism. It has been shown that recombination of DNA leading to the

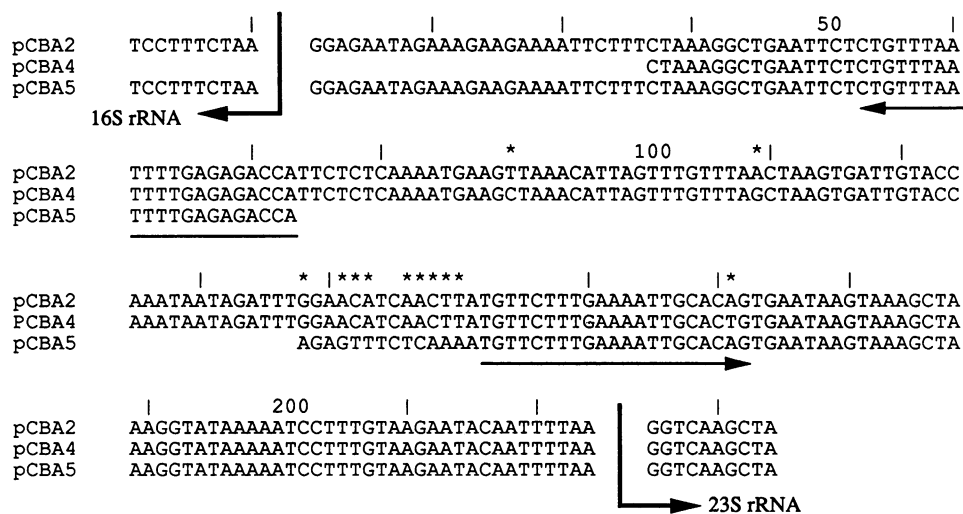


FIG. 2. Nucleotide sequence of 16S-23S rDNA intergenic spacer region of pCBA2, pCBA4, and pCBA5. Nucleotide differences are indicated by * above the sequence. Plasmids pCBA2 and pCBA5 contain an additional ~350 nucleotides of 16S rDNA (data not shown). The end of 16S rDNA and the start of 23S rDNA are indicated. Arrows under the sequence show inverted repeats believed to be important for processing rRNA (9).

```

1  GGTC AAGCTACA AAAGGGCGCATGGTGAATGCCT TGGCACTAGGAGCCGAAAGGACCGGATAAGCTGCGATAAGCCCTGGGTAGCCGAAATAGGCTGA
101  GATCCAGGGATTTC CCAATGAGGAACTCACATGGGTAACCCCATGTATCATGCATGAATACATAGGTGTATGAGGGTAAAC CCGGG GAACGAAACAT
201  CTAAGTACCCGGAGGAGAGAAAGAAAA ATCGAT TTCCCTAAGTAGCGGGCAGCGAACGGGAAAGAGCCCAACCAGAACTTTTCTGGGGTTCGGGAT
301  AGATCATAAAAAGAGGATCTTAATCGAAGAGGGGTGGAACGCCCTACCATAGAAAGTAAATAGTCCCTGTAGGTA AAAAGAGAAAACTTCGAGATCTAA
401  TCCAGAGTACCACGAGACGCTGAAACCTTGTGGGAAGCAGGAGGACCCTCCCAAGGCTAAATACTACCTAGTGACCGATAGTGAAGCAGTACCCTG
501  AGGGAAGGAAAGAAAGTCCCGGAAGGAGTGAATAGAACCTGAAACCGTGTGCCACACCCGATCGGAGCACGTTAAAGTGTGACGATGTGCTTTTTGT
601  AGAACGAGCCAGCGAGTTACGCTATGTAGCAAGGTTA AGTACT TAAGGTATGGAGCCGAAAGGAAACCGAGTCTGAAAAGGGCGAAAAGTTGCATGGTGT
701  AGAC CCGAAACCGGGTGACCTA CCAT TGCCAGGTTGAAAGCAGAGTAAATCTCGTGAGGACCGCAACCACGTTGGTGTGAAAAACCATGGGATGAGC
801  TGTGGATAGCGGAGAAATCCAATCGAACTCGGAGATAGCTGGTTCCTCGAAATAGCTTTAGGGTAGCGTCGTGAATGAGTAATGGAGGTAGAGC
901  ACTGAATGGGCTAGGGGTATAGTAGTTACCGAACCCATCAAACCTCCGAATGCCATATACTTGTATCACGGCAGTCAGACTCGCAATGTAAGATCCGT
1001  AGTCAAAAGGAAACAGCCAGACCATCAGCTAAGTCCCAAAGTGAAGTTAAGTGGAAAAGGATGTGGGATTTCTAAGACAACCTAGGATGTGGCTTA
1101  GAAGCAGCCACTCATTTAAAGAGTGCCTAATAGCTCACTAGTCAAGAGATCCTGCGCGAAGATGTCGGGGCTCAAACCTACCACCGAAGCTATGGGTG
1201  TACACTATGTGTACGCGGTAGAGGAGCTTTCTGTATGGTTGAAGTCGTACCGTAAGGAGCGGTGGACTGTACAGAAGTGAGAATGCTGGCATAAGTAGC
1301  GAGAAATAAGTGAGAACTTTATGGCCGAAAACCTAAGGTTTCTGAGGAGGTTCTGCTCCCTCAGGGTTAGTCGGGACCTAAGCCGAGGCCGAAAGGTA
1401  GGTGATGGCAATCGGTTGATATCCGATACCGCTATTACGTTTGAGAAATGGGGTACGACAGTAGGATAAGATGTGCGCACTATTGGATGTGCGTCT
1501  AAGCACTTAGGCATGCTGTATAGGCAAATCCGTCGAGCTAAGCTGAGGTGTGATGGGAGGCCATTTATGGCAGGTTCTGATCCACACTGCCAAGAAA
1601  AGCCTCTCTCGAGTGAATAGGTGCCGTAACCGCAAACCGACACAGGTAGGTGAGGAGAGAATCCTAAGGCCATCGGAAGAATTGCTGTTAAGGAACTCGG
1701  CAAATTAACCCCGTAACTTCGGGAGAGGGGTCCACGAAAGTAGCCGACAGAGAATAGGCCCAAGCAACTGTTAGCAAAAACACAGGTCTCTGCTAA
1801  AGCGAAAGCTGATGTATAGGGGTGACGCTGCCCGTCTGGAAGGTAAAGGGGAACACTTAGCGTAAGCGAAGGTGTGAACCTAAGCCCGAGTAAACG
1901  GCGGCCGTAACATAACGGTCTCAAGTAGCGAAATCCCTGTGCGGTAAAGTCCGACCCGACGAATGGCGTAATGATTTGGGCAGTGTCTCAACAGCA
2001  AATCCGGCGAAATTTAGTGCAAGTGAATGCTTGCTACCCCGATTGGACGGAAAGACCCCGTAGAGCTTTACTGTAGCTTAGCACTGAATCTCGGTA
2101  TTGCTGTACAGGATAGGTGGGAGACTTGGAAACTTGGGCGTACGTGAGTGGAGTATCCTTTGGGATACCACCTGACAGCACTGGGGTTCTAACCGGC
2201  GGCCATGAATCTGGTCACGGACATTGTTAGTGGGCAGTTTACTGGGGCGTCCCTCTAAAAGTAAACGGAGCGCCAAAAGTCCCTCAGCGCG
2301  GTCGGAATCGCGGAAGAGTGCAAGGCAGAAGGAGCTTACTGCGACACATACAGTGGAGCAGGACGAAAGTCCGGCTTAGTGATCCGGTGGTAC
2401  CTGCTGGGAGGCCATCGCTCAACGGATAAAAGCTACCTCGGGATAACAGGCTGATCTCCCAAGAGTCCACATCGACGGGAGGTTTGGCACCTCGA
2501  TGTGGCTCGTGCATCCTGGGCTGAAGTAGGTCCTCAAGGGTTGGGCTGTTCGCCATTAAAGCGGCACGCGAGCTGGGTTAGAACGTCGTGAGACAG
2601  TTCGGTCCCTATCCGTCGCGGGCGTAGGAAATTTGAGAGGAGCTGCTTACTGACGAGAGGACCGGATGGACCAACCTCTGGTGCACCAAGTTGTACGC
2701  CAGTGGCA CAGCT GGATAGCTATGTTGGACTGGATAAACGCTGAAAGCATCTAAGCGTGAAGCCAACCTCAAGATGAGATTCCCATAGCGTAAGCTAG
2801  TAAGACTCTGGAAGAACACCAGGTTGATAGTTCAGAGATGTAAGCATGGCAACATGTTAAGTTGACTGACTACTAATAAGTCGAGGGCTTGACCAA

```

FIG. 3. Complete nucleotide sequence of 23S rRNA encoded by pCBA4. The nine nucleotide changes in pCBA5 are indicated by * above the sequence, with the base substituted shown below. Restriction sites shown in Fig. 1 are indicated by underlining.

creation of chimeric molecules is possible during PCR of very similar target genes, e.g., those of RNA viruses (18). Because there are multiple copies of rRNA operons, these genes fall into this category. Meyerhans et al. (18) estimated that the frequency of recombinant clones was at maximum 5% and showed that this could be reduced up to sixfold by a long elongation step in PCR. To reduce such recombinational events by minimizing incomplete elongation, a lengthy elongation step (5 min at 72°C) was used over a low number of cycles (25 cycles). Under such conditions, the possibility

of producing chimeric PCR products should be considered remote. Comparison of the *C. botulinum* 23S rRNA gene sequences with those of other low G+C gram-positive bacteria (e.g., *L. monocytogenes* [23], *Bacillus cereus* [1], *Bacillus subtilis* [11]) revealed 79.9 to 80.5% sequence identity. No large insertions or deletions in the 23S rRNA were evident. The 23S rRNA has a G+C content of 51% compared with a total genomic G+C content of 27%. This is similar to the situation in *C. perfringens* (9).

rRNA sequence data are being used increasingly to deter-

- analysis. *J. Clin. Microbiol.* **29**:1114–1124.
11. Green, C. J., G. C. Stewart, M. A. Hollis, B. S. Volt, and K. F. Bott. 1985. Nucleotide sequence of the *Bacillus subtilis* ribosomal RNA operon, *rnmB*. *Gene* **37**:261–266.
 12. Hauschild, A. H. W. 1989. *Clostridium botulinum*, p. 111–189. In M. P. Doyle (ed.), *Foodborne bacterial pathogens*. Marcel Dekker, Inc., New York.
 13. Iwami, M., A. Muto, F. Yamao, and S. Osowa. 1984. Nucleotide sequence of the *rnmB* 16S ribosomal RNA gene from *Mycoplasma capricolum*. *Mol. Gen. Genet.* **196**:317–322.
 14. Lee, W. H., and H. Riemann. 1970. Correlation of toxic and nontoxic strains of *Clostridium botulinum* by DNA composition and homology. *J. Gen. Microbiol.* **60**:117–123.
 15. Lee, W. H., and H. Riemann. 1970. The genetic relatedness of proteolytic *Clostridium botulinum* strains. *J. Gen. Microbiol.* **64**:85–90.
 16. McCroskey, L. M., C. L. Hatheway, L. Fenicia, B. Pasolini, and P. Aureli. 1986. Characterization of an organism that produces type E botulin toxin but which resembles *Clostridium butyricum* from the feces of an infant with type E botulism. *J. Clin. Microbiol.* **23**:201–202.
 17. McCroskey, L. M., C. L. Hatheway, B. A. Woodruff, J. A. Greenberg, and P. Jurgenson. 1991. Type F botulism due to neurotoxicogenic *Clostridium baratii* from an unknown source in an adult. *J. Clin. Microbiol.* **29**:2618–2620.
 18. Meyerhans, A., J.-P. Vartamian, and S. Wain-Hobson. 1990. DNA recombination during PCR. *Nucleic Acids Res.* **18**:1687–1691.
 19. Rasmussen, O. F., J. Frydenberg, and C. Christiansen. 1987. Analysis of the leader and spacer regions of the two rRNA operons of *Mycoplasma* PG50: two tRNA genes are located upstream of *rnmA*. *Mol. Gen. Genet.* **208**:23–29.
 20. Raue, H. A., A. Rosner, and R. J. Planta. 1977. Heterogeneity of the genes coding for 5S RNA in three related strains of the genus *Bacillus*. *Mol. Gen. Genet.* **156**:185–193.
 21. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
 22. Suen, J. C., C. L. Hatheway, A. G. Steigerwalt, and D. J. Brenner. 1988. *Clostridium argentinense* sp. nov.: a genetically homogeneous group composed of all strains of *Clostridium botulinum* toxin type G, and some nontoxic strains previously identified as *Clostridium subterminale* or *Clostridium hastiforme*. *Int. J. Syst. Bacteriol.* **38**:375–381.
 23. Thompson, D. E., J. T. Balsdon, J. Cai, and M. D. Collins. 1992. Studies on the ribosomal RNA operons of *Listeria monocytogenes*. *FEMS Microbiol. Lett.* **96**:219–224.
 24. Tindall, K. R., and T. A. Kunkel. 1988. Fidelity of DNA synthesis by *Thermus aquaticus* DNA polymerase. *Biochemistry* **27**:6008–6013.
 25. Woese, C. R. 1987. Bacterial evolution. *Microbiol. Rev.* **51**:221–271.
 26. Young, R. A., R. Macklis, and J. A. Steitz. 1979. Sequence of the 16S-23S spacer in two ribosomal RNA operons of *Escherichia coli*. *J. Biol. Chem.* **254**:3264–3271.
 27. Young, R. A., and J. A. Steitz. 1978. Complementary sequences 1700 nucleotides apart form a ribonuclease III site in *Escherichia coli* ribosomal precursor RNA. *Proc. Natl. Acad. Sci. USA* **75**:3593–3597.