# THE TOXIGENIC ELEMENT OF *CLOSTRIDIUM DIFFICILE* STRAIN 10463 AND ITS TRANSCRIPTIONAL ANALYSIS IN STRAINS WHICH DIFFER IN TOXIGENICITY

by

Georgia Ann Hammond

Dissertation submitted to the graduate faculty of
Virginia Polytechnic Institute and State University in
partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Anaerobic Microbiology

(Department of Biochemistry and Anaerobic Microbiology)

**APPROVED:**

J. L. Johnson, Chairman

D. R. Dean

N. R. Krieg

T. D. Wilkins

T. J. Larson

March, 1994

Blacksburg, Virginia

The Toxigenic Element Defined in *Clostridium difficile* VPI Strain 10463 and its Transcriptional Analysis in Strains Which Differ in Toxigenicity

by

Georgia Ann Hammond

Committee Chairman: John L. Johnson
Department of Biochemistry and Anaerobic Microbiology

(ABSTRACT)

*Clostridium difficile* is a Gram positive, anaerobic bacterium which produces two potent protein toxins, A and B. The genes for toxins A and B have been previously cloned and sequenced and lie within 1.4 kb of each other.

Upstream and downstream boundaries between sequences shared by both toxigenic and nontoxigenic strains and those sequences which are unique to toxigenic strains were established.

A toxigenic element was defined in *C. difficile* strain 10463 which is 19.6 kb in length and is comprised of five open reading frames, including the toxin A and B genes. One of these open reading frames is previously unidentified and is located upstream of toxin B.

Products of Polymerase Chain Reaction (PCR) amplification of three regions in the toxigenic element: the upstream boundary, the downstream boundary, and the region between the toxin A and B genes, were all identical in length in six toxigenic strains, indicating that the toxigenic element is conserved among these strains. A short fragment unique to nontoxigenic strains and occupying the same position on the chromosome as the toxigenic element was identified. PCR products of this region were identical in length in three nontoxigenic strains.

Transcriptional analyses were undertaken using probes to each of the five open reading frames in the toxigenic element. Transcripts were detected for four of the open reading frames which are contiguous and transcribed in the same direction. In addition, a very large transcript, corresponding to the length of the four open reading frames and processing intermediates were detected, indicating that the toxin genes are cotranscribed. A promoter region and processing sites were identified. Sizes were determined for each of the individual transcripts which correspond well with the sizes of the open reading frames.

Six toxigenic strains which vary considerably in toxin production were selected for analysis to determine whether DNA sequence variation could account for the observed differences in toxin production. DNA restriction fragment length polymorphisms were examined, toxin-specific transcripts were analyzed, and sequences of regulatory regions were determined and compared. Whereas quantitative differences in toxin-specific transcripts were found among the toxigenic strains, the remaining analyses showed that DNA sequences were conserved among these strains.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor, Dr. John L. Johnson, for the opportunity to learn, to work, and to pursue my interest in science in his laboratory. I appreciate very much his patience and encouragement. I would also like to thank each of my advisory committee members: Dr. Dennis Dean, Dr. Tracy Wilkins, Dr. Noel Krieg, and Dr. Timothy Larson for their guidance and support.

I would like to extend my sincere appreciation to Julianna Toth and Dr. Christopher Russell, as well as to Dr. Shu-Zhen Wang and Dr. Harold Zappe for their friendship and encouragement and their willingness to listen and share ideas.

Finally, I would like to express my deepest appreciation for the continual love and support of my family and friends, especially, my parents, George and Blanche Hammond, my husband, Dr. Karl Pedersen, and my sons, Jens and Kalle.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER ONE

CHAPTER THREE

LIST OF TABLES

# INTRODUCTION

The advent of the antibiotic era ushered in great hope for controlling debilitating bacterial infections. However, concommitant with the widespread use of antibiotics came a number of antibiotic-related maladies, including both antibiotic-associated diarrhea (AAD) and pseudomembranous colitis (PMC). Estimates indicate that *Clostridium difficile* may be the cause of approximately 25% of the reported cases of antibiotic-associated diarrhea (Lyerly et al., 1988). Furthermore, *C. difficile* has been established as the causative agent for pseudomembranous colitis.

*Clostridium difficile* produces two potent toxin proteins (Taylor et al., 1981; Sullivan et al., 1982; Banno et al., 1984). Production of these toxins in the gut ultimately leads to the disease. Pseudomembranous colitis is characterized by initial diarrhea which then progresses to severe inflammation of the mucosal cells which line the colon. The disease, which can be fatal, takes its name from the fact that ultimately, during the course of the disease, sheath-like pseudomembranes are formed which are composed of mucous, fibrin, white blood cells, and epithelial cells. The majority of cases of pseudomembranous colitis are nosocomial or hospital-acquired (McFarland et al., 1989). Diarrhea, which is symptomatic of the disease, facilitates the spread of *C. difficile*. In addition, the Gram positive *C. difficile* forms spores, which can persist for long periods of time (Mulligan et al., 1979).

Initially, pseudomembranous colitis was correlated with the use of certain antibiotics, in particular, clindamycin, an antibiotic which had been developed for treating anaerobic infections (Gurwith et al., 1977; Swartzberg et al., 1977; Allo et al., 1979; Bartlett, 1979). However, it has now been demonstrated that treatment with almost any antibiotic can result in the disease. In fact, antibiotic therapy is not an essential prerequisite for pseudomembranous colitis. One of the first reports of the disease was in 1893 (Finney), long before antibiotic therapy was in effect. Any procedure which causes a change in the normal protective bacterial flora such as intestinal surgery or obstruction, thereby allowing *C. difficile* to colonize the gut, can result in pseudomembranous colitis.

*Clostridium difficile* was identified as the causative agent for pseudomembranous colitis during the process of characterization of the disease (George et al., 1978a; George et al., 1978b; Larson et al., 1978). In the late 1970s cytotoxic activity was detected in stool samples of patients with pseudomembranous colitis. A search for the source of this activity ultimately led to a toxin-producing bacterium (Bartlett et al., 1978; Bartlett et al., 1979). The finding that the cytotoxic activity was neutralized by gas gangrene antiserum pointed to the genus *Clostridium*, since gas gangrene itself is caused by a number of clostridial species and therefore the antiserum is actually comprised of a mixture of antisera. Antiserum to *Clostridium sordellii* neutralized the cytotoxic activity associated with stools from patients with pseudomembranous colitis, but *C. sordellii* was not isolated in

significant numbers from pseudomembranous colitis patients (Larson and Price, 1977; Rifkin et al., 1977; Chang et al., 1978; Rifkin et al., 1978a; Allo et al., 1979). *Clostridium difficile* was found in high numbers in stools from patients with pseudomembranous colitis and investigators focused on *C. difficile* as the causative agent of pseudomembranous colitis. The cross reaction between *C. sordellii* antiserum and *C. difficile* toxin can be explained by the fact that both *C. difficile* and *C. sordellii* produce toxins with a high degree of similarity (Martinez and Wilkins, 1988; Martinez and Wilkins, 1992).

Both Taylor et al. (1980) and Banno et al. (1981) provided evidence for the existence of two toxins from *Clostridum difficile* when they demonstrated an enterotoxic activity which was separable from the cytotoxic activity by anion-exchange chromatography. The two toxins were designated A and B, (D-1 and D-2, respectively, of Banno et al., 1981) based on their order of elution from the anion exchange resin. Toxin A, the enterotoxin, binds less tightly to the resin and therefore elutes prior to toxin B.

## BIOLOGICAL ACTIVITIES OF TOXINS A AND B

Toxin A and toxin B are both cytotoxins, although it is the toxin B activity which is detected in stool samples since it is the stronger cytotoxin of the two (Lyerly, Krivan, and Wilkins, 1988). Both toxin A and toxin B cause a similar

rounding effect on tissue culture cells. Toxin A and B are active against all of the mammalian cell lines which have been tested (Borriello and Welch, 1984; Lyerly and Wilkins, 1986), although different cell lines vary in sensitivity to the toxins (Lyerly and Wilkins, 1988). Most of the research directed toward elucidating the mechanism whereby the *C. difficile* toxins cause cell rounding has focused on toxin B. Toxin B disrupts the cellular actin microfilament system leading to a change in cell shape (Thelestam and Bronnegard, 1980; Wedel et al., 1983; Thelestam and Florin, 1984; Mitchell et al., 1987).

Both toxins are lethal toxins in experimental animals. Intraperitoneal or subcutaneous injections of either toxin A or B cause death in mice, rats, hamsters, and Rhesus monkeys (Taylor et al., 1981; Sullivan et al., 1982; Arnon et al., 1984; Lyerly et al., 1986; Lyerly et al., 1988). Since death in these experimental animals is unaccompanied by significant pathology, the manner in which the toxins act as lethal toxins remains to be elucidated. Both toxins act as lethal toxins when injected intracecally in hamsters (Lyerly and Wilkins, 1988). Intragastric doses of toxin A cause hemorrhage, fluid accumulation, diarrhea, and death in hamsters; however, toxin B, given intragastrically, elicits no response in hamsters unless it is administered along with toxin A, or the intestine is damaged in some manner (Lyerly et al., 1985). These data point to the possibility that toxin A and B may act in concert, with toxin A causing the tissue damage and facilitating the entry of toxin B, which can then act as a lethal toxin.

The enterotoxic activity attributed to toxin A (Taylor et al., 1980; Banno et al., 1981) was first detected through the use of ligated rabbit intestinal loops (Rifkin et al., 1978b; Humphrey et al., 1979). Ligated ileal loops treated with toxin A are filled with hemorrhagic fluid, apparently a result of the significant tissue damage to the mucosal cells caused by toxin A (Lyerly, Krivan, and Wilkins, 1988). *Clostridium difficile* toxin A is as potent an enterotoxin as *Vibrio cholerae* toxin in loop assays; however, the hemorrhagic fluid response elicited by toxin A differs from the clear, watery fluid which accumulates subsequent to cholera toxin injection (Lyerly et al., 1990). The basis for this difference lies in the mechanism of action of the two toxins. Cholera toxin acts by stimulating the cellular adenyl cyclase system, resulting in fluid accumulation with essentially no tissue damage. Conversely, *C. difficile* toxin A is responsible for extensive damage to the gut mucosa, and it is the tissue damage which results in the fluid response, although the precise manner in which the toxin acts remains to be elucidated.

## CHARACTERIZATION OF THE TOXIN A AND B PROTEINS

The toxin A and B proteins from *Clostridium difficile* strain 10463 have been purified and characterized (Banno et al., 1981; Taylor et al., 1981; Sullivan et al., 1982; Banno et al., 1984; Rihn et al., 1984; Lyerly et al., 1986; Pothoulakis et al., 1986; Krivan and Wilkins, 1987). Toxins A and B are very large proteins. Under

denaturing conditions, both toxins demonstrate an apparent molecular weight ($M_r$) of 250,000 - 300,000. These data are in good agreement with the sizes of the open reading frames for both of the toxins (Barroso et al., 1990, Dove et al., 1990; Eichel-Streiber et al., 1990; Sauerborn and Eichel-Streiber, 1990).

Both of the toxins are acidic molecules. Lyerly et al. (1986) determined isoelectric points for toxin A of 5.2-5.7 and toxin B of 4.1-4.5.

There are four conserved cysteine residues between toxin A and toxin B (Wilkins, pers. comm.). Treatment of either toxin A or B with oxidizing agents can lead to inactivation. The inactivation of oxidized toxin is irreversible; however, addition of reducing agents such as dithiothreitol prior to or concurrent with the oxidizing agents can protect against the inactivation (Lyerly et al., 1986a). In this same study cytotoxic activity of the toxins was assayed in the presence of sulfhydryl-inactivating agents as well. The biological activities of the toxins remained unaffected by either sulfhydryl-inactivating or reducing agents, indicating that sulfhydryl moieties are not involved in the activity of either toxin.

A hydrophobic region comprised of approximately fifty amino acids occurs near the midpoint of each protein (Eichel-Streiber et al., 1992a; Wilkins, pers. comm.). It remains to be determined whether this region serves a membrane-spanning function for toxins A and B.

The two toxins show an overall 44.8 percent amino acid identity. Including structurally equivalent residues in this calculation, increases the amino acid similarity

of the two toxins to 63 percent (Eichel-Streiber et al., 1992a). These data point to evolutionary conservation between the two toxin proteins.

The carboxy-terminal portion of toxin A is comprised of a series of repeating units (Dove et al., 1990; Eichel-Streiber et al., 1990). Toxin B, likewise, consists of a repetitive carboxy-terminus and a nonrepetitive amino-terminus. Eichel-Streiber (1992) has proposed that both toxin A and toxin B show structural and functional dualism; that is, that they are each comprised of an amino-terminal nonrepetitive toxic portion and a carboxy-terminal repetitive ligand region. Eichel-Streiber et al. (1992b) compared the predicted amino acid sequence of toxin A and toxin B with the glucosyltransferases GTF-I and GTF-SI from *Streptococcus mutans* and GTF-I and GTF-S from *S. downei*. The sequences showing similarity mapped in the carboxy-terminal repeat region of each of the proteins. Interestingly, results from experiments which were directed at deletion of this repeat region in the streptococcal glucosyltransferases indicate that this repeat region serves as a carbohydrate-binding region (Feretti et al., 1987; Kato and Kuramitsu, 1990).

Toxins A and B are immunologically distinct molecules. Polyclonal antibodies to toxin A and toxin B do not cross-react. Antibodies against toxin A are not capable of neutralizing the cytotoxic activity of toxin B. Similarly, antibodies to toxin B do not neutralize the cytotoxic and enterotoxic activity of toxin A (Libby and Wilkins, 1982).

## CLONING OF THE TOXIN A GENE

Initially, Muldrow et al. (1987) reported cloning of a fragment of the toxin A gene in the bacteriophage expression vector lambda gt11. The cloned *Taq*I fragment of Muldrow et al. (1987) was small, 0.3 kb; however, the expressed peptide reacted with polyvalent antiserum to toxin A. In addition, when the cloned fragment was used as a labelled hybridization probe, it hybridized to a *Pst*I fragment of *Clostridium difficile* genomic DNA which was estimated to be 4.5 kb. Wren et al. (1987a) also reported a lambda gt11 clone of toxin A. Their expressed polypeptide had an estimated molecular weight of 235,000. Both the clones of Muldrow et al. and Wren et al. proved unstable in *Escherichia coli* host cells (Eichel-Streiber, 1993).

Price et al. (1987) were successful in cloning 4.7 kb of the toxin A gene. Genomic DNA from *C. difficile* strain 10463 was digested with the restriction endonuclease *Pst*I and ligated into the vector pBR322. *Escherichia coli* Chi 1776 served as the host. From the 2,000 colonies which were screened, five reacted with affinity-purified antibody to toxin A. The five clones were identical, each containing a 4.7 kb fragment of the toxin A gene. One recombinant clone was chosen for characterization and designated pCD11. Cell lysates and supernatant fluid from the recombinant clones were not toxic, either in cytotoxicity assays with Chinese Hamster Ovary cells or in enterotoxicity assays with hamsters; however, the

same lysates reacted with the monoclonal antibody to toxin A, PCG-4, in an enzyme-linked immunosorbent assay (ELISA) (Price et al., 1987). Furthermore, there was partial identity between the protein expressed by the recombinant clone and purified toxin A in an Ouchterlony assay (Price et al., 1987). The *Pst*I fragment identified by Muldrow et al. (1987) was apparently the same fragment as that which was cloned by Price et al. (1987). Chromosome walking techniques were employed to clone the remainder of the toxin A gene utilizing the pCD11 clone as a probe (Johnson et al., 1988; Dove et al., 1990). The five overlapping clones covered a 15 kb region and included the toxin A gene in its entirety. The (5') clone contained 1.2 kb of the toxin B gene, a small open reading frame (approximately 500 bp) located between the toxin B and A genes, and the 5' end of the toxin A gene. The clone containing the 3' portion of the toxin A gene included approximately 4.1 kb of sequence downstream of the toxin A gene. In addition to the 3' end of the toxin A gene, two small open reading frames and a portion of a third open reading frame were contained within this sequence. DNA sequencing of the primary clones yielded an open reading frame for toxin A which was 8133 nucleotides in length (Dove et al., 1990).

Eichel-Streiber et al. (1989) generated a library from *C. difficile* strain 10463 genomic DNA that had been digested with both *Alu*I and *Hae*III restriction endonucleases. Fragments in the size range of 3-8 kb were selected and cloned into the vector pUC12. The library was screened with polyclonal antiserum to toxin A.

Positive clones identified in this manner and an additional clone identified by chromosome walking techniques encompassed the complete toxin A gene (Eichel-Streiber, 1990). The toxin A gene was also sequenced by this group (Sauerborn and Eichel-Streiber, 1990). The deduced amino acid sequences for toxin A from the two groups are identical.

The toxin A gene is 8133 nucleotides in length, 26.9 % G + C, and encodes a polypeptide of 2710 amino acid residues (Dove et al., 1990). The deduced protein has a molecular weight of 308,103 (Dove et al., 1990). Results from an amino-terminal amino acid analysis of toxin A are identical with the amino acid sequence deduced from DNA sequencing, indicating that processing does not occur in this portion of the protein (Dove et al., 1990). Likewise, a hydropathy plot derived for the deduced protein provides no evidence for a signal peptide at the amino-terminal end of toxin A (Dove et al., 1990). Conversely, Eichel-Streiber and Sauerborn (1990) identified a hydrophobic region at the carboxy-terminus of toxin A and point out the potential for a carboxy-terminal signal for export.

Both Dove et al. (1990) and Eichel-Streiber and Sauerborn (1990) analyzed the untranslated upstream region of toxin A for regulatory sequences. Both groups identified a Shine-Dalgarno sequence (AGGAGG), the first nucleotide of which is located 11 bases upstream of the ATG initiation codon. Through sequence inspection Eichel-Streiber and Sauerborn (1990) also proposed promoter and transcription termination signals for the toxin A gene.

Approximately one third of toxin A is comprised of contiguous, repeating DNA sequences located at the 3' end of the gene. Dove et al. (1990) designated 38 repeat units which they assigned to one of two classes, I or II, based on DNA sequence similarity. There are seven class I repeat units which are 90 nucleotides in length. The 31 members of class II are further divided into four subclasses. With one exception, class II repeat units are either 60 or 63 nucleotides long. A comparison of deduced amino acid sequences for the class I repeat units reveals that seventy percent of the amino acids are conserved among the seven class I units. The two amino acids tyrosine and phenylalanine are found in all thirty eight of the repeat units at the same position within the members of each class. The repeating units are arranged in such a manner that one class I combined with either three, four, or five class II repeating units comprise a large repeating unit.

The classification scheme for the repeating units derived by Eichel-Streiber et al. (1990a) differs from the Dove et al. (1990) classification in the number of repeating units. There are fewer repeating units in the Eichel-Streiber et al. (1990a) classification due to the fact that sequences of one of the five classes designated by Eichel-Streiber et al. as ALICE are comprised of contiguous class I and class IIA sequences of Dove et al. (1990) linked together. In addition, the 3' repeating unit $IID_6$ of Dove et al. (1990) is eliminated from the classification of Eichel-Streiber et al. (1990a).

The repeating units in the carboxy-terminus of toxin A are significant in that

several of the properties of the toxin can be attributed to this region. The recombinant protein expressed by the 3' end of the toxin A gene binds to the trisaccharide sequence, Galα1,3Galβ1,4GlcNAc (Price et al., 1987). This carbohydrate sequence is found on brush border membranes in hamsters, experimental animals which are extremely sensitive to toxin A. Krivan et al. (1986) have demonstrated that this carbohydrate moiety is a receptor for toxin A. The monoclonal antibodies, PCG-4 (Lyerly et al., 1986b) and 1337C8 (Eichel-Streiber et al., 1987), inhibit the binding of toxin A to this trisaccharide sequence. Furthermore, both the PCG-4 antibody and the 1337C8 antibody neutralize the enterotoxic activity of toxin A (Lyerly et al., 1986b; Eichel-Streiber, 1993). In addition, Tucker and Wilkins (1991) have identified carbohydrate antigens on human intestinal epithelial cells which are similar in structure to the Galα1,3Galβ1,4GlcNAc moiety which binds toxin A. These data point to the carboxy-terminal repeat region of toxin A as the binding or ligand portion of the toxin which interacts with the cellular receptor for the toxin.

## CLONING OF THE TOXIN B GENE

The 5' clone from the cloning of toxin A, pCD19, contained both a small open reading frame upstream of the toxin A gene and a partial open reading frame (Dove et al., 1990). The protein expressed from the partial open reading frame

reacted with antibodies to toxin B (Johnson et al., 1990). Johnson et al. (1990), utilizing chromosome walking techniques and a fragment of pCD19 as a hybridization probe, cloned a 6.8 kb *Xba*I / *Sph*I fragment which they designated pCD19L. The lysates from the *Escherichia coli* host cells which contained the recombinant plasmid were both cytotoxic and lethal in the respective assays and were neutralized by antibody to toxin B (Johnson et al., 1990). The two clones pCD19 and pCD19L contained the toxin B gene in its entirety and the DNA sequence for the toxin B gene was completed by this group (Barroso et al., 1990).

Schulze and Eichel-Streiber (1990) designed oligonucleotide probes based on a published amino-terminal amino acid sequence of toxin B (Meador and Tweten, 1988). Through a combination of techniques which included oligonucleotide screening of a genomic library and chromosome walking, they produced three clones encompassing the toxin B gene that were subsequently sequenced (Eichel-Streiber et al., 1990b). The toxin B sequences from both groups are identical (Eichel-Streiber, 1993).

The toxin B gene is 7098 bp in length, 27.4% G + C, and encodes a polypeptide of 2366 amino acid residues. The deduced protein has a molecular weight of 269,696 (Barroso et al., 1990).

The 5' untranslated region of toxin B was examined for the presence of regulatory signals. The first nucleotide of a Shine-Dalgarno sequence (AGGAGA) is positioned 13 nucleotides upstream of the ATG translation initiation codon. A

promoter-like structure and a potential transcription terminator stem-loop structure for toxin B were identified through DNA sequence inspection (Sauerborn and Eichel-Streiber, 1990; Eichel-Streiber et al., 1992; Eichel-Streiber, 1993).

## TOXINS A AND B IN STRAINS OF *CLOSTRIDIUM DIFFICILE*

The vast majority of investigations with toxigenic *Clostridium difficile* have focused on strain 10463 from the Anaerobe Collection, Department of Biochemistry and Anaerobic Microbiology, Virginia Polytechnic Institute and State University. This strain produces high titers of both toxins A and B, as analyzed by ELISA and cytotoxicity assays, respectively, and the genes are located in close proximity (less than 1.5 kb from each other) on the chromosome.

A large number of additional strains of *Clostridium difficile* have been examined for the presence or absence of the toxin genes and proteins. Lyerly et al. (1983) and Laughon et al. (1984) surveyed a large number of toxigenic isolates using assays for the two toxin proteins. In both studies, all toxigenic isolates produced both toxin A and B.

Price et al. (1987), using a fragment of the toxin A gene as a hybridization probe, showed that the toxin A gene was present in six toxigenic strains and absent in the four nontoxigenic strains examined. Fluit et al. (1991), using oligonucleotide primers for toxin A taken from the sequence which precedes the repetitive region of

the gene and for toxin B from the 5' end of the gene, generated amplification products for the two genes from a number of strains. In all thirty nine strains which were toxigenic, as determined by cytotoxicity assay, amplification products for both toxin A and B were found. In the twenty strains which were nontoxigenic, neither amplification product was found.

Similarly, McMillin et al. (1991) using both Polymerase Chain Reaction amplification and DNA hybridization examined a number of toxigenic and nontoxigenic strains for the presence of the toxin A and B genes. The DNA fragments specific for the toxins were present in all thirty seven toxigenic strains and absent in nine of ten nontoxigenic strains. One strain was not cytotoxic by the assay criteria used by this group; however, it appeared to contain toxin-specific fragments.

A single strain of *Clostridium difficile*, CCUG 8864, has been identified which produces only toxin B; however, this strain retains the capability to elicit diarrhea and death in antibiotic-treated hamsters (Torres, 1991; Borriello et al., 1992). Lyerly et al. (1992) present DNA hybridization data which indicate that the 5' end of the toxin A gene is present in strain 8864 and that the 3' end of the toxin A gene is absent.

Wren et al. (1987b) examined 172 strains of *C. difficile* for production of toxin A and B. The toxins were detected in 144 of these strains and these strains were classified as either high or low toxin producers. With only a few exceptions, the strains included in their study produced correspondingly high or low levels of

both toxin A and B. Lyerly and Wilkins (1986) also showed a positive correlation between toxin A production and toxin B production among toxigenic strains.

## REGULATION OF THE TOXIN GENES

Toxigenic strains of *Clostridium difficile* can vary in toxin production over a range of $10^6$ (Lyerly and Wilkins, 1986). The regulatory mechanisms involved in this variation are unknown. One possibility is that there are sequence differences in essential regions of DNA among these strains. Variation in the coding regions for the toxin genes could be responsible for differences in toxigenicity. Conversely, DNA sequence differences in intergenic regulatory regions such as promoter regions of the toxin genes might account for the observed variation in toxigenicity.

Results from several investigations (Lyerly and Wilkins, 1986; Wren et al., 1987b) indicate that strains which produce large amounts of one toxin also produce large amounts of the other toxin. Likewise, strains which produce relatively little of one toxin produce correspondingly little of the other toxin (Lyerly and Wilkins, 1986; Wren et al., 1987b). These data indicate that the toxin genes are coregulated.

There are several possibilities whereby the toxin genes might be coregulated. The toxin genes could be cotranscribed as a single unit or operon and therefore be

under the control of the same promoter. Alternatively, the toxin genes could be transcribed individually, each with its own promoter, but subject to coordinate regulation by the same trans-acting factor.

Trans-acting factors are proteins which exert regulatory effects on genes at some distance from their own locus. They can act as either positive or negative regulators of gene expression. Trans-acting factors are well-documented as regulators of expression of bacterial virulence genes (DiRita and Mekalanos, 1989; Miller et al., 1989; Mekalanos, 1992).

*Vibrio cholerae* is an example of a bacterium in which several trans-acting factors coordinately regulate expression of virulence genes. The Tox R protein is a trans-acting factor which controls the expression of a set of genes in *V. cholerae* designated the Tox R regulon, which includes the cholera toxin operon, *ctx AB* (Miller et al., 1987). Together with the Tox S protein, the Tox R protein binds to a DNA sequence which is repeated within the promoter for *ctx AB*, and activates transcription of the cholera toxin genes (Miller et al., 1987; DiRita and Mekalanos, 1991). Expression of many of the other genes in the Tox R regulon is only indirectly controlled by Tox R. The Tox R protein activates transcription of *tox T* (DiRita et al., 1991). It is the Tox T protein that then serves as a second trans-acting regulatory factor, activating transcription of the set of genes under Tox R control (DiRita et al., 1991).

## RESEARCH GOALS

My research focused on elucidating those DNA sequences which are unique to toxigenic strains of *Clostridium difficile* and determining whether variation exists among these DNA sequences that might account for the large variation in toxin A and B production within a set of toxigenic strains.

In order to determine the junction between the sequences specific for toxigenic strains and sequences held in common by both toxigenic and nontoxigenic strains, DNA fragments that overlap the toxin genes were used as probes in DNA hybridization analyses of toxigenic and nontoxigenic strain DNA. Hybridization analysis showed regions shared by both toxigenic and nontoxigenic strains downstream from the toxin genes. The corresponding region of nontoxigenic DNA was cloned and sequenced, and the toxigenic and nontoxigenic sequences were aligned to precisely determine the downstream boundary. There was no such region held in common by both toxigenic and nontoxigenic strains upstream of the toxin genes. Therefore, chromosome walking techniques were employed to clone regions in this direction. Each newly cloned region was analyzed by DNA hybridization to determine whether it was unique to toxigenic or shared by both types of strains. Sequences common to both nontoxigenic and toxigenic strains were identified upstream of the toxin genes and the entire region was sequenced and aligned with nontoxigenic sequence to identify an upstream boundary.

Once established, the boundaries were examined in a number of strains which vary in toxigenicity to determine whether they were the same for each strain. The boundaries of the toxigenic element were examined for evidence of insertion sequences. In addition, the boundary data permitted identification of the open reading frames and adjacent sequences that are unique to the toxigenic element.

Transcription analyses were undertaken to examine transcription of the genes for toxin A and toxin B as well as the remaining open reading frames which comprise the toxigenic element. Transcription initiation sites were determined and regulatory regions identified.

RNA from a number of toxigenic strains which vary in toxigenicity was analyzed to determine whether quantitative differences in toxin-specific transcripts exist among these strains. Quantitative differences were found among these strains, and the DNA sequence for the regulatory regions of the toxin genes in these strains was determined and compared.

## LITERATURE CITED

1. Allo, M., J. Silva, Jr., R. Fekety, G. D. Rifkin, and H. Waskin. 1979. Prevention of clindamycin-induced colitis in hamsters by *Clostridium sordellii* antitoxin. Gastroenterology 76: 352-355.

2. Arnon, S. S., D. C. Mills, P. A. Day, R. V. Henrickson, N. M. Sullivan, and T. D. Wilkins. 1984. Rapid death of infant rhesus monkeys injected with *Clostridium difficile* toxins A and B: physiologic and pathologic basis. J. Pediatr. 101: 34-40.

3. Banno, Y., T. Kobayashi, H. Kono, K. Watanabe, K. Ueno, and Y. Nozawa. 1984. Biochemical characterization and biologic actions of two toxins (D-1 and D-2) from *Clostridium difficile*. Rev. Infect. Dis. 6: S11-S20.

4. Banno, Y., T. Kobayashi, K. Watanabe, K. Ueno, and Y. Nozawa. 1981. Two toxins (D-1 and D-2) of *Clostridium difficile* causing antibiotic-associated colitis: purification and some characterization. Biochem. Int. 2: 629-635.

5. Barroso. L. A., S.-Z. Wang, C. J. Phelps, J. L. Johnson, and T. D. Wilkins. 1990. Nucleotide sequence of the *Clostridium difficile* toxin B gene. Nucleic Acids Res. 18: 4004.

6. Bartlett, J. 1979. Antibiotic-associated pseudomembranous colitis. Rev. Infect. Dis. 1: 370-378.

7. Bartlett, J. G., T. W. Chang, M. Gurwith, S. L. Gorbach, and A. B. Onderdonk. 1978. Antibiotic-associated pseudomembranous colitis due to toxin-producing clostridia. N. Engl. J. Med. 298: 531-534.

8. Bartlett, J. G., T. W. Chang, N. S. Taylor, and A. B. Onderdonk. 1979. Colitis induced by *Clostridium difficile*. Rev. Infect. Dis. 1: 370-378.

9. Borriello, S. P. and A. R. Welch. 1984. Detection of *Clostridium difficile* toxins. pp. 49-56, *In* S. P. Borriello (ed.), Antibiotic-associated diarrhoea and colitis. Martinus Nijhoff, Boston, MA

10. Borriello, S. P., B. W. Wren, S. Hyde, S. V. Seddon, P. Sibbons, M. M. Krishna, S. Tabaqchali, S. Manek, and A. B. Price. 1992. Molecular, immunological, and biological characterization of a toxin A negative, toxin B positive strain of *Clostridium difficile*. Infect. Immun. 60: 4192-4199.

11. Chang, T.-W., S. L. Gorbach, and J. G. Bartlett. 1978. Neutralization of *Clostridium difficile* toxin by *Clostridium sordellii* antitoxins. Infect. Immun. 22: 418-422.

12. DiRita, V. J. and J. J. Mekalanos. 1989. Genetic regulation of bacterial virulence. Ann. Rev. Genet. 23: 455-482.

13. DiRita, V. J. and J. J. Mekalanos. 1991. Periplasmic interaction between two membrane regulatory proteins, ToxR and ToxS, results in signal transduction and transcriptional activation. Cell 64: 29-37.

14. DiRita, V. J., C. Parsot, G. Jander, and J. J. Mekalanos. 1991. Regulatory cascade controls virulence in *Vibrio cholerae*. Proc. Natl. Acad. Sci. USA 88: 5403-5407.

15. Dove, C. H., S.-Z. Wang, S. B. Price, C. J. Phelps, D. M. Lyerly, T. D. Wilkins, and J. L. Johnson. 1990. Molecular characterization of the *Clostridium difficile* toxin A gene. Infect. Immun. 58: 480-488.

16. Eichel-Streiber, C. v. 1992. A dual model for the architecture of *Clostridium difficile* toxins A and B. pp.113-121, *In* Witholt et al. (eds.) Bacterial protein toxins. Zentralbl. Bakteriol. Suppl. 23. Gustav Fischer, Stuttgart.

17. Eichel-Streiber, C. v. 1993. Molecular biology of the *Clostridium difficile* toxins. pp. 264-289, *In* M. Sebald (ed.) Genetics and molecular biology of anaerobic bacteria. Springer-Verlag, New York.

18. Eichel-Streiber, C. v., U. Harperath, D. Bosse, and U. Hadding. 1987. Purification of two high molecular weight toxins of *Clostridium difficile* which are antigenically related. Microb. Pathogen. 2: 307-318.

19. Eichel-Streiber, C. v., R. Laufenberg-Feldmann, S. Sartingen, J. Schulze, and M. Sauerborn. 1990b. Cloning of *Clostridium difficile* toxin B gene and demonstration of high N-terminal homology between toxin A and B. Med. Microbiol. Immuno. 179: 271-279.

20. Eichel-Streiber, C. v., R. Laufenberg-Feldmann, S. Sartingen, J. Schulze, and M. Sauerborn. 1992a. Comparative sequence analysis of the *Clostridium difficile* toxins A and B. Mol. Gen. Genet. 233: 260-268.

21. Eichel-Streiber, C. v., R. Laufenberg-Feldmann, and M. Sauerborn. 1990a. The analysis of the *Clostridium difficile* toxin A reveals the 3' end to be composed of a 2499 bp repetitive structure. pp. 115-116, *In* Bacterial protein toxins. Fischer-Verlag, Stuttgart.

22. Eichel-Streiber, C. v. and M. Sauerborn. 1990. *Clostridium difficile* toxin A carries a C-terminal repetitive structure homologous to the carbohydrate binding region of streptococcal glycosyltransferases. Gene 96: 107-113.

23. Eichel-Streiber, C. v., M. Sauerborn, and H. K. Kuramitsu. 1992b. Evidence for a modular structure of the homologous repetitive C-terminal carbohydrate-binding sites of *Clostridium difficile* toxins and *Streptococcus mutans*

glycosyltransferases. J. Bacteriol. 174: 6707-6710.

24. Eichel-Streiber, C. v., D. Suckau, M. Wachter, and U. Hadding. 1989. Cloning and characterization of overlapping DNA fragments of the toxin A gene of *Clostridium difficile*. J. of Gen. Microbiol. 135: 55-64.

25. Ferretti, J. J., M. L. Glipin, and R. R. B. Russell. 1987. Nucleotide sequence of a glucosyltransferase gene from *Streptococcus sobrinus* MFe28. J Bacteriol. 69: 4271-4278.

26. Finney, J. M. 1893. Gastro-enterostomy for cicatrizing ulcer of the pylorus. Johns Hopkins Hosp. Bull. 11: 53-55.

27. Fluit, A. C., M. J. H. M. Wolfhagen, G. P. H. T. Verdonk, M. Jansze, R. Torensma, and J. Verhoef. 1991. Nontoxigenic strains of *Clostridium difficile* lack the genes for both toxin A and toxin B. J. Clin. Microbiol 29: 2666-2667.

28. George, R. H., J. M. Symonds, F. Dimock, J. D. Brown, Y. Arabi, N. Shinagawa, M. R. B. Keighley, J. Alexander-Williams, and D. W. Burdon. 1978. Identification of *Clostridium difficile* as a cause of pseudomembranous colitis. Br. Med. J. 1: 695.

29. George, W. L., E. J. C. Goldstein, V. L. Sutter, S. L. Ludwig, and S. M. Finegold. 1978b. Etiology of antimicrobial-agent-associated colitis. Lancet i: 802-803.

30. Gurwith, M. J., H. R. Rabin, and K. Love. 1977. Diarrhea associated with clindamycin and ampicillin therapy: preliminary results of a cooperative study. J. Infect. Dis. 135: S104-S110.

31. Humphrey, C. D., C. W. Condon, J. R. Cantey, and F. E. Pittman. 1979. Partial purification of a toxin found in hamsters with antibiotic-associated colitis. Gastroenterology 76: 478-486.

32. Johnson, J. L., C. H. Dove, S. B. Price, T. W. Sickles, C. J. Phelps, and T. D. Wilkins. 1988. The toxin A gene of *Clostridium difficile*. pp. 115-123. *In* J. M. Hardie and S. P. Borriello (eds.), Anaerobes today. John Wiley & Sons Ltd.

33. Johnson, J. L., C. Phelps, L. Barroso, M. D. Roberts, D. M. Lyerly, and T. D. Wilkins. 1990. Cloning and expression of the toxin B gene of *Clostridium difficile*. Curr. Microbiol. 20: 397-401.

34. Kato, C. and H. K. Kuramitsu. 1990. Carboxyl-terminal deletion analysis of the *Streptococcus mutans* glucosyltransferase-I enzyme. FEMS Microbiol. Lett. 72: 299-302.

35. Krivan, H. C., G. F. Clark, D. F. Smith, and T. D. Wilkins. 1986. Cell surface binding site for *Clostridium difficile* enterotoxin: evidence for a glycoconjugate containing the sequence Galα1-3Galβ1-4GlcNAc. Infect. Immun. 53:573-581.

36. Krivan, H. C. and T. D. Wilkins. 1987. Purification of *Clostridium difficile* toxin A by affinity chromatography on immobilized thyroglobulin. Infect. Immun. 55:1873-1877.

37. Larson, H. E., P. Honour, A. B. Price, and S. P. Borriello. 1978. *Clostridium difficile* and the etiology of pseudomembranous colitis. Lancet i: 1063-1066.

38. Larson, H. E. and A. B. Price. 1977. Pseudomembranous colitis: presence of clostridial toxin. Lancet ii: 1312-1314.

39. Laughon, B. E., R. P. Viscidi, S. L. Gdovin, R. H. Yolken, and J. G. Bartlett. 1984. Enzyme immunoassays for detection of *Clostridium difficile* toxins A and B in fecal specimens. J. Infect. Dis. 149: 781-788.

40. Libby, J. M. and T. D. Wilkins. 1982. Production of antitoxins to two toxins of *Clostridium difficile* and immunological comparison of the toxins by cross-neutralization studies. Infect. Immun. 35: 374-376.

41. Lyerly, D. M., L. A. Barroso, T. D. Wilkins, C. Depitre, and G. Corthier. 1992. Characterization of a toxin A-negative, toxin B-positive strain of *Clostridium difficile*. Infect. Immun. 60: 4633-4639.

42. Lyerly, D. M., J. L. Johnson, and T. D. Wilkins. 1990. The toxins of *Clostridium difficile*. pp. 137-146, *In* S. P. Borriello (ed.) Clinical and molecular aspects of anaerobes. Wrightson Biomedical Publishing Ltd., Petersfield.

43. Lyerly, D. M., H. C. Krivan, and T. D. Wilkins. 1988. *Clostridium difficile*: Its disease and toxins. Clin. Microbiol. Reviews 1: 1-18.

44. Lyerly, D. M., C. J. Phelps, J. Toth, and T. D. Wilkins. 1986b. Characterization of toxins A and B of *Clostridium difficile* with monoclonal

antibodies. Infect. Immun. 54: 70-76.

45. Lyerly, D. M., M. D. Roberts, C. J. Phelps, and T. D. Wilkins. 1986a. Properties of toxins A and B of *Clostridium difficile*. FEMS Microbiol. Lett. 33:31-35.

46. Lyerly, D. M., K. E. Saum, D. MacDonald, and T. D. Wilkins. 1985. Effect of toxins A and B given intragastrically to animals. Infect. Immun. 47:349-352.

47. Lyerly, D. M., N. M. Sullivan, and T. D. Wilkins. 1983. Enzyme-linked immunosorbent assay for *Clostridium difficile* toxin A. J. Clin. Microbiol. 17:72-78.

48. Lyerly, D. M. and T. D. Wilkins. 1986a. *Clostridium difficile* toxins. pp. 197-209, *In* R. B. Kohler (ed.) Antigen detection to diagnose bacterial infections. CRC Press, Boca Raton, Florida.

49. Lyerly, D. M. and T. D. Wilkins. 1986b. Commercial latex test for *Clostridium difficile* toxin A does not detect toxin A. J. Clin. Microbiol. 23: 622-623.

50. Lyerly, D. M. and T. D. Wilkins. 1988. Purification and properties of toxins A and B of *Clostridium difficile*. pp. 146-167. *In* R. D. Rolfe and S. M. Finegold (eds.), *Clostridium difficile*: Its role in intestinal disease. Academic Press, Inc., San Diego, CA.

51. Martinez, R. D. and T. D. Wilkins. 1988. Purification and characterization of *Clostridium sordellii* hemorrhagic toxin and cross-reactivity with *Clostridium difficile* toxin A (enterotoxin). Infect. Immun. 56:1215-1221.

52. Martinez, R. D. and T. D. Wilkins. 1992. Comparison of *Clostridium sordellii* toxins HT and LT with toxins A and B of *C. difficile*. J. Med. Microbiol. 36: 30-36.

53. McFarland, L. M. E. Mulligan, R. Y. Y. Kwok, and W. E. Stamm. 1989. Nosocomial acquisition of *Clostridium difficile* infection. N. Engl. J. Med. 320: 204-210.

54. McMillin, D. E., L. L. Muldrow, S. J. Leggette, Y. Abdulahi, and U. M. Ekanemesang. 1991. Molecular screening of *Clostridium difficile* toxins A and B

genetic determinants and identification of mutant strains. FEMS Microbiol. Lett. 78: 75-80.

55. Meador, J., III, and R. K. Tweten. 1988. Purification and characterization of toxin B from *Clostridium difficile*. Infect. Immun. 56: 1708-1714.

56. Mekalanos, J. J. 1992. Environmental signals controlling expression of virulence determinants in bacteria. J. Bacteriol. 174: 1-7.

57. Miller, J. F., J. J. Mekalanos, and S. Falkow. 1989. Coordinate regulation and sensory transduction in the control of bacterial virulence. Science 243: 916-922.

58. Miller, V. L., R. K. Taylor, and J. J. Mekalanos. 1987. Cholera toxin transcriptional activator *toxR* is a transmembrane DNA binding protein. Cell 48: 271-279.

59. Mitchell, M. J., B. E. Laughon, and S. Lin. 1987. Biochemical studies on the effect of *Clostridium difficile* toxin B on actin in vivo and in vitro. Infect. Immun. 55: 1610-1615.

60. Muldrow, L. L., G. C. Ibeanu, N. I. Lee, N. K. Bose, and J. Johnson. 1987. Molecular cloning of *Clostridium difficile* toxin A gene fragment in lambda gt11. FEBS Letters 213: 249-253.

61. Mulligan, M. E., R. D. Rolfe, S. M. Finegold, and W. L. George. 1979. Contamination of a hospital environment by *Clostridium difficile*. Curr. Microbiol. 3: 173-175.

62. Pothoulakis, C., L. M. Barone, R. Ely, B. Faris, M. E. Clark, C. Franzblau, and J. T. LaMont. 1986. Purification and properties of *Clostridium difficile* toxin B. J. Biol. Chem. 261: 1316-1321.

63. Price, S. B., C. J. Phelps, T. D. Wilkins, and J. L. Johnson. 1987. Cloning of the carbohydrate-binding portion of the toxin A gene of *Clostridium difficile*. Curr. Microbiol. 16: 55-60.

64. Rifkin, G. D., R. Fekety, and J. Silva. 1978a. Neutralization by *Clostridium sordellii* antitoxin of toxins implicated in clindamycin-induced cecitis in the hamster. Gastroenterology 75: 422-424.

65. Rifkin, G. D., J. Silva, Jr., and F. R. Fekety. 1978b. Gastrointestinal and systemic toxicity of fecal extracts from hamsters with clindamycin-induced colitis. Gastroenterology 74: 52-57.

66. Rifkin, G. D., J. Silva, Jr., F. R. Fekety, and R. B. Sack. 1977. Antibiotic-induced colitis. Implification of a toxin neutralized by *Clostridium sordellii* antitoxin. Lancet ii: 1103-1106.

67. Rihn, B., J. M. Scheftel, R. Girardot, and H. Monteil. 1984. A new purification procedure for *Clostridium difficile* enterotoxin. Biochem. Biophys. Res. Commun. 124: 690-695.

68. Sauerborn, M. and C. v. Eichel-Streiber. 1990. Nucleotide sequence of *Clostridium difficile* toxin A. Nucleic Acids Res. 18: 1629-1630.

69. Schulze, J. and C. v. Eichel-Streiber. 1990. Cloning and characterization of the *Clostridium difficile* toxin B gene. Zentralbl. Bakteriol. Mikrobiol. Hyg. A 273: 126.

70. Sullivan, N. M., S. Pellett, and T. D. Wilkins. 1982. Purification and characterization of toxins A and B of *Clostridium difficile*. Infect. Immun. 35: 1032-1040.

71. Swartzberg, J. E., R. M. Maresca, and J. S. Remington. 1977. Clinical study of gastrointestinal complications associated with clindamycin therapy. J. Infect. Dis. 135: S99-S103.

72. Taylor, N. S., G. M. Thorne, and J. G. Bartlett. 1981. Comparison of two toxins produced by *Clostridium difficile*. Infect. Immun. 34: 1036-1043.

73. Thelestam, M. and M. Bronnegard. 1980. Interaction of cytopathogenic toxin from *Clostridium difficile* with cells in tissue culture. Scand. J. Infect. Dis., Suppl. 22: 16-29.

74. Thelestam, M. and I. Florin. 1984. Cytopathogenic action of *Clostridium difficile* toxins. J. Toxicol., Toxin Rev. 3: 139-180.

75. Torres, J. F. 1991. Purification and characterization of toxin B from a strain of *Clostridium difficile* that does not produce toxin A. J. Med. Microbiol. 35: 40-44.

76. Tucker, K. D. and T. D. Wilkins. 1991. Toxin A of *Clostridium difficile* binds to the human carbohydrate antigens I, X, and Y. Infect. Immun. 59: 73-78.

77. Wedel, N., P. Toselli, C. Pothoulakis, B. Faris, P. Oliver, C. Franzblau, and T. LaMont. 1983. Ultrastructural effects of *Clostridium difficile* toxin B on smooth muscle cells and fibroblasts. Exp. Cell. Res. 148: 413-422.

78. Wren, B. W., C. L. Clayton, P. P. Mullany, and S. Tabaqchali. 1987a. Molecular cloning and expression of *Clostridium difficile* toxin A in *Escherichia coli* K12. FEBS Lett. 225: 82-86.

79. Wren, B., S. R. Heard, and Soad Tabaqchali. 1987b. Association between production of toxins A and B and types of *Clostridium difficile*. J. Clin. Pathol. 40: 1397-1401.

# CHAPTER ONE: The Toxigenic Element of *Clostridium difficile* Strain 10463

## INTRODUCTION

*Clostridium difficile* is a Gram positive, anaerobic bacterium, responsible for pseudomembranous colitis and a significant proportion of antibiotic-associated diarrhea. *C. difficile* produces two very large toxin proteins designated A and B which are held to be largely responsible for the disease symptoms (Lyerly et al., 1985; Wilkins and Lyerly, 1990).

Both the genes for toxin A (Price et al., 1987; Johnson et al., 1988; Eichel-Streiber et al., 1989; Dove et al., 1990; Sauerborn and Eichel-Streiber, 1990) and toxin B (Barroso et al., 1990; Eichel-Streiber et al., 1990; Johnson et al., 1990) have been cloned and sequenced from strain 10463. The genes for toxins A and B lie within 1.4 kb of each other with the gene for toxin B situated upstream (5') to the gene for toxin A. A small open reading frame, translated in the same direction as toxins B and A, is located in the region between the two genes (Dove et al., 1990; Eichel-Streiber and Sauerborn, 1990). Additional open reading frames, translated in the opposite direction and located downstream of the toxin A gene, have been identified (Dove et al., 1990; Eichel-Streiber et al., 1992; Eichel-Streiber, 1993).

Johnson et al. (1988) provided evidence that the location of the toxin genes is chromosomal. Nontoxigenic strains of *C. difficile* lack the genes for toxin A and B

(Fluit et al., 1991). With some rare exceptions, toxigenic strains produce both toxin A and B (McMillan et al., 1991; Torres, 1991; Borriello et al., 1992; Lyerly et al., 1992). There is considerable variation in toxin production among toxigenic strains; however, toxigenic strains which produce relatively high levels of toxin A also produce high levels of toxin B. Likewise, strains which produce low levels of one toxin produce correspondingly low levels of the other toxin (Lyerly and Wilkins, 1986; Wren et al., 1987).

The purpose of the present study is to determine the junction between the toxigenic sequences and the remainder of the chromosome in *C. difficile* strain 10463 by establishing both upstream (5') and downstream (3') boundaries for the sequences which are unique to the toxigenic strain. The toxigenic element is characterized as to the number and location of open reading frames associated with the toxin genes. Additional strains were examined to determine variations among their respective toxigenic elements.

## MATERIALS AND METHODS

### Bacterial strains, plasmids, bacteriophages, and growth conditions

*Clostridium difficile* strains were obtained from either the Anaerobe Collection of the Department of Biochemistry and Anaerobic Microbiology, or from

the laboratory collection of Dr. Tracy Wilkins, Department of Biochemistry and Anaerobic Microbiology, Virginia Polytechnic Institute and State University, Blacksburg, VA.

| Toxigenic | Toxicity | Nontoxigenic |
|---|---|---|
| 10463 | $10^6$ | 2037 |
| 26689 | $10^5$ | 11186 |
| 7698 | $10^3$ | 5199 |
| 11011 | $10^2$ | |
| Tox R | $10^2$ | |
| 4474 | $10^1$ | |

Toxicity values for toxigenic strains are the cytotoxicity titers for toxin B taken from Lyerly and Wilkins (1986). The *C. difficile* strains were grown anaerobically in peptone, yeast extract, glucose (PYG) medium (Anaerobe Laboratory Manual) supplemented with 1% dehydrated brain heart infusion, BHI, (Difco) at 37°C.

The plasmids pUC18 and pUC19, obtained from Bethesda Research Laboratories (BRL), served as vectors for cloning genomic DNA (Yanisch-Perron et al., 1985). *E. coli* DH5α obtained from BRL was used as the host for the cloning experiments and was grown aerobically at 37°C in Terrific Broth (TB) medium (Tartof and Hobbs, 1987) containing 200 μg/ml ampicillin.

The bacteriophage M13mp19 was utilized as the vector for cloning DNA for sequencing. *E. coli* DH5αF' obtained from BRL was used as the host with the M13mp19 vector and was grown aerobically at 37°C in 2XYT, yeast extract, tryptone medium (Sambrook et al., 1989).

DNA isolation and manipulations

Genomic DNA of *C. difficile* was isolated by the method of Marmur (1961) as modified by Johnson (1991). Plasmid isolation from *E. coli* followed the alkaline procedure of Birnboim and Doly (1979).

Restriction endonucleases were purchased from either Promega or BRL and used as described in Sambrook et al. (1989). Calf intestine alkaline phosphatase was purchased from Boehringer Mannheim Biochemicals and used according to the manufacturer's instructions. T4 DNA ligase was obtained from either BRL or Promega and was used according to the method of Struhl (1985). The source for the Klenow fragment of DNA Polymerase I, Exonuclease III, and Exonuclease VII was BRL.

DNA was electrophoresed on 11 X 14 cm 0.7% agarose (SeaKem LE, FMC Bioproducts) gels in Tris, acetate, EDTA (TAE) buffer (Sambrook et al., 1989).

DNA fragments were electroeluted using Schleicher and Schuell NA-45 DEAE membrane following the manufacturer's instructions.

DNA fragments were labelled as hybridization probes using the nonradioactive ECL Gene Detection System from Amersham, or $\alpha^{35}$S-dATP, 1000 Ci/mMol from DuPont NEN Research Products and either the Prime-It random primer labelling kit from Stratagene or the Prime Time random primer labelling kit from IBI Technologies.

Southern transfer followed the method of Southern (1975), using 0.45 μm nitrocellulose membranes from Schleicher and Schuell. The buffer for hybridization of Southern transfer membranes with radioactively labelled probes consisted of the following components in their final concentrations: 25% deionized formamide, 0.5% Sodium dodecyl sulfate (SDS), 5X Denhardt's Reagent (Sambrook et al., 1989), 6X Standard Saline-Phosphate-EDTA (SSPE) (Sambrook et al., 1989). Chemicals were obtained from Fisher Scientific or Sigma. Hybridizations were carried out overnight at 50°C. Posthybridization buffer consisted of 2X SSPE, 0.5% SDS heated to 50°C. Initial posthybridization washes were carried out at 50°C on a slowly moving shaker (50-70 rpm). Final posthybridization washes were at room temperature. Membranes were dried and exposed to Kodak SB5 X-ray film. Hybridizations to screen colony lift membranes followed this same set of procedures. Hybridizations with the ECL Gene Detection System followed the manufacturer's directions.

For the cloning of genomic fragments, restriction endonuclease-digested genomic DNA was electrophoresed on 0.7% SeaPlaque Low Melting Temperature agarose (FMC Bioproducts) gels and regions of the appropriate size were cut out and utilized in ligation reactions (Struhl, 1985). Ligation reactions were liquefied at 65°C for 15 min. and diluted 1:5 with distilled water. A twenty five μl aliquot of competent *E. coli* DH5α cells (Dower, 1988) was transformed with 1 μl of the diluted ligation mixture by electroporation, using a BTX Transfector 100 from Biotechnologies and Experimental Research, Inc. A 0.5 mm gap electrode was used

and the apparatus was set to 650 Volts. The pulse was delivered for 5 millisec. Colonies were replica plated onto 82 mm Nitro ME nitrocellulose circles, 0.45 μm pore size, from Micron Separations, Inc. The bacterial colonies were lysed and the DNA fixed to the nitrocellulose according to Maniatis et al. (1982).

DNA sequencing

DNA fragments were cloned into M13mp19 following the procedures outlined by Amersham in the M13 Cloning and Sequencing Handbook. Sets of nested deletions for sequencing were generated by using Exonuclease III and Exonuclease VII, along with the Klenow fragment (Yanisch-Perron, 1985). Ligation, preparation of competent DH5αF' cells, transformation, plating of transformed cells, and preparation of single-stranded template for sequencing followed the guidelines in Amersham's M13 Cloning and Sequencing Handbook. The 17 bp universal M13 primer was used for single-stranded sequencing with T7 DNA Polymerase, Sequenase Version 1.0 from United States Biochemical (USB). Sequencing reactions followed the methods established for single-stranded sequencing with Sequenase by DuPont which is a modification of the Sanger DNA sequencing techniques (Sanger et al., 1977). Sequencing reactions were run on 6% polyacrylamide gels on the DuPont Genesis 2000 Sequencer. DNA sequences were analyzed with Kodak MacVector software. Database searches included GenBank,

EMBL, SWISSPROT, and the National Biomedical Research Foundation Protein Sequence Data Base.

PCR amplifications

Polymerase Chain Reaction (PCR) amplifications were performed on a GTC-1 Genetic Thermal Cycler from Precision Scientific using Taq DNA Polymerase from Perkin Elmer, and in general followed the method of Saiki et al. (1988). Oligonucleotide primers were synthesized on a 381A DNA Synthesizer from Applied Biosystems. The primers for the downstream boundary of the toxigenic element were the following: 3'ToxAS 5'-AGCCCCTGGGATATATGGC and NT2A 5'-ACTGAGTCACTTAATTACATCA; primers for the upstream boundary for the toxigenic element were as follows: HdHcAS 5'-CAGGCAAGTGTATGTATTATAC and Tox B Primer Ext. 5'-CATCTTCTTGAGTACGAAATC; primers for the region in between toxin B and toxin A were the following: TOXA3' 5'-CTATATGCGAGTTTTATTAAC and TOXB5' 5'-ATCCTGATACAGCTCAATTAG; primers for the insertion region were HdHCAS and NT2A.

Symmetric PCR amplifications were run for 40 cycles. The initial denaturation step was at 95°C for 1.5 min., followed by annealing at 45°C for 1 min. In the remaining 39 cycles the denaturation temperature was lowered to 94°C for 30

sec., followed by annealing at 45°C for 30 sec. The elongation step for all 40 cycles was at 70°C for a range of 1-3 min. based on the length of the PCR product.

Asymmetric amplification reactions, using the symmetric amplification product as the template, to generate single-stranded DNA for sequencing were run for 30 cycles. The initial denaturation step was at 95°C for 1.5 min., followed by annealing at 45°C for 1 min. Subsequently, denaturation was at 94°C for 30 sec. and annealing was at 45°C for 30 sec. A 70°C, 2 min. elongation step was maintained for all 30 cycles. The single-stranded asymmetric amplification products were purified on NACS 52 Ion-Exchange Resin minicolumns (BRL) prior to being used in sequencing reactions.

## RESULTS

### Cloning and sequencing the regions 5' and 3' to the toxin genes

In order to define the toxigenic element in *Clostridium difficile* strain 10463, it was necessary to establish the boundaries between sequences that were unique to the toxigenic strain and those that were shared by both the toxigenic strain and a nontoxigenic strain. The nontoxigenic strain *C. difficile* 2037 was chosen. I utilized fragments of the primary sequencing clones of the toxin genes in order to generate overlapping clones that permitted sequencing both upstream (5') and downstream

(3') of the toxin genes. An *Xba*I - *Hin*cII fragment from pCD19L (Barroso et al., 1990) was used to generate overlapping clones in the region upstream of the toxin B gene. This 2.1 kb fragment hybridized to Southern transfers of both *Hin*dIII and *Hin*cII digests of genomic DNA from toxigenic strain 10463. The 5.1 kb *Hin*dIII and the 5.2 kb *Hin*cII fragments were cloned into the vector pUC19 and are designated pCD21 and pCD23, respectively. A map of these clones relative to the toxin B gene is presented in Fig. 1. When pCD21 was used as a probe, it hybridized only to digests of toxigenic strain DNA (Fig. 2). However, pCD23 which overlaps pCD21 and extends upstream of the *Hin*dIII site (Fig. 1) hybridized to both toxigenic and nontoxigenic strain DNA (Fig. 3). Therefore, I proposed that a region centered around the *Hin*dIII site comprised the 5' or upstream boundary of the toxigenic element. The insert regions from both pCD21 and pCD23 were subcloned into M13mp19 for sequencing this upstream region.

A *Pst*I - *Hin*dIII fragment from pCD11R6, which extends downstream of the toxin A gene (Dove et al., 1990), hybridized to Southern transfers of *Hin*dIII digests of genomic DNA from both toxigenic and nontoxigenic strains (Fig. 4). The 3.1 kb *Hin*d III fragment from the nontoxigenic strain 2037 was cloned into pUC18 and designated pCDNT2. Restriction endonuclease mapping and partial sequencing of pCDNT2 (data not shown) allowed alignment of the sequence of pCDNT2 with the sequence downstream of toxin A from strain 10463 (unpublished data). There was identity between the two DNA sequences which ended abruptly at a position 1263

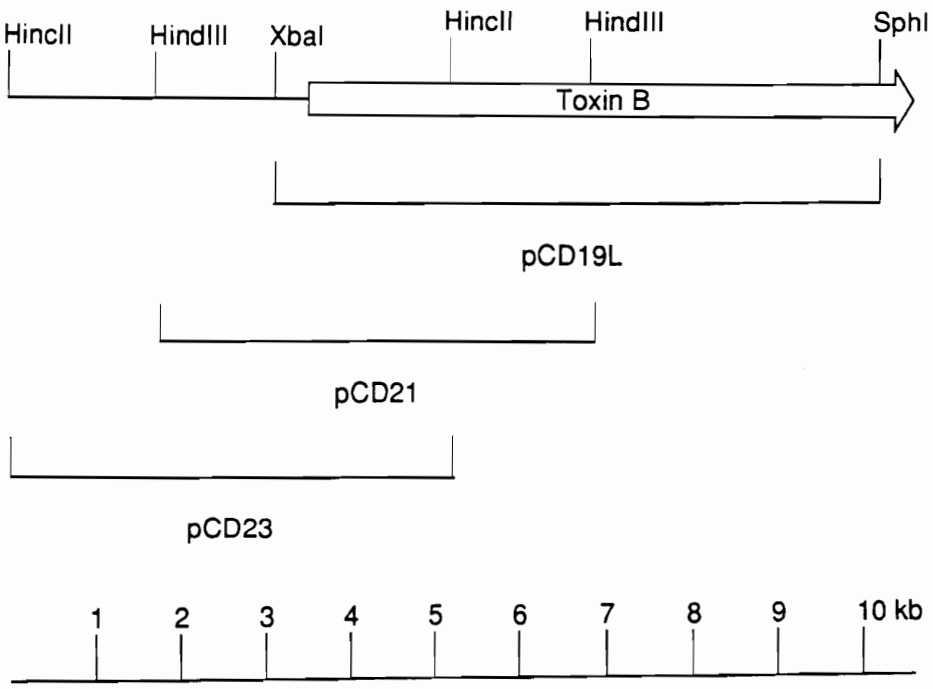


**FIG. 1. Restriction sites used in cloning and locations of the cloned 5.1 kb Hind III and the 5.2 kb Hinc II DNA fragments designated pCD21 and pCD23, respectively.** pCD19L is one of the clones for the toxin B gene (Johnson et al., 1990). Arrow indicates direction of transcription of toxin B gene.

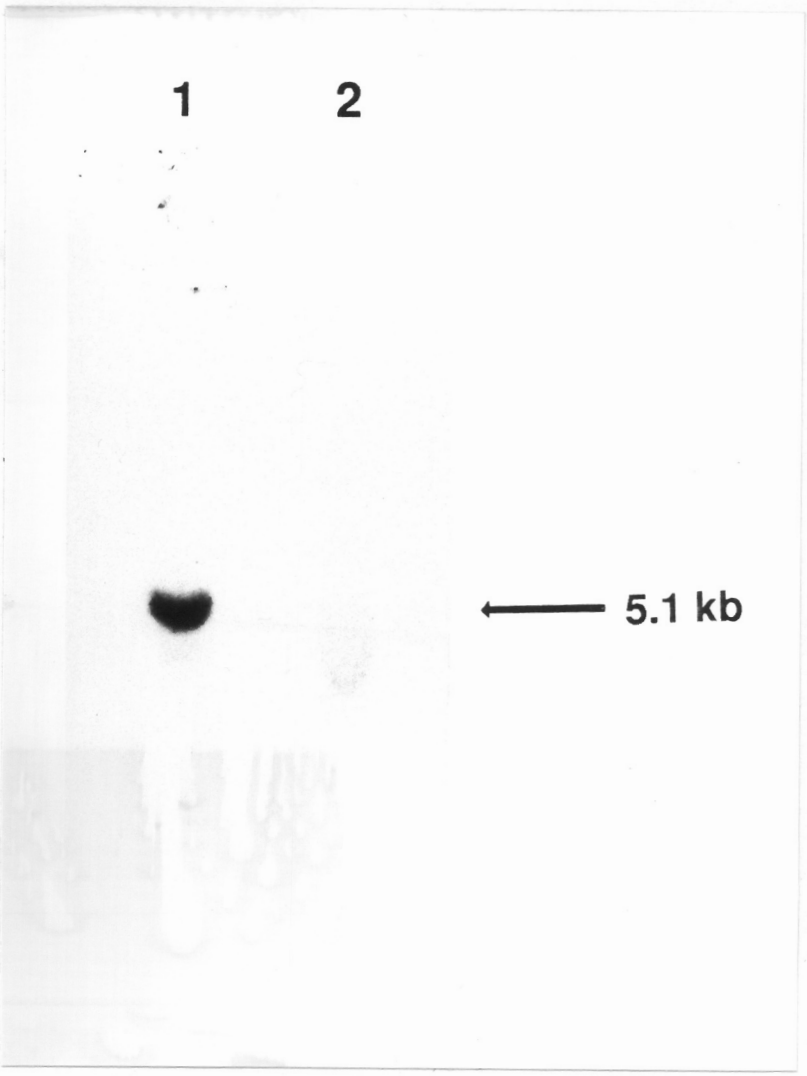


**FIG. 2. Southern blot analysis of restriction endonuclease digested *C. difficile* genomic DNA using pCD21 DNA as a probe.** Toxigenic strain 10463 DNA digested with Hind III (lane 1), nontoxigenic strain 2037 DNA digested with Hind III (lane 2). Size is indicated with arrow.

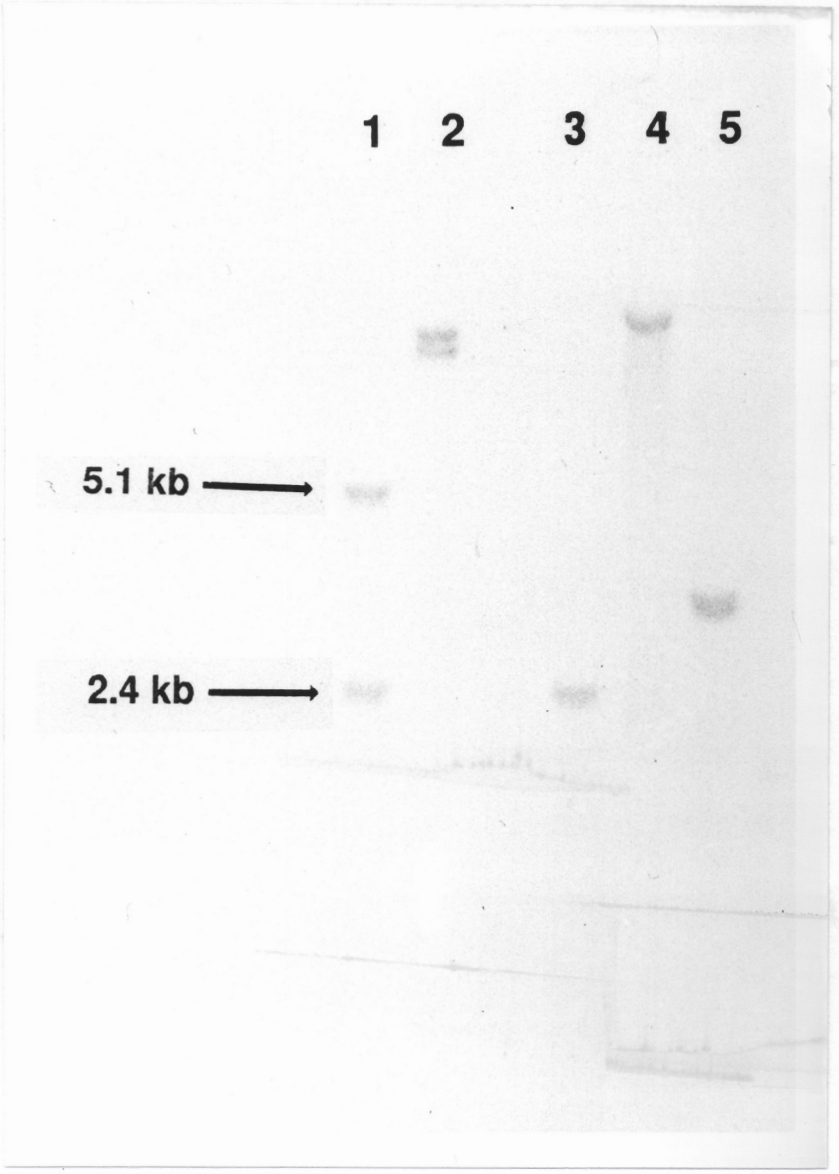


**FIG. 3. Southern blot analysis of restriction endonuclease digested *C. difficile* genomic DNA using pCD23 DNA as a probe.** Toxigenic strain 10463 DNA digested with Hind III (lane 1) and Xba I (lane 2). Nontoxigenic strain 2037 DNA digested with Hind III (lane 3), Xba I (lane 4), and Xba I/Acc I (lane 5). Sizes are indicated with arrows.

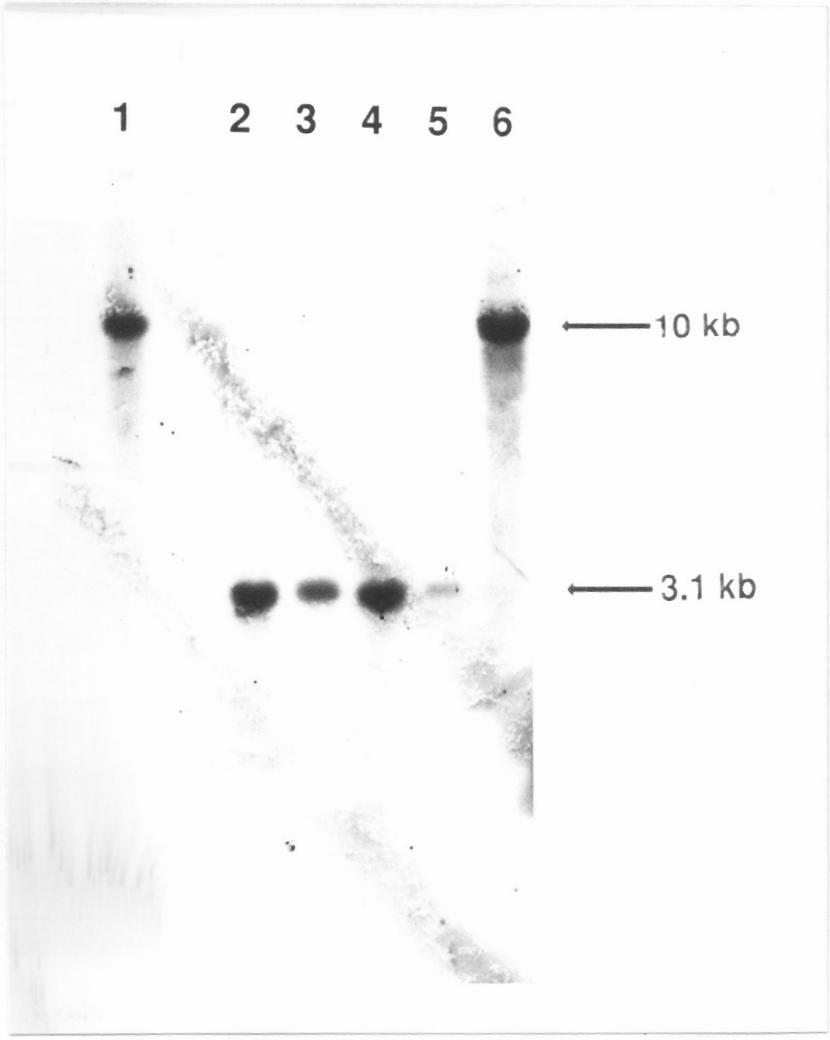


**FIG. 4. Southern blot analysis of Hind III digested *C. difficile* and *C. sordellii* genomic DNA showing shared sequence between toxigenic and nontoxigenic strains.** The probe is a Pst I-Hind III fragment from pCD11R6 (Dove et al., 1990). Toxigenic strain 10463 (lane 1); nontoxigenic strains 5199, 11186, and 2037 in lanes 2-4, respectively; *C. sordellii* strain 9048 (lane 5); toxigenic strain 26689 (lane 6). Sizes are indicated with arrows.

bp 3' to the toxin A open reading frame. With these data I established the 3' boundary for the toxigenic element.

Based on these alignment data, it appeared that pCDNT2 contained the insertion region for the toxigenic element; that is, sequences that comprised both the 5' and the 3' boundaries for the element, but the overlap between pCDNT2 and pCD21 was a very short sequence. Therefore, I generated a Polymerase Chain Reaction (PCR) product utilizing nontoxigenic strain 2037 DNA and a primer upstream of the *Hin*dIII site which is the 5' boundary of pCD21 and a primer downstream of the 3' boundary which was established for the toxigenic element (see Fig. 9). The DNA sequence of the PCR product was aligned with the sequence surrounding the *Hin*dIII site in toxigenic strain 10463 and the sequence for pCDNT2 (data not shown). The alignments of the three sequences showed that the *Hin*dIII site which comprised the 5' boundary of the nontoxigenic insert was the same *Hin*dIII site which comprised the 5' boundary of the pCD21 insert. These data confirmed the fact that the *Hin*dIII site formed the 5' boundary of the toxigenic element. In addition, these data established that pCDNT2 comprised the insertion region for the toxigenic element.

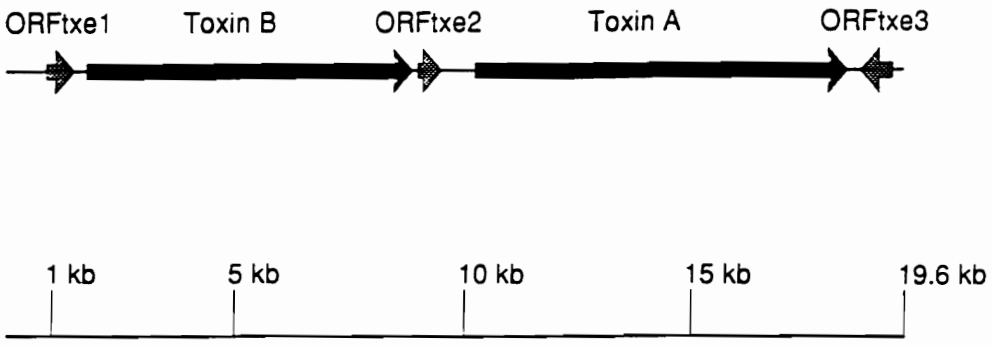


**FIG. 5. The toxigenic element of *C. difficile* strain 10463.** The open reading frames of the toxigenic element are the following: ORFtxe1, this study. Toxin B, (Barroso et al., 1990; Eichel-Streiber et al., 1990). ORFtxe2, toxin A, and ORFtxe3 from Dove et al. (1990); Sauerborn and Eichel-Streiber (1990). ORFtxe2 and ORFtxe3 are utx A and dtx A, respectively, of Sauerborn and Eichel-Streiber (1990). Arrows indicate the direction of transcription.

Characterizing the toxigenic element

With both the 5' and 3' boundaries established, it was now possible to characterize the toxigenic element from strain 10463. The toxigenic element consists of 19.6 kb of DNA and contains five open reading frames (Fig. 5). One of these open reading frames has not been previously identified. It is located approximately 300 bp upstream of the ATG start codon for toxin B, and I have designated it ORFtxe1, for the first open reading frame in the toxigenic element. ORFtxe1 is 555 bp in length and encodes a putative protein of 184 amino acid residues with a molecular mass of 22,158 Da. ORFtxe1 has a potential ribosome-binding site located 11 bp upstream of its ATG start codon. There are a number of positively charged amino acid residues present at the carboxy-terminus of ORFtxe1. Searches in databases have failed to reveal any significant similarity with other known proteins. Both the nucleotide and the deduced amino acid sequences for ORFtxe1 are presented in Fig. 6.

Examining the toxigenic element in other toxigenic strains

In order to determine whether the toxigenic element is unique to *C. difficile* strain 10463 or whether it exists in the same form in other toxigenic strains, I decided to examine both the 5' and the 3' boundaries of the toxigenic element in

```
        10          20          30          40          50          60
AAGCTTATGT ATAAATGTTA TTATTTTTAT TACAATTTAG ATAATTTTGG TTTTGTAAGC

        70          80          90         100         110         120
TCTGCAACTA TTTTTAGATG GTTGCAGAGT TTATTTTTAT TTTTTATTCC CTTAAGCAAT

       130         140         150         160         170         180
ACTTAAAAGT AAACGCATTT ATATTTACTA ACTTCAATAA CTTCTAATAC TTAAATTACT

       190         200         210         220         230         240
TTTTAAAAAT ATTTTGATAT GTCATCTCTA TAGTAATCTA GAACTTGTTT CTATTCTTTA

       250         260         270         280         290         300
TATAAAGTTT TTATGAATGA TATATATAAT ATTCGTTCAT TTGATTAAAA ATAACAAAAT

       310         320         330         340         350         360
ATTAAATAAT TCTACTCTGA TAGTTTGTTA AAAAAATAAT AAAAAATATT AATAAACAAA

       370         380         390         400         410         420
AAAAAAATTA TCTTAAGAGA CGAGAATGTT CTAAAATATA AAAAGGTTTC TAGATTTCAT

       430         440         450         460         470         480
AAAAGATACT ATTTTAGTCT TGAAAATATT TAGTTTGAAA AGATTTTAAT TTAATGATTG

       490         500         510         520         530         540
ATTAAATTAA AAATGTGTAT GTAATAAATT TTATTTTATT TATTGACTAA ATTATAAAGT

       550         560         570         580         590         600
TTACATAATT ATTTAATAAT TATGTAATTG TTACTTGAAA ATTGATCTAT TTTAAAATCT

       610         620         630         640         650         660
AGTTATAACT TCAAAAAAGA CTGAAAATTA AGAAAAAAGA AATATAAATA TAAAAATATG

       670         680         690         700         710         720
TCATATAGAT TTTTTTTATT TTACTTTAAT AAAATGATTT GTTTTTACAA TACTTTATTA

       730         740         750         760         770         780
ATATAAAGTT TATTGCTAAA ATACTTTATT TATTAGAAAA AGATTACTAA TTAATTATTA

       790         800         810         820         830         840
AAATTAATGT ATTCATAATG CATATTTTCA TATAAAATTT AATTTATTTG CCGATTATAT

       850         860         870         880         890         900
AATTATAATT ACTGATTTAA TTCCAATGTT GTCAAAATTT TCAAATAAAG CATCATTATA

       910         920           930           940           950
AATATAAGAG AGGGTGATTT T ATG CAA AAG TCT TTT TAT GAA TTA ATT GTT
                          Met Gln Lys Ser Phe Tyr Glu Leu Ile Val

         970           980           990          1000          101
TTA GCA AGA ATT AAC TCA GTA GAT GAT TTG CAA GAA ATT TTA TTT ATG
Leu Ala Arg Asn Asn Ser Val Asp Asp Leu Gln Glu Ile Leu Phe Met

0         1020          1030          1040          1050          1
TTT AAG CCA TTA GTA AAA AAA CTT AGT AGA GTT TTA CAT TAT GAA GAG
Phe Lys Pro Leu Val Lys Lys Leu Ser Arg Val Leu His Tyr Glu Glu

060          1070          1080          1090          1100
GGA GAA ACA GAT TTA ATA ATA TTT TTT ATT GAA TTA ATA AAA AAT ATT
Gly Glu Thr Asp Leu Ile Ile Phe Phe Ile Glu Leu Ile Lys Asn Ile

  1110          1120          1130          1140          1150
AAA TTA AGT AGC TTT TCA GAA AAA AGC GAT GCT ATT ATA GTC AAA TAT
Lys Leu Ser Ser Phe Ser Glu Lys Ser Asp Ala Ile Ile Val Lys Tyr

     1160          1170          1180          1190          1200
ATT CAT AAA TCA TTA CTG AAT AAG ACT TTT GAG TTG TCT AGA AGA TAT
Ile His Lys Ser Leu Leu Asn Lys Thr Phe Glu Leu Ser Arg Arg Tyr
```

```
       1210          1220          1230          1240          125
TCT AAA ATG AAG TTT AAT TTT GTA GAA TTT GAT GAA AAT ATC TTA AAT
Ser Lys Met Lys Phe Asn Phe Val Glu Phe Asp Glu Asn Ile Leu Asn

0         1260          1270          1280          1290          1
ATG AAA AAT AAT TAT CAA AGT AAG TCT GTT TTT GAG GAA GAT ATT TGT
Met Lys Asn Asn Tyr Gln Ser Lys Ser Val Phe Glu Glu Asp Ile Cys

300          1310          1320          1330          1340
TTT TTC GAA TAT ATT TTG AAA GAA TTA TCT GGT ATT CAA AGA AAA GTT
Phe Phe Glu Tyr Ile Leu Lys Glu Leu Ser Gly Ile Gln Arg Lys Val

  1350          1360          1370          1380          1390
ATT TTT TAT AAA TAT TTA AAA GGA TAT TCT GAT AGA GAA ATA TCA GTG
Ile Phe Tyr Lys Tyr Leu Lys Gly Tyr Ser Asp Arg Glu Ile Ser Val

     1400          1410          1420          1430          1440
AAA TTA AAA ATA TCT AGA CAA GCT GTT AAT AAG GCT AAA AAT AGA GCA
Lys Leu Lys Ile Ser Arg Gln Ala Val Asn Lys Ala Lys Asn Arg Ala

       1450          1460          1470          1480          1490
TTT AAA AAA ATA AAA AAA GAC TAT GAA AAT TAT TTT AAC TTG TAATTAA
Phe Lys Lys Ile Lys Lys Asp Tyr Glu Asn Tyr Phe Asn Leu

      1500        1510        1520        1530        1540        1550
TGAGCTTAAA GAGATATTTA TAATAGAAAT CAAATTTTAG AATTAACTTT ATTGTAAAAT

      1560        1570        1580        1590        1600        1610
CAATAACTTA ATCTAAGAAT ATCTTAATTT TTATATTTTA TATAGAACAA AGTTTACATA

      1620        1630        1640        1650        1660        1670
TTTATTTCAG ACAACGTCTT TATTCAATCG AAGAGCAAAT TAATCAACTG AGTGTCTTAA

      1680        1690        1700        1710        1720        1730
ATTTAAAATG TTAGGAAGTG AATGTATATG AAAACCTAAG TAGATATTAG TATATTTTAT

      1740        1750        1760        1770        1780        1790
AAATAGAAAG GAGGATATAT AAAAGAGTTT TAGCATTTAG ATGTAAAAAT ATTATAGTAA

      1800
AGGAGAAAAT TTTATG
```

**FIG. 6. Nucleotide and deduced amino acid sequence for ORFtxe1.** Nucleotides 1-1405 comprise the 5' end of pCD21. Nucleotides 1406-1806 comprise the 5' end of pCD19L (Johnson et al., 1990; Barroso et al., 1990).

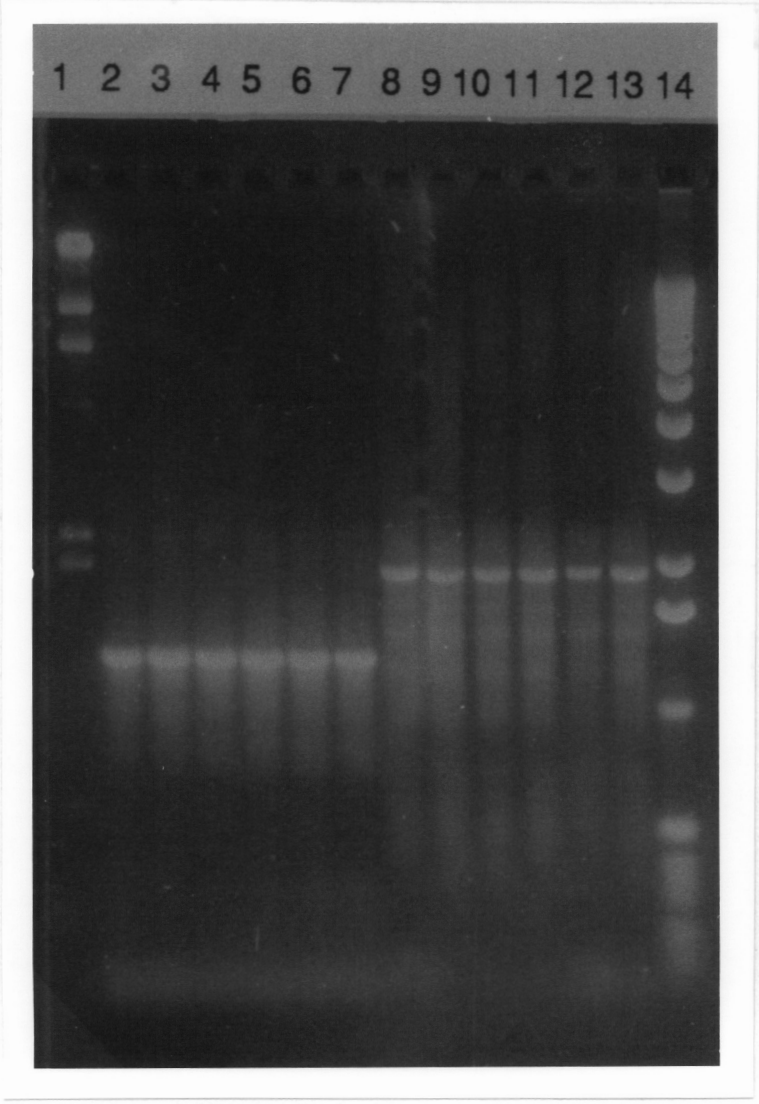


**FIG. 7. Agarose gel with products of PCR amplifications of the region from the 3' end of toxin A to the 3' boundary of the toxigenic element, and the region that includes the 5' boundary of the toxigenic element in 6 toxigenic strains of *Clostridium difficile*.** Lanes 1 and 14 are λ Hind III and 1 kb markers, respectively. PCR products from the 3' boundary region (lanes 2-7); PCR products from the 5' boundary region (lanes 8-13). Strain 10463 (lanes 2,8); strain 26689 (lanes 3,9); strain 7698 (lanes 4,10); strain ToxR (lanes 5,11); strain 11011 (lanes 6,12); strain 4474 (lanes 7,13).

other toxigenic strains. I selected 5 strains in addition to strain 10463 which varied in toxigenicity over a range of $10^6$ (Lyerly and Wilkins, 1986). Using a pair of oligonucleotide primers which allowed me to generate a PCR product corresponding to the region from the 3' end of the toxin A gene to the 3' boundary of the toxigenic element, amplification products were generated which were identical in length for each of six strains (Fig. 7). Using this same strategy, I selected a pair of oligonucleotide primers which were utilized to generate a PCR product for the region that comprises the 5' boundary of the toxigenic element and the 5' end of toxin B. Again, the amplification products which were generated were identical in length for each of the six strains (Fig. 7). Likewise, amplification products for the region that lies between the two toxin genes (from the 3' end of the toxin B gene to the 5' end of the toxin A gene) were identical in length for each of these strains (Fig. 8). These data, coupled with results from Southern analyses of restriction endonuclease fragment length polymorphisms using probes from the toxigenic element (Chapter 3), indicate that the 19.6 kb toxigenic element is conserved among these six toxigenic strains.

Characterizing the insertion region

Alignment of toxigenic strain 10463 and nontoxigenic strain 2037 DNA sequences indicated that there was additional sequence located within the insertion

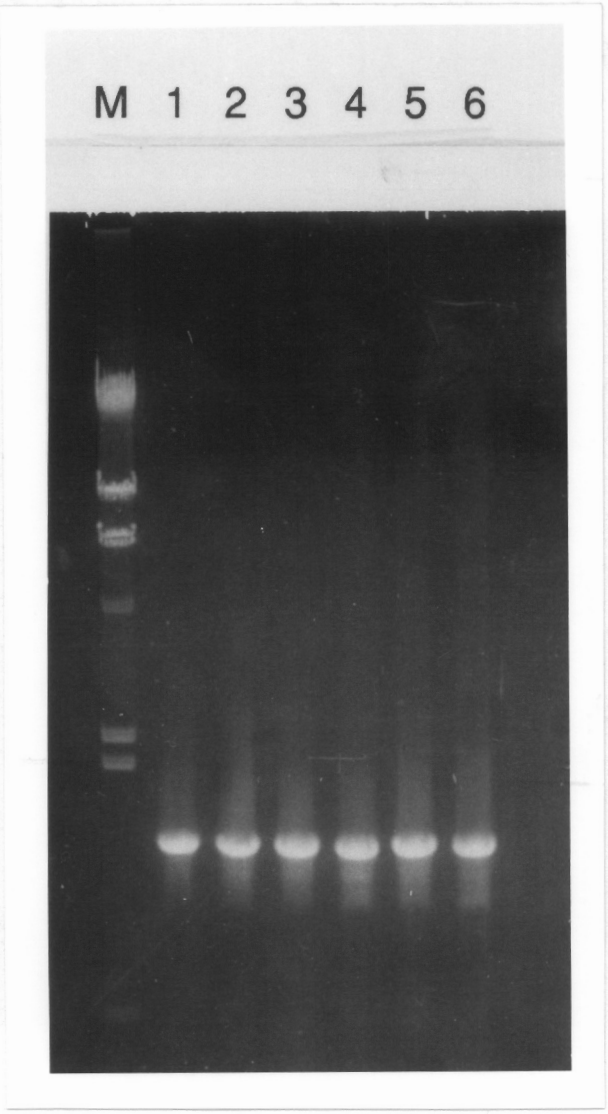


**FIG. 8. Agarose gel with products of PCR amplification of the region from the 3' end of the toxin B gene to the 5' end of the toxin A gene in 6 toxigenic strains of *Clostridium difficile*.** Marker is λ Hind III (M). Strain 10463 (lane 1); strain 26689 (lane 2); strain 11011 (lane 3); strain 7698 (lane 4); strain 4474 (lane 5); strain ToxR (lane 6).

region in the nontoxigenic strain which was not found in the toxigenic strain. In order to demonstrate this I prepared oligonucleotides from the region upstream of the 5' toxigenic element boundary and downstream of the 3' toxigenic element boundary. In the toxigenic strain 10463 these two primers are separated by a distance of 19.6 kb. I predicted that only a very short distance separated these two primers in the nontoxigenic strain. I utilized this primer pair along with genomic DNA from toxigenic strain 10463 and three nontoxigenic strains in a PCR amplification of the region. Due to the reaction parameters and the length of our anticipated product, a PCR product using toxigenic strain DNA was not generated; however, a short fragment was generated from each of the nontoxigenic strains. The PCR products were identical in length for each of the nontoxigenic strains (Fig. 9).

Using the symmetric PCR product from nontoxigenic strain 2037 as the template, asymmetric PCR products were generated for sequencing. This DNA sequence was then aligned with the upstream and the downstream sequences from the toxigenic strain 10463. The results demonstrate the presence of a unique 127 bp fragment in nontoxigenic strains. The 19.6 kb which comprise the toxigenic element occupy this same position in toxigenic strain 10463 (Fig. 10).

The short fragment unique to the nontoxigenic strains has some characteristics of an insertion element. Six base pair AT-rich repeats are found at the ends of the fragment. Internally, there are two inverted and one direct repeat regions (Fig. 11).

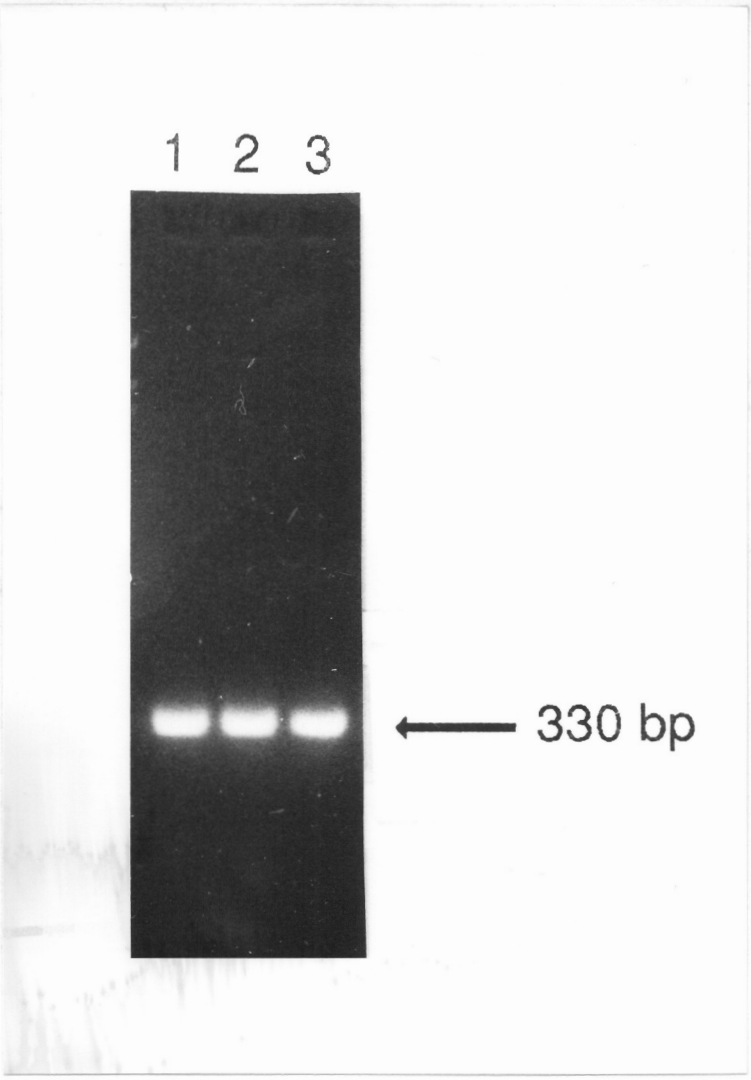


**FIG. 9. Agarose gel with products of a PCR amplification of the region in nontoxigenic strains which occupies the same position as the toxigenic element.** Strain 2037 (lane 1); strain 5199 (lane 2); strain 11186 (lane 3). Size is indicated by arrow. Oligonucleotide primers were HdHcAS and NT2A (see Materials and Methods).

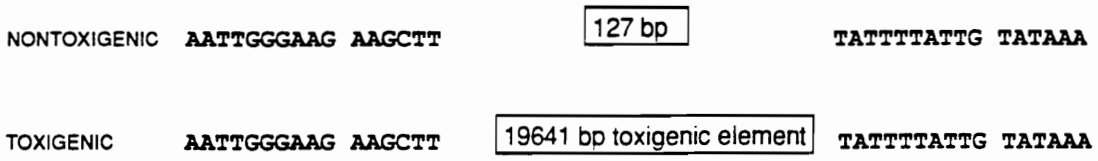


**FIG. 10. The 127 bp sequence unique to nontoxigenic strains which occupies the same chromosomal location as the 19,641 bp toxigenic element.** The upstream flanking sequence contains the upstream Hind III site of pCD21. See Fig. 11 for DNA sequence of the 127 bp fragment.

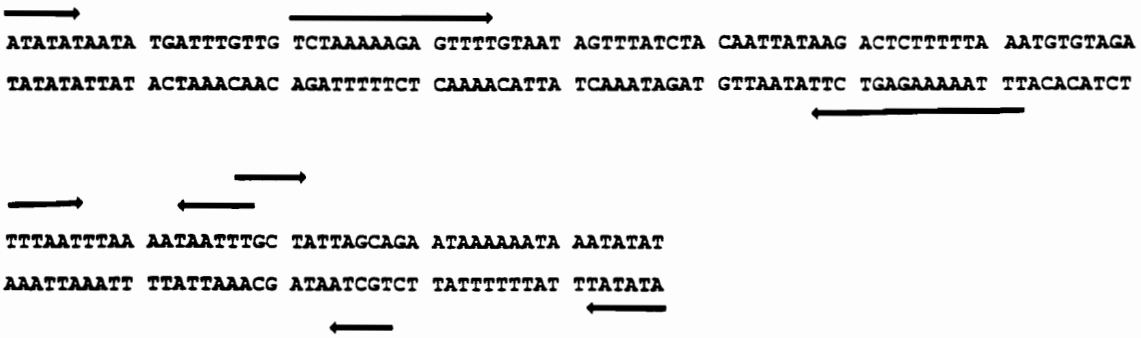


**FIG. 11. Nucleotide sequence of the nontoxigenic 127 bp fragment which occupies the same chromosomal location as the toxigenic element.** Repeat regions are indicated by arrows above the sequence.

DISCUSSION

By using DNA fragments that flank the toxin genes as probes, I have defined the toxigenic element in *Clostridium difficile* strain 10463. I have established upstream (5') and downstream (3') boundaries between DNA sequences that are unique to the toxigenic strain and those that are held in common among both toxigenic and nontoxigenic strains.

The toxigenic element is 19.6 kb in length and contains five open reading frames. Four of these open reading frames, including toxin A and B have been identified previously (Barroso et al., 1990; Dove et al., 1990; Eichel-Streiber and Sauerborn, 1990; Eichel-Streiber et al., 1992; Eichel-Streiber, 1993). The open reading frames which I have designated ORFtxe2 and ORFtxe3 are designated as utxA and dtxA by Eichel-Streiber. The fifth open reading frame is identified herein. It is a small open reading frame, 555 nucleotides in length, located upstream of the toxin B gene which I have designated ORFtxe1.

Now that the toxigenic element is defined in strain 10463, it will be possible to further examine the role of these sequences adjacent to the toxin genes. In this regard, each of the toxin genes has been cloned separately. The recombinant proteins expressed by these clones demonstrate the same properties as the native proteins (Johnson et al., 1990; Phelps et al., 1991). Therefore, expression of the three remaining open reading frames in the toxigenic element most likely plays no

role in the activity of either toxin A or B; however, the proteins expressed from these open reading frames may have some additional function in virulence of *Clostridium difficile.* I will present data from Northern analyses which indicate that two of the three open reading frames, ORFtxe1 and ORFtxe2, are expressed (Chapter 2).

The fact that regions held in common among both toxigenic and nontoxigenic strains were detected, supports the chromosomal location of the toxin genes. Johnson et al. (1988) had previously concluded that the toxin A gene was chromosomal when they hybridized toxin A probe with Southern transfers of high molecular weight DNA which had been subjected to pulse field electrphoresis. The hybridization pattern appeared the same as that of the ethidium bromide-stained gel.

Southern blot analyses of six strains which vary considerably in their toxigenicity (Chapter 3), coupled with PCR amplification analyses of both the upstream and the downstream boundaries, as well as the region in between the toxin genes, revealed that the toxigenic element is conserved among these strains. Similarly, the insertion region, the small DNA fragment which occupies the same position as the toxigenic element, was conserved in the three nontoxigenic strains which were examined. These data support that the toxigenic element is located at a specific position within the chromosome.

A 127 bp fragment unique to the three nontoxigenic strains which we examined occupies the same position in the chromosome as the 19.6 kb toxigenic

element. The DNA sequences at the junctions of the toxigenic element were examined in order to determine whether they showed evidence of insertional type sequences such as inverted repeat regions at the termini. The junctions lack terminal inverted repeat regions; however, the insertion region from the nontoxigenic strains does contain terminal inverted repeat regions. In addition, there are several internal repeat regions that may serve a recombination function.

## LITERATURE CITED


1. Barroso, L. A., S.-Z. Wang, C. J. Phelps, J. L. Johnson, and T. D. Wilkins. 1990. Nucleotide sequence of the *Clostridium difficile* toxin B gene. Nucleic Acids Res. 18: 4004.

2. Birnboim, H. C. and J. Doly. 1979. A rapid alkaline extraction procedure for recombinant plasmid DNA. Nucleic Acids Res. 7: 1513-1523.

3. Borriello, S. P., B. W. Wren, S. Hyde, S. V. Seddon, P. Sibbons, M. M. Krishna, S. Tabaqchali, S. Manek, and A. B. Price. 1992. Molecular, immunological, and biological characterization of a toxin A negative, toxin B positive strain of *Clostridium difficile.* Infect. and Immun. 60: 4192-4199.

4. Dove, C. H., S.-Z. Wang, S. B. Price, C. J. Phelps, D. M. Lyerly, T. D. Wilkins, and J. L. Johnson. 1990. Molecular characterization of the *Clostridium difficile* toxin A gene. Infect. and Immun. 58: 480-488.

5. Dower, W. J. 1988. High efficiency transformation of *E. coli* by high voltage electroporation. Nucleic Acids Res. 16: 6127-6145.

6. Eichel-Streiber, C. v. 1993. Molecular biology of the *Clostridium difficile* toxins. pp. 264-289. *In* M. Sebald (ed.), Genetics and molecular biology of anaerobic bacteria. Springer-Verlag, New York.

7. Eichel-Streiber, C. v., R. Laufenberg-Feldmann, S. Sartingen, J. Schulze, and M. Sauerborn. 1990. Cloning of *Clostridium difficile* toxin B gene and demonstration of high N-terminal homology between toxin A and B. Med. Microbiol. Immuno. 179: 271-279.

8. Eichel-Streiber, C. v., R. Laufenberg-Feldmann, S. Sartingen, J. Schulze, and M. Sauerborn. 1992. Comparative sequence analysis of the *Clostridium difficile* toxins A and B. Mol. Gen. Genet. 233: 260-268.

9. Eichel-Streiber, C. v. and M. Sauerborn. 1990. *Clostridium difficile* toxin A carries a C-terminal repetitive structure homologous to the carbohydrate binding region of streptococcal glycosyltransferases. Gene 96: 107-113.

10. Eichel-Streiber, C. v., D. Suckau, M. Wachter, and U. Hadding. 1989. Cloning and characterization of overlapping DNA fragments of the toxin A gene of *Clostridium difficile.* J. Gen. Microbiol. 135: 55-64.

11. Fluit, A. C., M. J. H. M. Wolfhagen, G. P. H. T. Verdonk, M. Jansze, R. Torensma, and J. Verhoef. 1991. Nontoxigenic strains of *Clostridium difficile* lack the genes for both toxin A and toxin B. J. of Clin. Microbiol. 29: 2666-2667.

12. Holdeman, L. V., E. P. Cato, and W. E. C. Moore, (eds.). 1977. Anaerobe Laboratory Manual. V. P. I. Anaerobe Laboratory, Virginia Polytechnic Institute and State University, Blacksburg, VA.

13. Johnson, J. L. 1991. DNA reassociation experiments. pp.21-41. *In* E. Stackebrandt and M. Goodfellow (eds.), Nucleic acid techniques in bacterial systematics. John Wiley & Sons Ltd., West Sussex, England.

14. Johnson, J. L., C. H. Dove, S. B. Price, T. W. Sickles, C. J. Phelps, and T. D. Wilkins. 1988. The toxin A gene of *Clostridium difficile.* pp. 115-123. *In* J. M. Hardie and S. P. Borriello (eds.), Anaerobes today. John Wiley & Sons Ltd., West Sussex, England.

15. Johnson, J. L., C. Phelps, L. Barroso, M. D. Roberts, D. M. Lyerly, and T. D. Wilkins. 1990. Cloning and expression of the toxin B gene of *Clostridium difficile.* Curr. Microbiol. 20: 397-401.

16. Lyerly, D. M., L. A. Barroso, T. D. Wilkins, C. Depitre, and G. Corthier. 1992. Characterization of a toxin A-negative, toxin B-positive strain of *Clostridium difficile*. Infect. and Immun. 60:4633-4639.

17. Lyerly, D. M., K. E. Saum, D. K. MacDonald, and T. D. Wilkins. 1985. Effects of *Clostridium difficile* given intragastrically to animals. Infect. and Immun. 47: 349-352.

18. Lyerly, D. M. and T. D. Wilkins. 1986. Commercial latex test for *Clostridium difficile* toxin A does not detect toxin A. J. of Clin. Microbiol. 23: 622-623.

19. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

20. Marmur, J. 1961. A procedure for the isolation of deoxyribonucleic acid from microorganisms. J. Mol. Biol. 3: 209-218.

21. McMillin, D. E., L. L. Muldrow, S. J. Leggette, Y. Abdulahi, and U. M. Ekanemesang. 1991. Molecular screening of *Clostridium difficile* toxins A and B genetic determinants and identification of mutant strains. FEMS Microbiol. Lett. 78: 75-80.

22. Phelps, C. J., D. L. Lyerly, J. L. Johnson, and T. D. Wilkins. 1991. Construction and expression of the complete *Clostridium difficile* toxin A gene in *Escherichia coli*. Infect. and Immun. 59: 150-153.

23. Price, S. B., C. J. Phelps, T. D. Wilkins, and J. L. Johnson. 1987. Cloning of the carbohydrate-binding portion of the toxin A gene of *Clostridium difficile*. Curr. Microbiol. 16: 55-60.

24. Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H.A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239: 487-491.

25. Sambrook, J. Fritsch, E. F., and T. Maniatis. 1989. Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, New York.

26. Sanger, F. S. Nickelsen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74: 5463-5467.

27. Sauerborn, M. and C. v. Eichel-Streiber. 1990. Nucleotide sequence of *Clostridium difficile* toxin A. Nucleic Acids Res. 18: 1629-1630.

28. Southern E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98: 503-517.

29. Struhl, K. 1985. A rapid method for creating recombinant DNA molecules. Biotechniques 3: 452-453.

30. Tartof, K. D. and C. A. Hobbs. 1987. Improved media for growing plasmid and cosmid clones. Focus 9: No. 2, p. 12. Bethesda Research Laboratories, Life Technologies, Inc.

31. Torres, J. F. 1991. Purification and characterization of toxin B from a strain of *Clostridium difficile* that does not produce toxin A. J. Med. Microbiol. 35: 40-44.

32. Wilkins, T. D. and D. M. Lyerly. 1990. Toxins of *Clostridium difficile* and *Clostridium sordellii*. pp. 45-53. *In* J.-C. Rambaud and R. Ducluzeau (eds.), *Clostridium difficile*-associated intestinal diseases. Springer-Verlag, France.

33. Wren, B., S. R. Heard, and S. Tabaqchali. 1987. Association between production of toxins A and B and types of *Clostridium difficile*. J. Clin. Pathol. 40: 1397-1401.

34. Yanisch-Perron, C. J. Vieira, and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33: 103-119.

CHAPTER TWO: Transcriptional Analysis of the Toxigenic Element

## INTRODUCTION

A toxigenic element from *Clostridium difficile* strain 10463 has been defined (Chapter 1). This toxigenic element is conserved among the six toxigenic strains which were examined. The toxigenic element, illustrated in Fig. 1, is 19.6 kb in size, and comprised of five open reading frames which include the toxin A and toxin B genes. Four of the open reading frames which are contiguous are transcribed in the same direction. The fifth, ORFtxe3, is downstream from the others and oriented in the opposite direction. The genes for toxin A (Price et al., 1987; Johnson et al., 1988; Eichel-Streiber et al., 1989; Dove et al., 1990; Sauerborn and Eichel-Streiber, 1990) and toxin B (Barroso et al., 1990; Eichel-Streiber et al., 1990; Johnson et al., 1990) have been cloned and sequenced and lie within 1.4 kb of each other on the chromosome. Based on DNA sequence inspection, putative promoters and transcription terminators have been identified for the toxin A gene (Eichel-Streiber and Sauerborn, 1990; Eichel-Streiber, 1993) and the toxin B gene (Eichel-Streiber et al., 1992; Eichel-Streiber, 1993).

With rare exceptions (McMillan et al., 1991; Torres, 1991; Borriello et al., 1992; Lyerly et al., 1992), toxigenic strains of *C. difficile* produce both toxins.

Toxigenic strains that produce high levels of the toxin A protein, also produce high levels of toxin B protein. Conversely, toxigenic strains that produce relatively low levels of toxin A also produce relatively low levels of toxin B (Lyerly and Wilkins, 1986; Wren et al., 1987). For these reasons, it has been generally accepted that the two toxin genes are co-regulated.

The goals of the present study were to analyze transcription of the toxigenic element and the open reading frames contained therein, to determine the size of the transcripts, to identify transcriptional promoter regions, and to examine the role of bacterial growth state on expression of the toxin genes.

## MATERIALS AND METHODS

### Bacterial strains, bacteriophages, and growth conditions

*Clostridium difficile* strain 10463 was obtained from the laboratory collection of Dr. Tracy Wilkins, Department of Biochemistry and Anaerobic Microbiology, Virginia Polytechnic Institute and State University, Blacksburg, VA. Cultures for the isolation of RNA were grown anaerobically at 37°C in BHI, Brain-heart infusion (Difco) dialysis flasks as described in Lyerly and Wilkins (1991). Incubation times ranged from 7 - 22 hours.

Cultures for the isolation of DNA were grown anaerobically at 37°C in PYG

medium (Anaerobe Laboratory Manual) supplemented with 1% dehydrated BHI.

The bacteriophage M13mp19 was utilized as the vector for cloning DNA for the purposes of sequencing. *E. coli* DH5αF' obtained from BRL served as the host with the M13mp19 vector and was grown aerobically at 37°C in 2XYT medium (Sambrook et al., 1989).

DNA isolation

Genomic DNA of *C. difficile* was isolated by the method of Marmur (1961) as modified by Johnson (1991).

RNA isolation and manipulations

Total RNA was isolated according to the guanidinium thiocyanate method of Puissant and Houdebine (1990) with the following exceptions. Once the initial cell pellet was resuspended in sucrose buffer, cells were lysed with lysozyme. Following phenol-chloroform extraction and isopropanol precipitation, the pellet was resuspended in Tris EDTA (TE) buffer, pH 7.5. A sufficient volume of 6 M lithium chloride (LiCl) was added to make the final solution 4 M LiCl for the LiCl extraction step. The final RNA pellet was dissolved in a small volume of diethyl pyrocarbonate (DEP)-treated water to which was added the equivalent of 4 volumes

of 95% ethanol to make an RNA-ethanol solution (Sambrook et al., 1989) which was stored at -20°C. RNA concentration was measured at 260 nm on a Gilford 2400 spectrophotometer.

To obtain RNA samples for gel electrophoresis, either 10 or 25 μg quantities of RNA were removed from the RNA-ethanol solution and precipitated with 0.1 volume of 3M sodium acetate. RNA sample preparation followed the method of Kroczek and Siebert (1990). RNA samples were electrophoresed on 12.7 X 15.9 cm 1.1% formaldehyde, 1X MOPS (3-[N-Morpholino]propanesulfonic acid), 1% agarose (SeaKem LE, FMC Bioproducts) gels at 70 Volts for 4-5 hours on a horizontal gel electrophoresis apparatus (Jordan Scientific) which recirculated the 1X MOPS electrophoresis buffer. Chemicals for the isolation of RNA and for RNA electrophoresis were obtained from Amresco, Gibco BRL, Fisher Scientific, or Sigma.

When electrophoresis was complete, the RNA gel was placed on three layers of Whatman 3MM paper and dried under vacuum in a Bio-Rad slab gel dryer, Model 1125B, initially, for 30 min. at room temperature, and then for 2 hours at 60°C (Ahmad et al., 1990).

DNA fragments were labelled as hybridization probes using $\alpha^{32}$P-dATP, 3000 Ci/mMole from DuPont NEN Research Products and either the Prime-It or the Prime-It II random primer labelling kit from Stratagene.

To prepare the dried gel for hybridization, it was briefly hydrated in a glass

baking dish filled with DEP-treated water in order to remove the Whatman paper backing. The gel was then placed in hybridization buffer which consisted of the following components in their final concentrations: 25% deionized formamide, 0.5% sodium dodecyl sulfate (SDS), 5X Denhardt's Reagent (Sambrook et al., 1989), 6X Standard Saline-Phosphate-EDTA (SSPE) (Sambrook et al., 1989). Chemicals were obtained from either Fisher Scientific or Sigma. Heat-denatured, radioactively labelled probe was added (1 X $10^6$ cpm/ml) and hybridizations were carried out at 50°C overnight. Hybridizations were performed in a Techne Hybridiser HB-1 oven using FHB 11 bottles with 8 ml hybridization buffer.

All of the posthybridization buffers were heated to 50°C. Posthybridization washes took place at 37°C with slow agitation, and consisted of the following solutions: 2X Standard Saline Citrate (SSC), 0.1% SDS for 10 min; 0.5X SSC, 0.1% SDS for 20 min.; 0.1X SSC, 0.1% SDS, 3 times at 40 min. each; and 0.1X SSC for 30 min. Gels were covered with Saran Wrap and exposed to Kodak SB5 film at -80°C. For use as markers, the 1 kb ladder (BRL) was treated in the same manner as the RNA samples and electrophoresed along with them (Chaudhari, 1991). A separate aliquot of the ladder DNA was radioactivelly labelled and added to the hybridization mixture along with the RNA-specific probe. RNA sizes were determined using a linear regression analysis program.

RNA secondary structure was analyzed with the Genetics Computer Group program from the University of Wisconsin.

PCR amplification

DNA fragments which were used as hybridization probes were Polymerase Chain Reaction (PCR) products. PCR amplifications were performed on a GTC-1 Genetic Thermal Cycler from Precision Scientific using Taq DNA Polymerase from Perkin Elmer and in general followed the method of Saiki et al. (1988). Oligonucleotide primers were synthesized on a 381A DNA Synthesizer from Applied Biosystems. The symmetric amplifications were run for a total of 40 cycles. The 50 µl reaction contained 100 ng of genomic DNA along with 1.25 Units of Taq DNA Polymerase. The final concentration of oligonucleotide primer was 0.5µM. Initially, the template was denatured at 95°C for 1.5 min., followed by annealing at 45°C for 1 min. For the remaining 39 cycles the denaturing temperature was lowered to 94°C for 30 sec., followed by annealing for 30 sec. at 45°C. The elongation step for all 40 cycles was at 70°C, but ranged in duration from 1-3 min. based on the length of the PCR product.

Primer extension analysis

Primer extension reactions in general followed the method of Alam et al. (1986). Three oligonucleotide primers were used. The primer for the full length transcript, 5' GCAAATCATCTACTGAGTTATTTC 3' is located 61 bp downstream

of the translational start for ORFtxe1. The extension primer for toxin B, 5' CATCTTCTTGAGTACGAAATC 3', is found 67 bp 3' to the ATG start codon for toxin B. 5' GGTCTAATGCTATATGCGAG 3', the extension primer for A, is situated 53 bp downstream of the translational start site for toxin A.

Total RNA (in either 10 or 25 μg quantities) was pelleted and the pellets were resuspended in 1 μl primer (2.0 pmol) and 5.5 μl of a solution which consisted of 7.5 μM each of dGTP, dCTP, and dTTP (Boehringer Mannheim). The RNA was heat-denatured by boiling for 2 min.; 2 μl of 5X AMV Reverse Transcriptase buffer (Promega) were added and the mixture was incubated at 37°C for 5 min. to allow annealing with the primer. Once annealing was complete, 1 μl $\alpha^{32}$P-dATP, 500 Ci/mMole from DuPont NEN Research Products and 0.5 μl Promega AMV Reverse Transcriptase, 23 U/μl, were added. The reaction was incubated at 37°C for 30 min. As a chase solution 3 μl of 2.5 mM dNTPs were added and the incubation was continued for an additional 30 min. at 37°C. The reaction stop solution consisted of the following: 95% formamide, 20 mM EDTA, 0.05% Bromophenol Blue, and 0.05% Xylene Cyanol FF. Chemicals were obtained from either Fisher Scientific or Sigma.

DNA fragments were cloned into M13mp19 following the procedures outlined by Amersham in the M13 Cloning and Sequencing Handbook. Single-stranded template was prepared according to the Amersham protocol. DNA sequencing followed the method for Sequencing with Sequenase Version 2.0 from

USB which is a modification of the Sanger techniques (Sanger et al., 1977) with the following exceptions. Single-stranded DNA (2 μg) was mixed with 0.5 pmol of the 5' GCAAATCATCTACTGAGTTATTTC 3' primer and boiled for 2 min. 5X Sequenase buffer (USB) was added and the primer was allowed to anneal to the template at 37°C for 5 min. $\alpha^{32}$P-dATP, 500 Ci/mMole from DuPont NEN Research Products and Sequenase Version 1.0 (USB) were used. The sequencing termination mixes were modified to increase the amount of dATP and dTTP, due to the low %G + C of *C. difficile*.

Both the sequencing reactions and the primer extension products were analyzed on a 6% polyacrylamide sequencing gel run on an IBI Model STS45 Sequencing Gel Electrophoresis Apparatus at a constant power of 65 Watts. Subsequently, the gel was soaked in 7.5% acetic acid, 7.5% methanol for 30 min. and dried on a Bio-Rad Slab Gel Dryer Model 1125B at 80°C. The gel was exposed to Kodak XAR-5 or Kodak SB5 film at room temperature.

## RESULTS

### Transcription of the toxigenic element

Using sequence-specific primers, PCR amplification products were generated

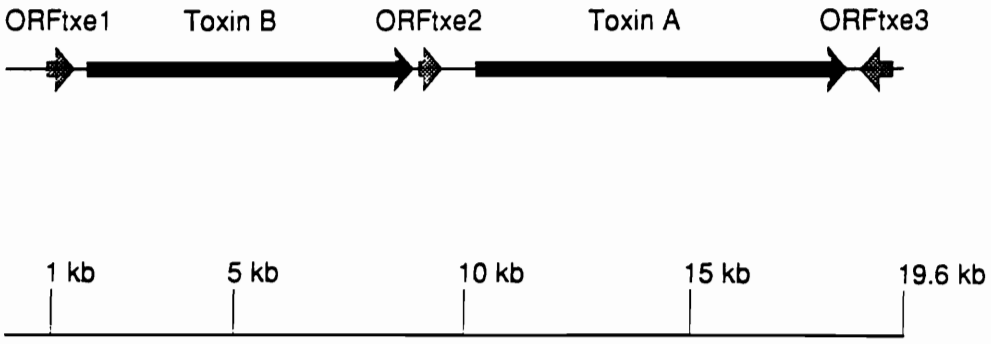


**FIG. 1. The toxigenic element of *C. difficile* strain 10463.** The open reading frames of the toxigenic element are the following: ORFtxe1, this study. Toxin B, (Barroso et al., 1990; Eichel-Streiber et al., 1990). ORFtxe2, (Dove et al., 1990); utxA of Sauerborn and Eichel-Streiber (1990). Toxin A, (Dove et al., 1990; Sauerborn and Eichel-Streiber, 1990). ORFtxe3, (Dove et al., 1990), dtxA of Sauerborn and Eichel-Streiber (1990). Arrows indicate the direction of transcription.

for each of the five open reading frames of the toxigenic element (Fig. 1), labelled, and used as hybridization probes. As a positive control in the hybridization experiments, a PCR product was generated and used as a probe for the transcript of the glutamate dehydrogenase gene of *C. difficile*. The probes for toxin A, toxin B, ORFtxe3, and glutamate dehydrogenase were all approximately 1 kb in size. Probes for ORFtxe2 and ORFtxe1 were smaller due to the size of the open reading frame for each.

The results of an in-gel hybridization of total RNA from 5 toxigenic and 1 nontoxigenic strains with the toxin A-specific probe are seen in Fig. 2A. The toxin A probe hybridized to two bands, estimated to be 15.5 kb and 8.4 kb in size. The length of the 8.4 kb transcript corresponds well with the length of the open reading frame for toxin A (8.1 kb). The results of an in-gel hybridization with a probe specific for toxin B are shown in Fig. 2B. There are three bands present, estimated to be 15.5 kb, 8.1 kb, and 7.4 kb. The length of the 7.4 kb transcript compares closely with the 7.1 kb open reading frame of toxin B. Additional bands and smears are the result of hybridization of the probes with both specific and nonspecific degradation products of toxins A and B.

The results of hybridizations with probes to two of the three remaining open reading frames: ORFtxe1, located 5' to the toxin B gene; and ORFtxe2, situated in between the genes for toxins B and A, appear in Fig. 3. Three bands were detected

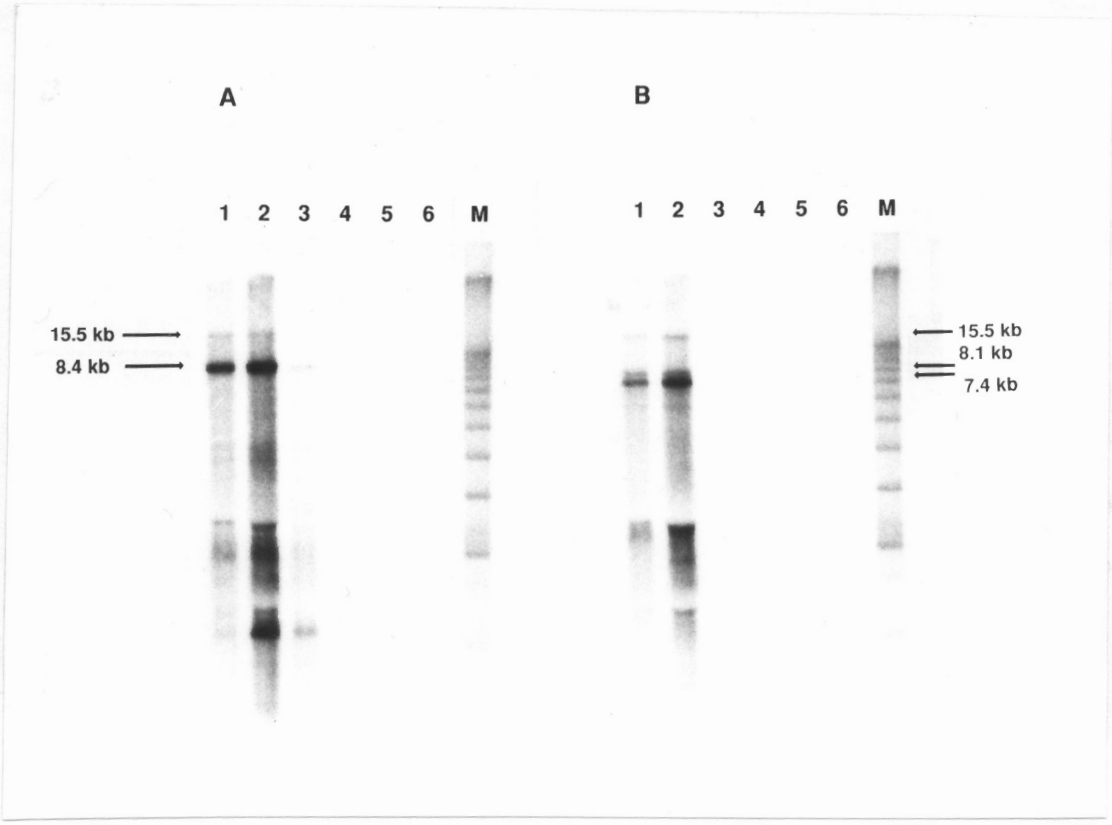


**FIG. 2. In-gel hybridization analysis of total RNA using toxin A and toxin B probes.** A. Toxin A probe hybridized to 10 μg total RNA from 12 hr. dialysis cultures of strain 10463 (lane 1), strain 26689 (lane 2), strain 7698 (Lane 3), strain 11011 (lane 4), strain 4474 (lane 5), strain 2037 (lane 6). **B.** Toxin B probe hybridized to same samples. Marker is 1 kb ladder (M). Transcripts are indicated by arrows.

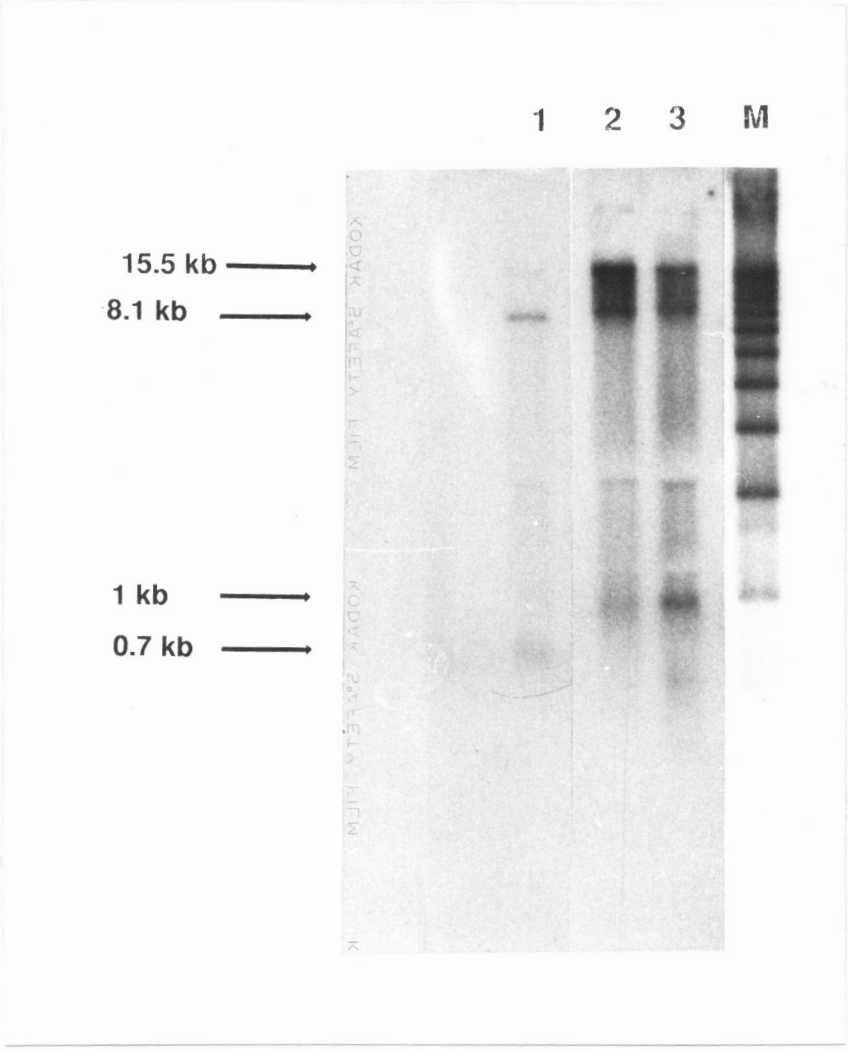


**FIG. 3. In-gel hybridization analysis of total RNA using ORFtxe1 and ORFtxe2 probes.** Twenty five μg total RNA from 12 hr. dialysis cultures. ORFtxe1 hybridized to strain 10463 (lane 1); ORFtxe2 hybridized to strain 10463 (lane 2), strain 26689 (lane 3). Marker is 1 kb ladder (M). Transcripts are indicated by arrows.

with both of the probes. In each case both a 15.5 kb and an 8.1 kb transcript were detected. Additionally, a 700 bp transcript was identified with the ORFtxe1 probe and a 1 kb transcript hybridized with the ORFtxe2 probe. I failed to detect a trancript with the probe for ORFtxe3, located 3' to the gene for toxin A and transcribed in the opposite direction to that of all of the other open reading frames (data not shown).

Identifying the promoter region and the major processing sites of the toxigenic element

In order to map the transcription initiation site for the toxigenic element, a primer extension analysis was undertaken utilizing material from the same RNA preparations which were used in the in-gel hybridizations. An oligonucleotide primer which had its 5' terminus located 61 bp downstream of the translational start site for ORFtxe1 was designed to be used to determine transcription initiation for the large transcript. Similarly, oligonucleotide primers were designed which were located 53 bp and 67 bp downstream of the ATG start codon for the toxin A and toxin B open reading frames, respectively, in order to identify the major processing sites for these transcripts.

The results of all three primer extension reactions are shown in Fig. 4. A transcription initiation site for the large transcript (15.5 kb) was identified 236 bp

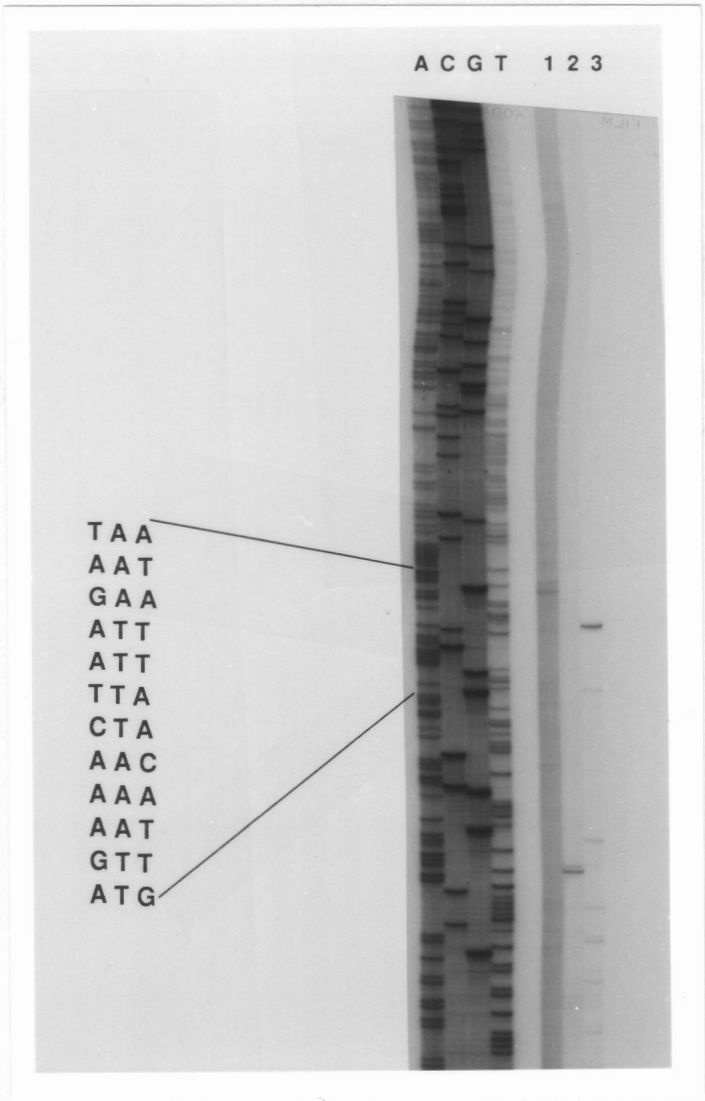


**FIG. 4. Primer extension analysis for toxigenic element and primer extension products for toxins A and B.** Primer extension product for toxigenic element (lane 1); primer extension product for toxin B (lane 2); primer extension product for toxin A (lane 3). DNA sequence is from the toxigenic element primer. The DNA sequence of the region which corresponds to the primer extension product for the toxigenic element is indicated. See Materials and Methods for primer sequences.

Table 1. *Clostridium difficile* Promoters

| Source | Gene | -35 | Nucleotide Sequence | -10 |
|---|---|---|---|---|
| Wren *et al.* (1989) | cat D | TTGAAA | AATTCACAAAAATGTGG | TATAAT |
| This study | toxigenic element promoter | TATGTC | ATATAGATTTTTTT | TATTTT |
| Daley and Schloemer (1988) | promoter element | TTCACA | TCCTCCACCTAAAGCAAATCCG | TTTACA |
| Eichel-Streiber and Sauerborn (1990) | tox A | TTAACA | AATTACTATCAGACAATCTCC | TTATCT |
| Eichel-Streiber and Sauerborn (1990) | utx A (referred to in this study as ORFtxe2) | TGCACA | GTAGTTCACCTTTTTATATT | TCTAAT |
| Eichel-Streiber (1993) | tox B | TTTACA | TATTTATTTCAGACAACGTCT | TTATTC |

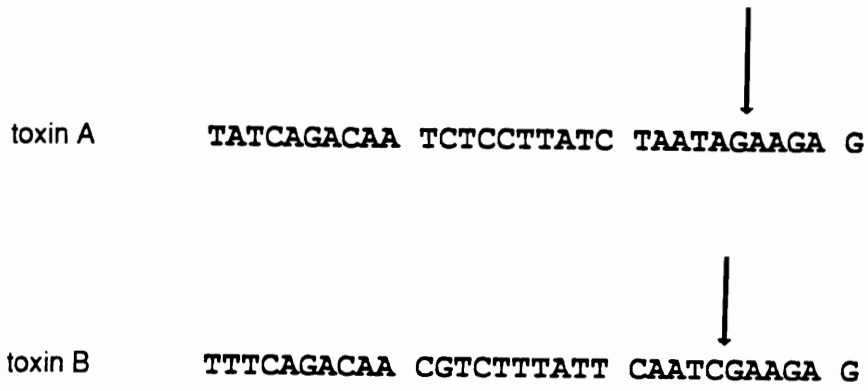


**FIG. 5. Nucleotide sequences for major processing sites for toxin A and toxin B transcripts as identified by primer extension analysis.** Arrows indicate major processing sites.

upstream of the translational start codon for ORFtxe1. A promoter region has been specified and it is presented along with other published *Clostridium difficile* promoters, including the promoters which have been designated for toxin B, ORFtxe2, and toxin A, based on sequence inspection, in Table 1. Major processing sites were identified 233 bp upstream of toxin A and 164 bp upstream of toxin B.

The DNA sequences comprising the major processing sites for both the toxin A and toxin B transcripts, as determined by primer extension analysis, show extensive sequence similarity (Fig. 5). RNAse processing occurs at the same site within each of these regions. An analysis of the 172 bp sequence which is located between toxin B and ORFtxe2; that is, 3' to the toxin B gene and 5' to ORFtxe2, (a region where processing also occurs to yield the ORFtxe2 and toxin B transcripts) reveals a sequence quite similar to the processing sites demonstrated upstream of toxins A and B (TCTATCTAATATGAAGA). A computer analysis of potential RNA structure downstream of each of the first three of the four open reading frames of the toxigenic element was undertaken. These correspond to the regions in which processing occurs. Extensive RNA secondary structure is predicted in all three regions downstream of ORFtxe1, toxin B, and ORFtxe2 (data not shown).

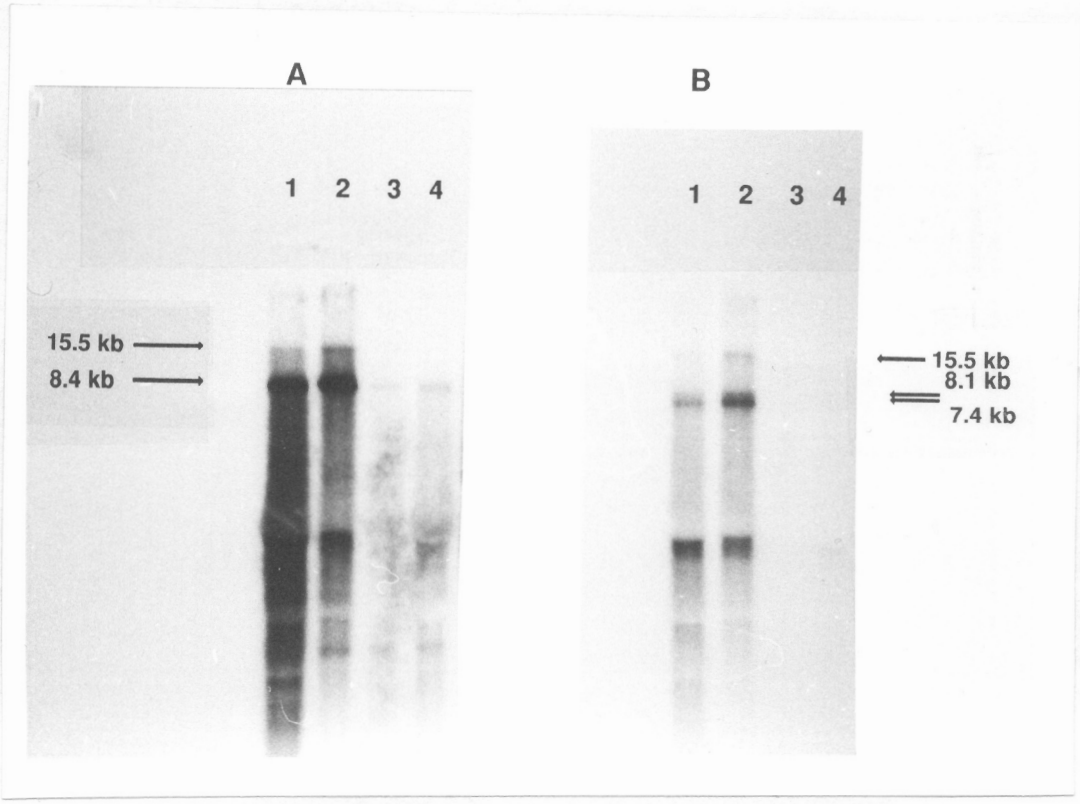


**FIG. 6. In-gel hybridization of total RNA isolated at different stages of growth with probes to toxin A and toxin B.** Twenty five μg total RNA from *C. difficile* strain 10463 isolated from dialysis or batch flask cultures. **A.** Toxin A probe hybridized to RNA from 18 hr. dialysis culture (lane 1); 12 hr. dialysis culture (lane 2); 7 hr. dialysis culture (lane 3); 9 hr. batch flask culture (lane 4). **B.** Toxin B probe hybridized to same samples. Transcripts are indicated by arrows.

Transcriptional regulation by bacterial growth state

In-gel hybridization analyses using toxin-specific probes with RNA isolated from dialysis cultures grown for 7 h, 12 h, and 18 h were performed to examine whether transcription was affected by bacterial growth state. The results of hybridizations of these RNA samples with probes for toxins A and B are presented in Fig. 6. Toxin-specific transcription appears to be minimal in the 7-h sample. Cells from the 7-h dialysis culture were small, quite uniform in size, and rapidly dividing. In the 12-h culture the bacterial cells showed diverse sizes and cell divisions were not as numerous, indicating that the culture was entering stationary phase growth. Toxin-specific transcription had increased dramatically in the 12-h culture. This relatively high level of transcription was maintained in the RNA isolated from the 18-h cultures; however, there appeared to be greater degradation in the total RNA from the 18-h culture.

## DISCUSSION

### Transcription of the toxigenic element

In order to clarify transcription of the toxin genes, I undertook an in-gel hybridization analysis of RNA from *Clostridium difficile* strain 10463. The direct

in-gel hybridization was employed rather than a standard Northern blot analysis, since the anticipated transcripts are quite large (the ORF for toxin A is 8.1 kb and the ORF for toxin B is 7.1 kb), and, in general, large transcripts are not transferred as efficiently as are smaller transcripts (Ahmad et al., 1990; Ehtesham and Hasnain, 1991).

The results of the in-gel hybridizations with probes specific for each of the open reading frames show that transcription could be detected for four of the five open reading frames of the toxigenic element. The sizes for each of the four transcripts correspond well with the published sizes of their respective open reading frames. The ORFtxe1 transcript measured 700 bp in length and the size of the ORFtxe1 open reading frame is 555 bp (Chapter 1). The open reading frame for the toxin B gene is 7098 bp in length (Barroso et al., 1990), and a 7.4 kb transcript was detected with the toxin B probe. ORFtxe2, which has several potential translational start sites has been estimated to be 500 bp (Dove et al., 1990). Our hybridization analysis reveals a 1 kb transcript for ORFtxe2. The 8.4 kb transcript which was detected for toxin A compares closely with its 8.1 kb open reading frame (Dove et al., 1990).

The primer extension data allow even greater accuracy in correlating the size of the four individual genes with their respective transcripts. A transcription initiation site 236 bp upstream of the translational start site for ORFtxe1 is identified with the extension reaction for ORFtxe1. The primer extension for toxin B shows

Table 2. Size of Toxigenic Element Transcripts

| | ORFtxe1 | toxin B | ORFtxe2 | toxin A |
|---|---|---|---|---|
| | | Nucleotides | | |
| Size of deduced open reading frame[a] | 555 | 7098 | 501 | 8133 |
| Predicted size of transcript[b] | 944 | 7351 | 1028 | 8486 |
| Actual transcript size[c] | 700 | 7400 | 1000 | 8400 |

[a] Nucleotide sequence for toxin B, Barroso *et al.* (1990); ORFtxe2 and toxin A from Dove *et al.* (1990). ORFtxe1, this study

[b] Using primer extension data

[c] As measured on agarose gels

that processing occurs approximately 150 bp 3' to the open reading frame for ORFtxe1. The 172 bp region downstream of toxin B (upstream of ORFtxe2) was not analyzed by primer extension; however, a DNA sequence with significant similarity to the processing sites upstream of toxins B and A is found here. In addition, Eichel-Streiber et al. (1992) and Eichel-Streiber (1993) have proposed a stem-loop structure to serve as a transcription terminator for toxin B which is contiguous to this sequence. The primer extension analysis for toxin A revealed a processing site 233 bp upstream of toxin A. Eichel-Streiber and Sauerborn (1990) and Eichel-Streiber (1993) proposed a pallindromic sequence to serve as the transcriptional terminator for toxin A. This is located 68 bp downstream of toxin A. These data were combined with the open reading frame size to predict a transcript size for each of the four open reading frames. The sizes of the open reading frames, the predicted transcipt size, and the estimated transcript size as measured on gels are presented in Table 2.

Additional transcripts hybridized with the four toxigenic probes. A very large transcript, estimated to be 15.5 kb in length was detected with each of the four probes. However, a probe specific for the glutamate dehydrogenase gene of *C. difficile* (Lyerly et al., 1991), which was used as a control, failed to hybridize to a band in this size range (data not shown, see Fig. 3 in Chap. 3). Therefore, I conclude that this band corresponds to a full length transcript and is not the result of DNA contamination. The sum of the predicted sizes of the four individual

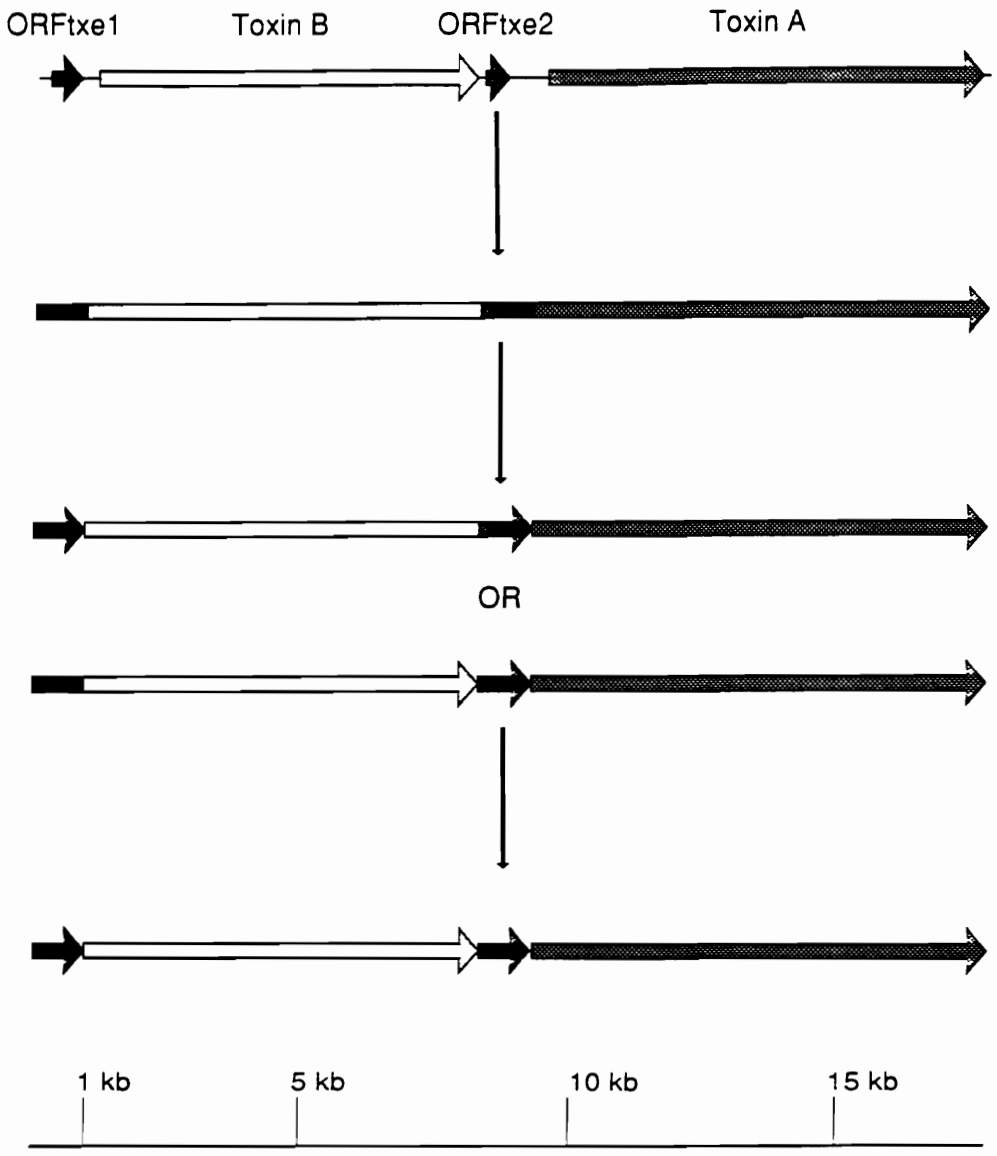


**FIG. 7. Proposed model for transcription and processing of the toxigenic element.** The full-length transcript is initially processed into toxin A and a processing intermediate. Toxin B comprises a processing intermediate with either ORFtxe2 or ORFtxe1. These are further processed into individual transcripts.

transcripts is 17.8 kb. This value is higher than the 15.5 kb full length transcript measured on the gels. The 15.5 kb size for the full length transcript probably represents an underestimate since the largest size marker that I used was 12 kb. Three of the four toxigenic probes, ORFtxe1, toxin B, and ORFtxe2 also hybridized with an 8.1 kb band.

The transcription data fit a model (Fig. 7) in which the toxigenic element is transcribed as a very large, full length transcript comprised of four contiguous open reading frames in the following order of transcription: ORFtxe1, toxin B, ORFtxe2, toxin A. Analysis of our hybridization data with the toxin A probe shows that the transcript for toxin A is immediately processed from the full length transcript. Conversely, the transcript for the toxin B gene comprises part of a processing intermediate which can consist of either one of the two following: ORFtxe1 linked with toxin B or toxin B coupled with ORFtxe2. This processing intermediate model fits well with the hybridization data for each of the ORFtxe1, ORFtxe2, and toxin B probes in which an 8.1 kb band was detected along with the full length transcript (see Figs. 2, 3). Each of the individual transcripts is ultimately derived from the processing intermediates. Thus, a 7.4 kb band was detected with the toxin B probe, a 700 bp transcript with the ORFtxe1 probe, and a 1 kb transcript with the ORFtxe2 probe.

Young et al. (1989a) in their evaluation of the clostridial genes which were characterized at the time, stated that most clostridial genes were monocistronic.

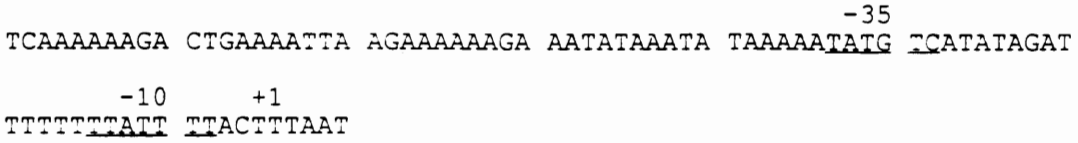


**FIG. 8. Promoter region for toxigenic element.** Transcription initiation (+1), -10 and -35 regions are indicated.

They then qualified their statement by enumerating examples of polycistronic transcripts, one of the most well-known being the nitrogenase genes of *Clostridium pasteurianum* (Chen et al., 1986). Recently, Perelle et al. (1993) have cloned the iota-toxin genes from *Clostridium perfringens*. The two genes (Ia and Ib) are located in close proximity to each other and only a single promoter region was identified by DNA sequence inspection. Therefore, they conclude that the iota-toxin genes form a polycistronic message. The model for *Clostridium difficile* transcription which is proposed herein specifies a polycistronic transcript for the toxigenic element as well.

The primer extension analyses which I performed were directed at three regions within the toxigenic element. The first region is situated upstream of ORFtxe1. The primer extension product identifies a transcription initiation site 236 bp upstream of ORFtxe1 and a promoter region is designated with the DNA sequences TATTTT and TATGTC, corresponding to the -10 and -35 promoter regions, respectively (Fig. 8). In terms of DNA sequence and spacing these compare well with a clostridial consensus promoter proposed by Young et al. (1989b). RNA polymerases with a number of sigma factors, each showing different promoter specificity, have been demonstrated in the Gram-positive bacteria (Doi and Wang, 1986). This implies that any consensus promoter, in particular, one that might be derived from sequences identified by inspection may be of limited usefulness (Young et al., 1989a).

The region extending upstream of the promoter for the toxigenic element (-40 to -60) is quite AT-rich with several stretches consisting entirely of A (see Fig. 8). AT-rich regions, in particular, A tracts, are found in a number of bacterial promoters, including some Gram positive promoters (Graves and Rabinowitz, 1986; Frisby and Zuber, 1991). Recently, Ross et al. (1993), working with several *E. coli* promoters, have identified AT-rich sequences in the -40 to -60 region as binding sites for the α subunit of RNA polymerase. These data support that the region mapped by primer extension is the promoter for the full length transcript.

Considerable RNA secondary structure exists downstream of ORFtxe1, toxin B, and ORFtxe2. Although a portion of the secondary structure may be due to the AT-richness of these intergenic regions, which on average is 80 percent A + T, these pallindromic sequences could serve as potential processing sites for RNAse cleavage. Analysis of the remaining two sites which were mapped by primer extension (upstream of toxin B and toxin A) indicate extensive sequence identity in regions of the promoter not typically conserved. This sequence similarity is maintained in a region upstream of ORFtxe2 (see Fig. 5), where RNAse processing also occurs. These data support the processing model which I have proposed.

Conversely, promoters for the toxin A gene, the toxin B gene, and ORFtxe2, as well as transcription terminators for toxin A and B have been proposed based solely on sequence analysis (Eichel-Streiber and Sauerborn, 1990; Eichel-Streiber et al., 1992; Eichel-Streiber, 1993). Interestingly, the pallindromic sequence proposed

as the transcription terminator for toxin B overlaps the promoter sequence for ORFtxe2. My primer extension products for toxins A and B mapped in the same region as the promoters proposed by Eichel-Streiber. Primer extension analyses are used to determine both transcription initiation and processing sites; however, they do not possess the potential to differentiate between these two possibilities. Therefore, since promoter-like sequences have been identified upstream of all four of the open reading frames, a model of transcription of the toxigenic element must encompass the fact that each of these genes could potentially be transcribed individually, in addition to being cotranscribed as part of a large polycistronic transcript, and thereafter processed through intermediates into individual transcripts.

Initially, I used PYG or BHI batch cultures for the isolation of RNA. When total RNA from these batch cultures was used in hybridization analyses, detection of either toxin A or toxin B was minimal. As an alternative to the batch cultures, I used a BHI dialysis flask culture which has been used for production of large amounts of toxin protein (Lyerly and Wilkins, 1991). The dialysis flask cultures yielded much more toxin-specific transcript per μg total RNA; however, in cultures grown for periods greater than 18 h, RNA extraction became very difficult due to the polysaccharide associated with the bacterial cells.

The hybridization analyses with RNA isolated from cultures at various stages of growth indicated that toxin A- and B-specific transcription increased dramatically subsequent to exponential growth. The effect of bacterial growth phase on toxin A

and B expression may be due to either a change in some environmental factor brought on by culture growth or to the presence or absence of a trans-acting factor whose own expression is modified in stationary phase growth.

With regard to late log or early stationary stage induction, Sauer and Durre (1992) examined codon usage in the sequenced genes of *Clostridium acetobutylicum*. They demonstrate that the ACG threonine codon is quite rare and is limited to either those genes which are involved in metabolism or uptake of minor carbon or nitrogen substrates or those genes expressed toward the end of exponential growth. The gene for glutamate dehydrogenase from *C. difficile* (Lyerly et al., 1991), an enzyme which is continually expressed, has no ACG codons. Interestingly, both the gene for toxin A and the gene for toxin B from *C. difficile* have five ACG codons (Barroso et al., 1990; Dove et al., 1990).

## LITERATURE CITED

1. Ahmad, I., J. A. Finkelstein, and A. W. Steggles. 1990. The analysis of RNA by *in situ* agarose gel hybridization is more sensitive than the equivalent Northern blot analysis. Biotechniques 8: 162-165.

2. Alam, J., R. A. Whitaker, D. W. Krogmann, and S. E. Curtis. 1986. Isolation and sequence of the gene for ferredoxin I from the cyanobacterium *Anabena* sp. strain PCC 7120. J. Bacteriol. 168: 1265-1271.

3. Barroso, L. A., S.-Z. Wang, C. J. Phelps, J. L. Johnson, and T. D. Wilkins. 1990. Nucleotide sequence of the *Clostridium difficile* toxin B gene. Nucleic Acids Res. 18: 4004.

4. Borriello, S. P., B. W. Wren, S. Hyde, S. V. Seddon, P. Sibbons, M. M. Krishna, S. Tabaqchali, S. Manek, and A. B. Price. 1992. Molecular, immunological and biological characterization of a toxin A negative, toxin B positive strain of *Clostridium difficile*. Infect. and Immun. 60: 4192-4199.

5. Chaudhari, N. 1991. Plasmid fragments as accurate markers for Northern blots. Biotechniques 10:601-605.

6. Chen, K. C.-K., J.-S. Chen, and J. L. Johnson. 1986. Structural features of multiple *nif* H-like sequences and very biased codon usage in nitrogenase genes of *Clostridium pasteurianum*. J. Bacteriol. 166: 162-172.

7. Dailey, D. C. and R. H. Schloemer. 1988. Identification and characterization of *Clostridium difficile* promoter element that is functional in *Escherichia coli*. Gene 70: 343-350.

8. Doi, R. H. and L. F. Wang. 1986. Multiple prokaryotic ribonucleic acid polymerase sigma factors. Microbiol. Rev. 50: 227-243.

9. Dove, C. H., S.-Z. Wang, S. B. Price, C. J. Phelps, D. M. Lyerly, T. D. Wilkins, and J. L. Johnson. 1990. Molecular characterization of the *Clostridium difficile* toxin A gene. Infect. and Immun. 58: 480-488.

10. Ehtesham, N. Z. and S. E. Hasnain. 1991. Direct in-gel hybridization, without blotting, using nick-translated cloned DNA probe. Biotechniques 11:718-721.

11. Eichel-Streiber, C. v. 1993. Molecular biology of the *Clostridium difficile* toxins, pp. 264-289. *In* M. Sebald (ed.), Genetics and molecular biology of anaerobic bacteria. Springer-Verlag, New York.

12. Eichel-Streiber, C. v., R. Laufenberg-Feldmann, S. Sartingen, J. Schulze, and M. Sauerborn. 1990. Cloning of the *Clostridium difficile* toxin B gene and demonstration of high N-terminal homology between toxin A and B. Med. Microbiol. Immuno. 179:271-279.

13. Eichel-Streiber, C. v., R. Laufenberg-Feldmann, S. Sartingen, J. Schulze, and M. Sauerborn. 1992. Comparative sequence analysis of the *Clostridium difficile* toxins A and B. Mol. Gen. Genet. 233: 260-268.

14. Eichel-Streiber, C. v. and M. Sauerborn. 1990. *Clostridium difficile* toxin A carries a C-terminal repetitive structure homologous to the carbohydrate binding region of streptococcal glycosyltransferases. Gene 96: 107-113.

15. Eichel-Streiber, C. v., D. Suckau, M. Wachter, and U. Hadding. 1989. Cloning and characterization of overlapping DNA fragments of the toxin A gene of *Clostridium difficile*. J. Gen. Microbiol. 135: 55-64.

16. Frisby, D. and P. Zuber. 1991. Analysis of the upstream activating sequence and site of carbon and nitrogen source repression in the promoter of an early-induced sporulation gene of *Bacillus subtilis*. J. Bacteriol. 173: 7557-7564.

17. Graves, M. C. and J. C. Rabinowitz. 1986. In vivo and in vitro transcription of the *Clostridium pasteurianum* ferredoxin gene. J. Biol. Chem. 261: 11409-11415.

18. Holdeman, L. V., E. P. Cato, and W. E. C. Moore (eds.), Anaerobe Laboratory Manual. 1977. VPI Anaerobe Laboratory, Virginia Polytechnic Institute and State University, Blacksburg, VA.

19. Johnson, J. L. 1991. DNA reassociation experiments. pp. 21-41 *In* E. Stackebrandt and M. Goodfellow (eds.), Nucleic Acid Techniques in Bacterial Systematics. John Wiley & Sons Ltd., West Sussex, England.

20. Johnson, J. L., C. H. Dove, S. B. Price, T. W. Sickles, C. J. Phelps, and T. D. Wilkins. 1988. The toxin A gene of *Clostridium difficile*. pp. 115-123. *In* J. M. Hardie and S. P. Borriello (eds.), Anaerobes today. John Wiley & Sons Ltd., West Sussex, England.

21. Johnson, J. L., C. Phelps, L. Barroso, M. D. Roberts, D. M. Lyerly, and T. D. Wilkins. 1990. Cloning and expression of the toxin B gene of *Clostridium difficile*. Curr. Microbiol. 20: 397-401.

22. Kroczek, R. A. and E. Siebert. 1990. Optimization of Northern analysis by vacuum-blotting, RNA-transfer visualization, and ultraviolet fixation. Anal. Biochem. 184: 90-95.

23. Lyerly, D. M., L. A. Barroso, and T. D. Wilkins. 1991. Identifiaction of the latex test-reactive protein of *Clostridium difficile* as glutamate dehydrogenase. J. Clin. Microbiol. 29: 2639-2642.

24. Lyerly, D. M., L. A. Barroso, T. D. Wilkins, C. Depitre, and G. Corthier. 1992. Characterization of a toxin A-negative, toxin B-positive strain of *Clostridium difficile*. Infect. and Immun. 60: 4633-4639.

25. Lyerly, D. M. and T. D. Wilkins. 1986. Commercial latex test for *Clostridium difficile* toxin A does not detect toxin A. J. Clin. Microbiol. 23: 622-623.

26. Lyerly, D. M. and T. D. Wilkins. 1991. Toxins of Anaerobes, pp. 163-181. *In* P. N. Levett (ed.), Anaerobic microbiology: a practical approach. IRL Press, Oxford.

27. Marmur, J. 1961. A procedure for the isolation of deoxyribonucleic acid from microorganisms. J. Mol. Biol. 3: 209-218.

28. McMillin, D. E., L. L. Muldrow, S. J. Leggette, Y. Abdulahi, and U. M. Ekanemesang. 1991. Molecular screening of *Clostridium difficile* toxins A and B genetic determinants and identification of mutant strains. FEMS Microbiol. Lett. 78: 75-80.

29. Perelle, S., M. Gibert, P. Boquet, and M. R. Popoff. 1993. Characterization of *Clostridium perfringens* iota-toxin genes and expression in *Escherichia coli*. Infect. and Immun. 61: 5147-5156.

30. Price, S. B., C. J. Phelps, T. D. Wilkins, and J. L. Johnson. 1987. Cloning of the carbohydrate-binding portion of the toxin A gene of *Clostridium difficile*. Curr. Microbiol. 16: 55-60.

31. Puissant, C. and L.-M. Houdebine. 1990. An improvement of the single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Biotechniques 8: 148-149.

32. Ross, W., K. K. Gosink, J. Salomon, K. Igarashi, C. Zou, A. Ishihama, K. Severinov, R. L. Gourse. 1993. A third recognition element in bacterial promoters: DNA binding by the αsubunit of RNA polymerase. Science 262:1407-1413.

33. Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239: 487-491.

34. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, New York.

35. Sanger, F., S. Nickelsen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74: 5463-5467.

36. Sauer, U. and P. Durre. 1992. Possible function of $tRNA^{thr}_{ACG}$ in regulation of solvent formation in *Clostridium acetobutylicum*. FEMS Microbiol. Lett. 100: 147-154.

37. Sauerborn, M. and C. v. Eichel-Streiber. 1990. Nucleotide sequence of *Clostridium difficile* toxin A. Nucleic Acids Res. 18: 1629-1630.

38. Torres J. F. 1991. Purification and characterization of toxin B from a strain of *Clostridium difficile* that does not produce toxin A. J. Med. Microbiol. 35: 40-44.

39. Wren B., S. R. Heard, and S. Tabaqchali. 1987. Association between production of toxins A and B and types of *Clostridium difficile*. J. Clin. Pathol. 40: 1397-1401.

40. Wren, B. W., P. Mullany, C. Clayton, and S. Tabaqchali. 1989. Nucleotide sequence of a chloramphenicol acetyl transferase gene from *Clostridium difficile*. Nucleic Acids Res. 17: 4877.

41. Young, M., N. P. Minton, and W. L. Staudenbauer. 1989b. Recent advances in the genetics of the clostridia. FEMS Microbiol. Rev. 63: 301-326.

42. Young, M., W. L. Staudenbauer, and N. P. Minton. 1989a. Genetics of *Clostridium*, pp. 63-103. *In* N. P. Minton and D. J. Clarke (eds.), Clostridia. Plenum Press, New York.

CHAPTER THREE: Quantitative Toxin-specific Transcript Differences Among Toxigenic Strains are Unaccompanied by Significant DNA Sequence Differences

## INTRODUCTION

Toxigenic strains of *Clostridum difficile*, a Gram positive anaerobic bacterium, make two very large toxin proteins, designated A and B. Nontoxigenic strains of *C. difficile* lack the toxin genes (Fluit et al., 1991). The toxin A and toxin B genes are separated by a distance of 1.4 kb, and together with two additional small open reading frames comprise the toxigenic element (Chapter 1).

Toxigenic strains of *C. difficile* can vary in toxin production over a one million-fold range (Lyerly and Wilkins, 1986). The reason for this variation in toxin production remains to be elucidated. The toxins are assumed to be coregulated since toxigenic strains which produce high amounts of toxin A also produce high amounts of toxin B. Similarly, toxigenic strains which produce relatively low levels of the toxin A protein also produce relatively low levels of the toxin B protein (Lyerly and Wilkins, 1986; Wren et al., 1987).

Six toxigenic strains, designated as either of high, intermediate, or low toxigenicity were selected for the present study. The purpose of the study was to determine whether there were nucleic acid differences among these six strains which might account for the differences in toxin production. Restriction fragment length

polymorphisms were examined, toxin-specific transcripts were analyzed, and sequences of regulatory regions were determined and compared among these toxigenic strains.

## MATERIALS AND METHODS

### Bacterial strains and growth conditions

*Clostridium difficile* strains were obtained from either the Anaerobe Collection of the Department of Biochemistry and Anaerobic Microbiology, or from the laboratory collection of Dr. Tracy Wilkins, Department of Biochemistry and Anaerobic Microbiology, Virginia Polytechnic Institute and State University, Blacksburg, VA.

| Toxigenic | Nontoxigenic |
|---|---|
| 10463 | 2037 |
| 26689 | 11186 |
| 7698 | 5199 |
| 11011 | |
| Tox R | |
| 4474 | |

For DNA isolation the *C. difficile* strains were grown anaerobically in PYG medium (Anaerobe Laboratory Manual) supplemented with 1% dehydrated BHI (Difco) at

37°C. Cultures for the isolation of RNA were grown anaerobically at 37°C in BHI dialysis flasks as described in Lyerly and Wilkins (1991).

DNA isolation and manipulations

Genomic DNA was isolated by the method of Marmur (1961) as modified by Johnson (1991).

Restriction endonucleases were purchased from either Promega or BRL and used as described in Sambrook et al. (1989).

DNA was electrophoresed on 11 X 14 cm 0.7% agarose (SeaKem LE, FMC Bioproducts) gels in TAE buffer (Sambrook et al., 1989).

DNA was labelled as a hybridization probe using either $\alpha^{35}$S-dATP, 1000 Ci/mMole from DuPont NEN Research Products and the BRL kit for Nick Translation or $\alpha^{32}$P-dATP, 3000 Ci/mMole from DuPont NEN Research Products and either the Prime-It or the Prime-It II random primer labelling kit from Stratagene.

Southern transfer followed the method of Southern (1975), using 0.45 μm nitrocellulose membranes from Schleicher and Schuell. The buffer for hybridization of Southern transfer membranes with radioactively labelled probes consisted of the following components in their final concentrations: 25% deionized formamide, 0.5% sodium dodecyl sulfate (SDS), 5X Denhardt's Reagent (Sambrook et al., 1989), 6X

Standard Saline-Phosphate-EDTA (SSPE) (Sambrook et al., 1989). Chemicals were obtained from either Fisher Scientific or Sigma. Hybridizations were carried out overnight at 50°C. Posthybridization buffer consisted of 2X SSPE, 0.5% SDS heated to 50°C. Initial posthybridization washes were carried out at 50°C on a slowly moving shaker (50-70 rpm). Final posthybridization washes were at room temperature. Membranes were dried and exposed to Kodak SB5 X-ray film.

RNA isolation and manipulations

Total RNA was isolated according to the guanidinium thiocyanate method of Puissant and Houdebine (1990) with the following exceptions. Once the initial cell pellet was resuspended in sucrose buffer, cells were lysed with lysozyme. Following phenol-chloroform extraction and isopropanol precipitation, the pellet was resuspended in Tris EDTA (TE) buffer, pH 7.5. A sufficient volume of 6 M lithium chloride was added to make the final solution 4 M LiCl for the LiCl extraction step. The final RNA pellet was dissolved in a small volume of diethyl pyrocarbonate (DEP)-treated water to which was added 4 volumes of 95% ethanol to make an RNA-ethanol solution (Sambrook et al., 1989) which was stored at -20°C. RNA concentration was measured at 260 nm on a Gilford 2400 spectrophotometer.

To obtain RNA samples for electrophoresis, 25 μg quantities of RNA were removed from the RNA-ethanol solution and precipitated with 0.1 volume of 3M

sodium acetate. RNA sample preparation followed the method of Kroczek and Siebert (1990). RNA samples were electrophoresed on 12.7 X 15.9 cm 1.1% formaldehyde, 1X MOPS, 1% agarose (SeaKem LE, FMC Bioproducts) gels at 70 Volts for 4-5 hours on a horizontal gel electrophoresis apparatus from Jordan Scientific Co. which recirculated the 1X MOPS electrophoresis buffer. Chemicals for the isolation of RNA and for RNA electrophoresis were obtained from Amresco, Gibco BRL, Fisher Scientific, or Sigma.

When electrophoresis was complete, the RNA gel was placed on three layers of Whatman 3MM paper and dried under vacuum in a Bio-Rad slab gel dryer Model 1125B, initially, for 30 min. at room temperature and then, for 2 hours at 60°C (Ahmad et al., 1990).

To prepare the dried gel for hybridization, it was briefly hydrated in a glass baking dish filled with DEP-treated water in order to remove the Whatman paper backing. The gel was then placed in hybridization buffer which consisted of the following components in their final concentrations: 25% deionized formamide, 0.5% SDS, 5X Denhardt's Reagent (Sambrook et al., 1989), 6X SSPE (Sambrook et al., 1989). Chemicals were obtained from either Fisher Scientific or Sigma. Heat-denatured, radioactively labelled probe was added (1 X $10^6$ cpm/ml) and hybridizations were carried out at 50°C overnight. Hybridizations were performed in a Techne Hybridiser HB-1 oven using FHB 11 bottles with 8 ml hybridization buffer.

All of the posthybridization buffers were heated to 50°C. Posthybridization washes took place at 37°C with slow agitation, and consisted of the following solutions: 2X Standard Saline Citrate (SSC), 0.1% SDS for 10 min.; 0.5X SSC, 0.1% SDS for 20 min.; 0.1X SSC, 0.1% SDS, 3 times at 40 min. each; and 0.1X SSC for 30 min. Gels were covered with Saran Wrap and exposed to Kodak SB5 film at -80°C.

For use as markers, the 1 kb ladder (BRL) was treated in the same manner as the RNA samples and electrophoresed along with them (Chaudhari, 1991). A separate aliquot of the ladder DNA was radioactively labelled and added to the hybridization mixture along with the RNA-specific probe. RNA sizes were determined using a linear regression analysis.

RNA samples for dot blots were removed from the RNA-ethanol solution and precipitated with 0.1 volume of 3 M sodium acetate. RNA pellets were then resuspended in DEP-treated water and denatured at 50°C for 90 sec. The RNA samples were loaded onto a nitrocellulose membrane (0.45μm) in a Schleicher and Schuell dot blot manifold in an excess of 20X SSC and thereafter prepared according to the method outlined in Sambrook et al. (1989). Once the RNA was loaded onto the nitrocellulose membranes, the membranes were crosslinked in 10X SSC in a Stratagene UV Stratalinker 2400. Hybridizations and posthybridizations were carried out as previously described for RNA gels. Individual dots were counted in Ecolume scintillation fluid (ICN Biomedicals) in a Beckman LS 8100

scintillation spectrometer.

Toxin protein analyses

Enzyme-linked immunosorbent assay (ELISA) titers for *Clostridium difficile* toxin A and cytotoxicity assays for toxin B were performed as described in Ehrich et al. (1980) and Lyerly et al. (1989).

PCR amplifications

DNA fragments which were used as hybridization probes were Polymerase Chain Reaction (PCR) products. PCR amplifications were performed on a GTC-1 Genetic Thermal Cycler from Precision Scientific using Taq DNA Polymerase from Perkin Elmer and in general followed the method of Saiki et al. (1988). Oligonucleotide primers were synthesized on a 381A DNA Synthesizer from Applied Biosystems. The symmetric amplifications were run for a total of 40 cycles. The 50 µl reaction contained 100 ng of genomic DNA along with 1.25 Units of Taq DNA Polymerase. The final concentration of oligonucleotide primer was 0.5 µM. Initially, the template was denatured at 95°C for 1.5 min., followed by annealing at 45°C for 1 min. For the remaining 39 cycles the denaturing temperature was lowered to 94°C for 30 sec., followed by annealing for 30 sec. at

45°C. The elongation step for all 40 cycles was at 70°C, but ranged in duration from 1-3 min. based on the length of the PCR product.

Symmetric PCR products were also generated for the three regulatory regions of the toxigenic element in the manner outlined above. Oligonucleotide primers for the regulatory regions were as follows:

BREGS1 5'-CTAGACAAGCTGTTAATAAGGC

BREGS2 5'-CAATCGAAGAGCAAATTAATC

BREGA1 5'-CATCTTCTTGAGTACGAAATC

BREGA2 5'-GATTAATTTGCTCTTCGATTG

AREGS1 5'-GATGAAATGAATGCTAAGGATG

AREGS2 5'-CTATTAAGCTACATTAGTTAC

AREGS3 5'-GTCAATTAACTAATTGAGTATC

AREGA1 5'-GGTCTAATGCTATATGCGAG

AREGA2 5'-GTAACTAATGTAGCTTAATAG

AREGA3 5'-GATACTCAATTAGTTAATTGAC

ELEREGS1 5'-GGGTAGTAAATTGGGAAGAAGC

ELEREGS2 5'-GTCATCTCTATAGTAATCTAG

ELEREGS3 5'-GAAAAGATTTTAATTTAATGATTG

ELEREGS4 5'-CTTTATTAATATAAAGTTTATTGC

ELEREGA1 5'-GCAAATCATCTACTGAGTTATTTC

ELEREGA2 5'-GCAATAAACTTTATATTAATAAAG

ELEREGA3 5'-CAATCATTAAATTAAAATCTTTTC

ELEREGA4 5'-CTAGATTACTATAGAGATGAC

BREGS1 and BREGA1, AREGS1 and AREGA1, and ELEREGS1 and ELEREGA1 were the oligonucleotide pairs used in the PCR amplification reactions. The remaining oligonucleotides were used as sequencing primers with the asymmetric PCR product templates.

DNA sequencing

Sequence-specific oligonucleotide primers were designed for each regulatory region (see above). These were used for single-stranded sequencing with T7 DNA Polymerase, Sequenase, Version 1.0, from USB. Sequencing reactions followed the methods established for single-stranded sequencing with Sequenase by DuPont which is a modification of the Sanger DNA sequencing techniques (Sanger et al., 1977). Sequencing reactions were run on 6% polyacrylamide gels on the DuPont Genesis 2000 Sequencer. DNA sequences were analyzed with Kodak MacVector software.

RESULTS

Restriction fragment length polymorphisms

In order to examine whether restriction fragment length polymorphisms exist within six toxigenic strains which range from high to low toxigenicity, genomic DNA was prepared from each of these strains and from several nontoxigenic strains. The genomic DNA from each strain was digested with each one of a set of the following restriction endonucleases: *Pst*I, *Eco*RI, *Eco*RV, *Hin*cII, and *Hin*dIII. Southern transfers of these restriction endonuclease-digested genomic DNAs were hybridized with toxin-specific probes derived from the primary clones for toxins A and B. Variability among the DNAs should be apparent as restriction fragment length polymorphisms. Essentially, no variability was detected among five of the six strains examined (data not shown); however, a restriction fragment length polymorphism was found using the pCD19L probe (see Fig. 1 in Chap. 1) in *Clostridium difficile* strain 26689 genomic DNA digested with *Hin*dIII, as shown in Fig. 1. This polymorphism is situated within the coding region of the toxin B gene.


**FIG. 1. Southern blot analysis of Hind III-digested chromosomal DNA from six toxigenic strains of *C. difficile* hybridized with pCD19L probe.** Two μg of DNA from each strain were applied to the gel. Lane 1, strain 10463; lane 2, strain 11011; lane 3, strain 4474; lane 4, strain ToxR; lane 5, strain 7698; lane 6, strain 26689. Sizes are indicated by arrows.

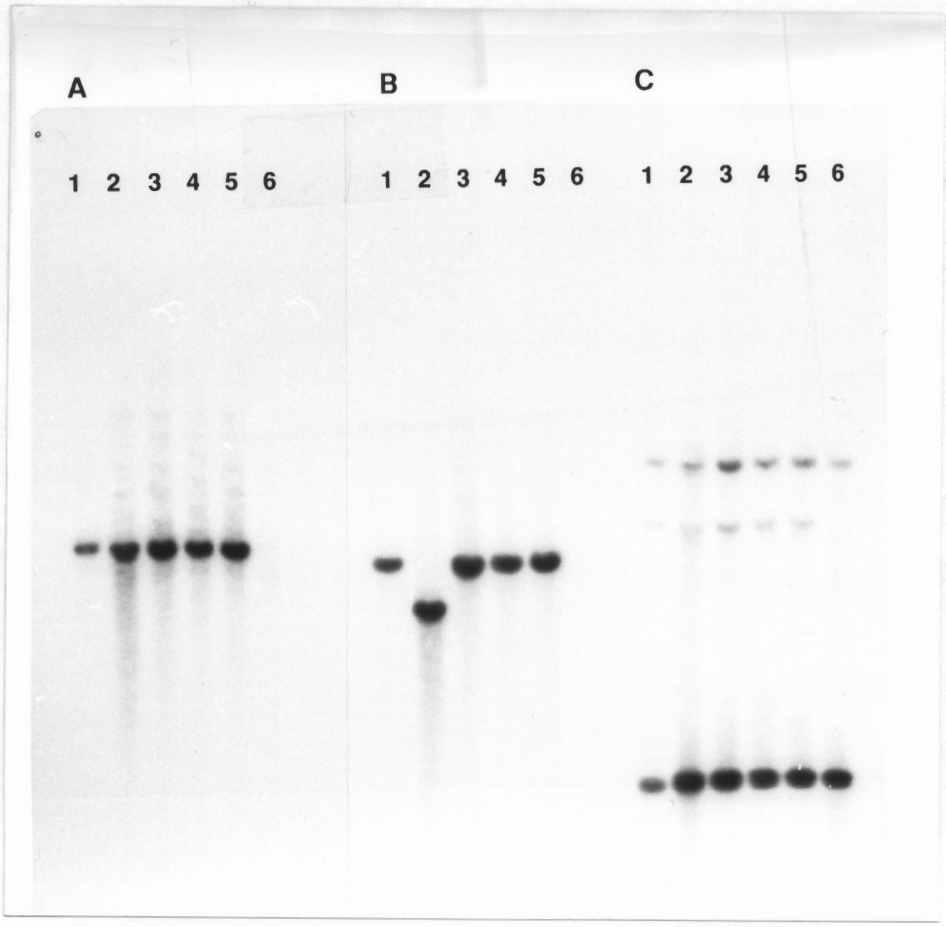


**FIG. 2. Southern blot analysis of genomic DNA from 5 toxigenic and 1 nontoxigenic strains of *C. difficile* using PCR amplified probes for toxin A, toxin B, and glutamate dehydrogenase. A. Toxin A probe.** Lane 1, strain 10463; lane 2, strain 26689; lane 3, strain 7698; lane 4, strain 11011; lane 5, strain 4474; lane 6, strain 2037. **B. Toxin B probe.** Samples as in A. **C. Glutamate dehydrogenase probe.** Samples as in A.

Transcription of toxins A and B in toxigenic strains

Using sequence-specific primers, PCR amplification products were generated for toxin A and toxin B. As a positive control in the hybridization experiments, a PCR product was generated and used as a probe for the glutamate dehydrogenase gene of *Clostridium difficile*. All three of these PCR products were approximately 1 kb in length. Initially, as a control for subsequent RNA hybridizations, the probes were utilized in Southern hybridizations of toxigenic strain genomic DNA digested with *Hin*dIII. The results of the Southern hybridizations are presented in Fig. 2 and demonstrate no differences among the toxigenic strains with respect to the hybridization probes with the exception that the toxin B probe hybridizes in the region of the *Hin*dIII polymorphism in strain 26689 (see Fig. 1).

RNA was isolated at 12 hr. of dialysis culture growth from five toxigenic strains which vary from high to low toxigenicity. Concurrently, samples were removed to quantitate toxin A and toxin B protein. The results of in-gel hybridizations with these RNAs using the toxin-specific probes are seen in figure 3. With the toxin A probe transcripts were detected in three of the five toxigenic strains, corresponding to the three strains highest in toxigenicity. The sizes of the transcripts were identical, and correlated well with the previously determined size of the toxin A transcript from strain 10463 (Chapter 2). Transcripts were detected in the two strains highest in toxigenicity in hybridizations with the toxin B probe.

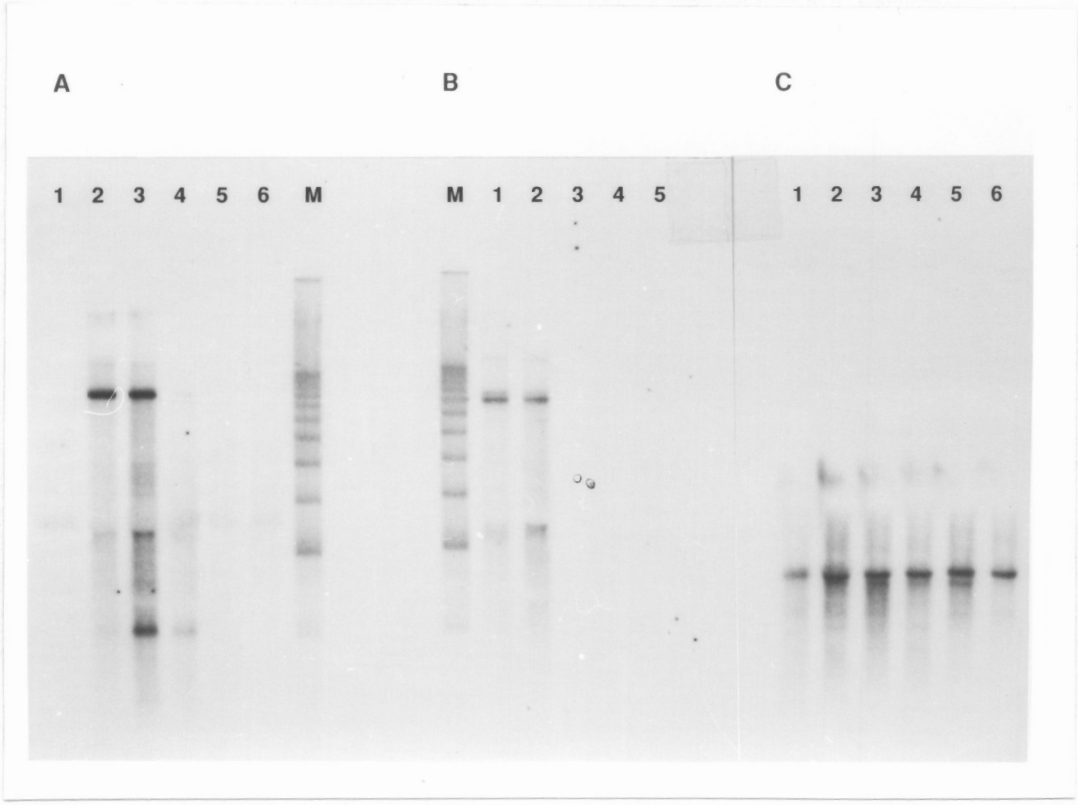


**FIG. 3. In-gel hybridizations of total RNA from dialysis cultures of 5 toxigenic and 1 nontoxigenic strains of *C. difficile* with PCR- amplified probes to toxin A, toxin B, and glutamate dehydrogenase.** Twenty five μg of total RNA were electrophoresed for each sample. **A. Toxin A probe.** Lane 1, strain 2037; lane 2, strain 10463; lane 3, strain 26689; lane 4, strain 7698; lane 5, strain 11011; lane 6, strain 4474. **B. Toxin B probe.** Lane 1, strain 10463; lane 2, strain 26689; lane 3, strain 7698; lane 4, strain 11011; lane 5, strain 4474. **C. Glutamate dehydrogenase probe.** Lane 1, strain 10463; lane 2, strain 26689; lane 3, strain 7698; lane 4, strain 11011; lane 5, strain 4474; lane 6, strain 2037. Marker is 1 kb ladder (M).

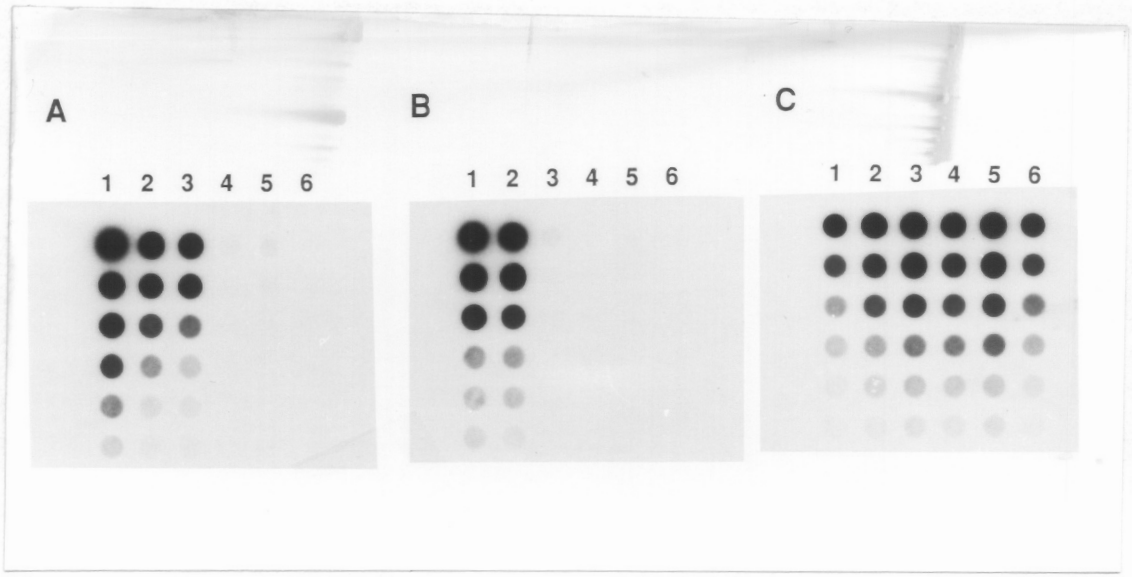


**FIG. 4. RNA dot blot hybridization of serial two-fold dilutions of RNA from 5 toxigenic and 1 nontoxigenic strains of *C. difficile* with PCR-amplified probes for toxin A, toxin B, and glutamate dehydrogenase.** Serial two-fold dilutions of RNA started with 25 μg. **A. Toxin A probe.** Lane 1, strain 10463; lane 2, strain 26689; lane 3, strain 7698; lane 4, strain 11011; lane 5, strain 4474; lane 6, strain 2037. **B. Toxin B probe.** Samples as in A. **C. Glutamate dehydrogenase probe.** Samples as in A.

Table 1. Quantitation of RNA Dot Blot Hybridizations[a]

| | *C. difficile* Strains | | | | | |
|---|---|---|---|---|---|---|
| | 10463 | 26689 | 7698 | 11011 | 4474 | 2037 |
| | cpm | | | | | |
| Toxin A | 26630 | 9697 | 6691 | 182 | 229 | 94 |
| | 7923 | 4300 | 3581 | 114 | 127 | 74 |
| | 4663 | 2130 | 1456 | 76 | 82 | 58 |
| | 2486 | 1013 | 612 | 53 | 59 | 49 |
| | 1228 | 417 | 308 | 40 | 47 | 33 |
| | 535 | 214 | 200 | 38 | 36 | 27 |
| Toxin B | 19533 | 15076 | 202 | 67 | 72 | 61 |
| | 9330 | 6656 | 127 | 48 | 49 | 41 |
| | 4381 | 3229 | 72 | 38 | 38 | 28 |
| | 1059 | 945 | 41 | 38 | 31 | 26 |
| | 672 | 610 | 33 | 32 | 31 | 29 |
| | 464 | 298 | 33 | 23 | 25 | 24 |
| Glutamate Dehyd-rogenase | 4136 | 6574 | 9100 | 6894 | 8794 | 4237 |
| | 2031 | 3450 | 5607 | 3366 | 5441 | 2245 |
| | 994 | 1994 | 2907 | 2098 | 2764 | 1420 |
| | 680 | 994 | 1386 | 1263 | 1620 | 841 |
| | 286 | 560 | 737 | 611 | 685 | 373 |
| | 155 | 252 | 393 | 328 | 459 | 196 |

[a] Two-fold serial dilutions starting with 25 µg total RNA

Table 2. Toxin A and Toxin B Titers[a]

| *C. difficile* Strains | Toxin A ELISA Titer[b] | | Toxin B Cytotoxic Titer[c] | |
|---|---|---|---|---|
| | Measured | Reported[d] | Measured | Reported[d] |
| 10463 | $10^2$ | $10^4$ | $10^4$ | $10^6$ |
| 26689 | $10^3$ | $10^3$ | $10^4$ | $10^5$ |
| 7698 | $10^1$ | $10^1$ | $10^2$ | $10^3$ |
| 11011 | $10^1$ | - | $10^2$ | $10^2$ |
| 4474 | - | - | $10^1$ | $10^1$ |
| 2037 | - | - | - | - |

[a] Cell samples from same cultures as in Table 1

[b] ELISA titers assays were performed as per Lyerly *et al.* (1989)

[c] Cytotoxic assays were as described by Ehrich *et al.* (1980)

[d] Reported values from Lyerly and Wilkins (1986)

Likewise, the size of these two transcripts agreed with the size of the toxin B transcript determined for strain 10463 (Chapter 2).

RNA dot blot analyses were undertaken using these same RNA samples and toxin-specific probes. Hybridizations of serial two-fold dilutions of total RNA with the toxin A probe, the toxin B probe, and the glutamate dehydrogenase probe are shown in Fig. 4. Following autoradiography, individual dots from the dot blots were counted in a scintillation counter. The data for these cpm are found in Table 1. Quantitation data for toxins A and B from these same samples are presented along with previously published toxin A and B titers for each of the strains in Table 2.

PCR amplification of regulatory regions in strains which vary in toxigenicity

Oligonucleotide primers were selected which allowed me to generate Polymerase Chain Reaction (PCR) products for three intergenic regions of the toxigenic element which lie upstream of ORFtxe1, upstream of the toxin B gene, and upstream of the toxin A gene. These regions were selected because each contains potential regulatory signals for transcription and translation. The regions are designated Tox Ele Reg, Tox B Reg, and Tox A Reg, corresponding to upstream of ORFtxe1, upstream of toxin B, and upstream of toxin A, respectively.

PCR amplification reactions utilizing these primers and genomic DNA from

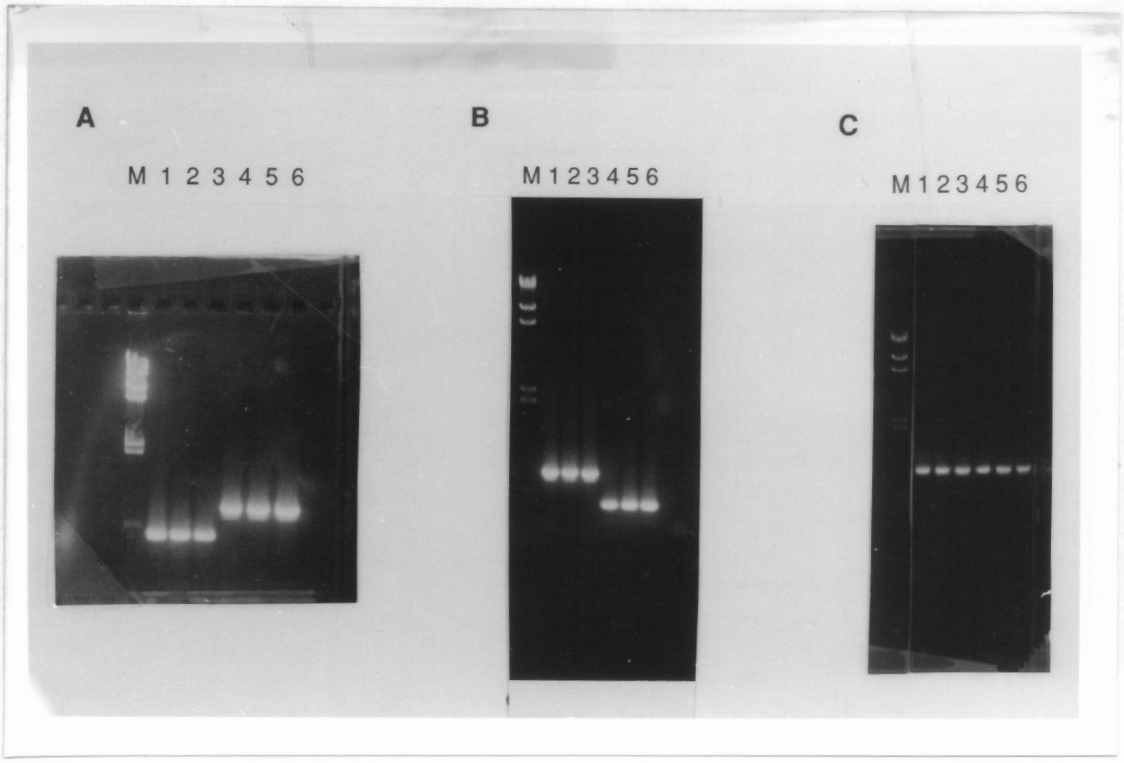


**FIG. 5. Agarose gels of PCR products from three regulatory regions of the toxigenic element in 6 toxigenic strains of *C. difficile*. A. TOX B REG and TOX A REG in 3 strains.** Lane 1, strain 10463, Tox B Reg; lane 2, strain 26689, Tox B Reg; lane 3, strain 7698, Tox B Reg. Lanes 4-6, samples as in lanes 1-3, Tox A Reg. **B. TOX A REG and TOX B REG in 3 strains.** Lane 1, strain ToxR, Tox A Reg; lane 2, strain 11011, Tox A Reg; lane 3, strain 4474, Tox A Reg. Lanes 4-6, samples as in lanes 1-3, Tox B Reg. **C. TOX ELE REG in 6 strains.** Lane 1, strain 10463; lane 2, strain 26689; lane 3, strain 7698; lane 4, strain ToxR; lane 5, strain 11011; lane 6, strain 4474. Marker is λ Hind III (M).

six toxigenic strains were undertaken. The amplification products for each of the regions appeared to be identical in length in each of the six toxigenic strains examined (Fig. 5).

Sequencing regulatory regions in strains which vary in toxigenicity

Two strains of *Clostridum difficile*, one intermediate in relative toxigenicity, strain 7698, and one with very low levels of both toxins A and B, strain 4474, were selected for sequencing of the three regulatory regions: Tox Ele Reg, Tox B Reg, and Tox A Reg. The sequence of *C. difficile* strain 10463 served as a basis of comparison.

The symmetric PCR amplification products from the three regulatory regions in each strain were used as templates to generate single-stranded asymmetric products for sequencing. The sequence for the 800 bp Tox A Reg region from *C. difficile* strain 7698 was identical to that of strain 10463, whereas the Tox A Reg sequence from strain 4474 differed by only 1 bp from that of strain 10463 (data not presented). Within the 350 bp Tox B Reg region two nucleotides differed in both *C. difficile* strain 7698 and strain 4474 relative to strain 10463 (Fig. 6). Both the location and the base change are conserved in each strain. Several conserved changes were found in the 1100 bp Tox Ele Reg region in each of the two *C. difficile* strains, 7698 and 4474. These are shown in Fig. 7.

## DISCUSSION

This study was undertaken to determine whether there were significant changes in the DNA within the toxigenic element which would account for a one-million fold difference in toxin production among toxigenic strains. Six such strains which vary in toxin production over this broad range (Lyerly and Wilkins, 1986) were selected for study. The RNA analyses were conducted with five of the same six strains. The toxigenic element is conserved among these six strains. Polymerase Chain Reaction (PCR) amplification of both the upstream (5') and the downstream (3') boundaries of the toxigenic element yielded products of identical length for each of these strains (Chapter 1).

The results of the restriction fragment length polymorphism analysis point to the fact that there are no major rearrangements of the toxigenic element DNA in these six strains. Even in the case of *C. difficile* strain 26689, which exhibits the only polymorphism which was detected, all remaining restriction sites were conserved. In addition, the PCR amplification product of the region upstream of toxin B from strain 26689 is apparently identical in length to amplification products from 5 other strains (Fig. 7 in Chap. 1). These data indicate that the polymorphism is due to a single base change which gives rise to a *Hin*dIII restriction site. The regions which were examined for polymorphisms, for the most part, consisted of the genes themselves, since the hybridization probes were fragments of the primary

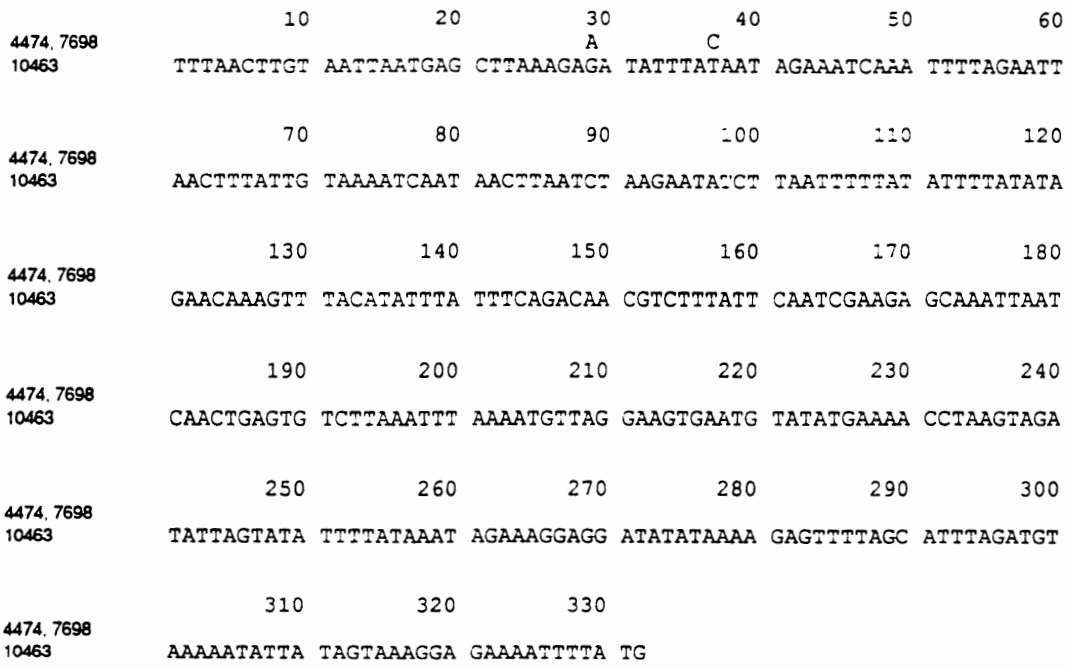


**FIG. 6. Nucleotide sequence of Toxin B Regulatory (Tox B Reg) region.** Sequence presented is the sense strand of strain 10463. Nucleotides that differ from this sequence in strain 7698 or strain 4474 are indicated.

```
                10         20         30         40         50         60
4474, 7698                                                   G
10463      CATATTTTCA TATAAAATTT AATTTATTTG CCGATTATAT AATTATAATT ACTGATTTAA

                70         80         90        100        110        120
4474, 7698                            A                               A
10463      TTCCAATGTT GTCAAAATTT TCAAATAAAG CATCATTATA AATATAAGAG AGGGTGATTT

4474, 7698
10463      TATG
```

**FIG. 7. Nucleotide sequence of Toxigenic Element Regulatory (Tox Ele Reg) region.** Sequence presented is the sense strand of strain 10463. Nucleotides that differ from this sequence in strain 7698 or strain 4474 are indicated. Sequence presented is the portion of the 1100 bp region where differences are found.

clones of toxins A and B. Lyerly et al. (1990) examined toxins A and B from twenty toxigenic strains of *Clostridum difficile* and found that in each case the toxin proteins were identical in size to the toxins from strain 10463.

Both the in-gel hybridizations and the RNA dot blot analyses show quantitative differences in toxin-specific messages among the toxigenic strains selected for this study. Since the RNA dot blot analysis is a more sensitive, quantitative measure of specific transcripts, toxin A transcripts are detected for all five of the toxigenic strains, albeit at a very low level for the two strains for which transcripts were not detected with the in-gel hybridizations. Likewise, transcripts for toxin B were detected in only two of the strains using in-gel hybridization. However, both the dot blot hybridization and the measured radioactivity indicate that toxin B transcripts are present in at least 3 of the strains.

Attempts to measure transcription accurately by quantitating levels of a specific mRNA are difficult. Furthermore, the toxin protein titer as measured in a culture supernatant is not necessarily an accurate gauge of transcription of the toxin gene. However, the transcription data and the RNA dot blot analyses do show a positive correlation between the abundance of toxin-specific transcript present in the RNA sample from a specific toxigenic strain and the titer of toxin protein from that same sample.

In the present study I have demonstrated quantitative differences in toxin-specific transcripts. The positive correlation between the abundance of toxin-specific

message and the toxin protein titer holds true for both toxins A and B from a specific toxigenic strain. When a probe for the glutamate dehydrogenase gene of *C. difficile* (Lyerly et al., 1991) is used in hybridizations, quantitative differences were, at maximum, two-fold (see Figs. 3, 4, Table 1). Therefore, these data indicate that at least a portion of the regulation of the toxin genes is at the transcriptional level.

Toxigenic strains that produce high levels of the toxin A protein, also produce high levels of toxin B protein. Likewise, toxigenic strains that produce relatively low levels of toxin A also produce relatively low levels of toxin B (Lyerly and Wilkins, 1986; Wren et al., 1987). I have demonstrated a very large transcript which is comprised of both the toxin B and the toxin A messages, in which processing occurs to yield the individual transcripts (Chapter 2). These data fit with a model in which the toxin genes are co-regulated.

An examination of the three regulatory regions in the toxigenic element in three toxigenic strains typified as low, intermediate, and high toxigenicity strains revealed that DNA sequences in the regions containing the transcriptional and translational regulatory signals contained very few changes among the three strains. These regions are intergenic regions and might be expected to demonstrate greater sequence variability than the toxin genes themselves. Some rare changes in the DNA sequence do exist, and in almost all cases the changes are conserved as to the location and the base change. For the most part these conserved changes are far removed from the regulatory signals per se. However, one conserved change is

located within the ribosome binding site of ORFtxe1, the open reading frame located at the 5' end of the toxigenic element (Fig. 7).

The processes whereby the amount of toxins A and B which are produced by a toxigenic strain of *C. difficile* are regulated are most likely quite complex. This study has demonstrated that at least a portion of the regulation of the toxin genes is at the transcriptional level. The restriction fragment length polymorphism data coupled with the regulatory region sequence analysis in several strains show that the DNA sequence of the toxigenic element is conserved and indicate that this transcriptional regulation is not due to DNA sequence differences in the regulatory regions. Instead, the transcriptional regulation may involve a trans-acting factor which may itself be subject to regulation.

## LITERATURE CITED

1. Ahmad, I., J. A. Finkelstein, and A. W. Steggles. 1990. The analysis of RNA by *in situ* agarose gel hybridization is more sensitive than the equivalent Northern blot analysis. Biotechniques 8: 162-165.

2. Chaudhari, N. 1991. Plasmid fragments as accurate markers for Northern blots. Biotechniques 10: 601-605.

3. Ehrich, M. R., L. van Tassell, J. M. Libby, and T. D. Wilkins. 1980. Production of *Clostridium difficile* antitoxin. Infect. and Immun. 28: 1041-1043.

4. Fluit, A. C., M. J. H. M. Wolfhagen, G. P. H. T. Verdonk, M. Jansze, R. Torensma, and J. Verhoef. 1991. Nontoxigenic strains of *Clostridium difficile* lack the genes for both toxin A and toxin B. J. Clin. Microbiol. 29: 2666-2667.

5. Holdeman, L. V., E. P. Cato, and W. E. C. Moore (eds.). 1977. Anaerobe Laboratory Manual. VPI Anaerobe Laboratory, Virginia Polytechnic Institute and State University, Blacksburg, VA.

6. Johnson, J. L. 1991. DNA reassociation experiments. pp. 21-41, *In* E. Stackebrandt and M. Goodfellow (eds.), Nucleic Acid Techniques in Bacterial Systematics. John Wiley & Sons Ltd. West Sussex, England.

7. Kroczek, R. A. and E. Siebert. 1990. Optimization of Northern analysis by vacuum-blotting, RNA-transfer visualization, and ultraviolet fixation. Anal. Biochem. 184: 90-95.

8. Lyerly, D. M., L. A. Barroso, and T. D. Wilkins. 1991. Identification of the latex test-reactive protein of *Clostridium difficile* as glutamate dehydrogenase. J. Clin. Microbiol. 29: 2639-2642.

9. Lyerly, D. M., J. L. Johnson, and T. D. Wilkins. 1990. The toxins of *Clostridium difficile*. pp. 137-146, *In* S. P. Borriello (ed.), Clinical and molecular aspects of anaerobes. Wrightson Biomedical Publishing Ltd., Petersfield.

10. Lyerly, D. M., C. J. Phelps, and T. D. Wilkins. 1989. Monoclonal antibodies against *Clostridium difficile* and its toxins: use in diagnostic microbiology. pp. 621-636, *In* B. Swaminathan and G. Prakash (eds.), Nucleic acid and monoclonal antibody probes. Applications in diagnostic microbiology. Marcel Dekker, Inc.

11. Lyerly, D. M. and T. D. Wilkins. 1986. Commercial latex test for *Clostridium difficile* toxin A does not detect toxin A. J. of Clin. Microbiol. 23: 622-623.

12. Lyerly, D. M. and T. D. Wilkins. 1991. Toxins of anaerobes. pp. 163-182, *In* P. N. Levett (ed.), Anaerobic microbiology: a practical approach. IRL Press, Oxford.

13. Marmur, J. 1961. A procedure for the isolation of deoxyribonucleic acid from microorganisms. J. Mol. Biol. 3: 209-218.

14. Puissant, C. and L.-M. Huidebine. 1990. An improvement of the single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Biotechniques 8: 148-149.

15. Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239: 487-491.

16. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning. A laboratory manual. Cold Spring Harbor Press. Cold Spring Harbor, New York.

17. Sanger, F., S. Nickelsen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74: 5463-5467.

18. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98: 503-517.

19. Wren, B., S. R. Heard, and S. Tabaqchali. 1987. Association between production of toxins A and B and types of *Clostridium difficile*. J. Clin. Pathol. 40: 1397-1401.

# CONCLUSIONS

Many toxin genes are either plasmid encoded or carried by bacteriophage. By establishing boundaries between toxigenic sequences and those sequences shared by nontoxigenic and toxigenic strains, this work confirms the chromosomal location of the toxin genes. The junctions of the toxigenic element were examined for evidence of insertion type sequences, but none was found. PCR amplification products that encompassed either the upstream or the downstream boundary were identical in length for each of six toxigenic strains, indicating that the boundaries are conserved among these strains.

In the three nontoxigenic strains examined, a short fragment (127 bp) occupies the same chromosomal location as the large 19.6 kb toxigenic element. This nontoxigenic fragment has terminal inverted repeat regions as well as internal repeat regions. These could potentially serve a role in recombination. The fact that the toxigenic element boundaries are conserved among six toxigenic strains, coupled with the presence of fragments of the toxigenic element insertion region which are identical in length in the three nontoxigenic strains examined indicates that the toxigenic element occupies a specific location on the chromosome.

Since the boundaries of the toxigenic element have been precisely identified in this work, it is now clear which open reading frames are associated with the toxin genes. One small open reading frame is identified for the first time in this work. It

is located upstream of toxin B.

Transcriptional analyses demonstrate that two open reading frames, in addition to the toxin A and toxin B genes are transcribed. A large transcript was detected in addition to processing intermediates, and individual transcripts. These data support a model in which the toxin A and B genes are cotranscribed along with the two remaining open reading frames, ORFtxe1 and ORFtxe2. The toxin A transcript is immediately processed from the full length transcript, leaving one of two processing intermediates. Either ORFtxe1 is removed, leaving toxin B joined with ORFtxe2, or, conversely, ORFtxe2 is processed, leaving toxin B attached to ORFtxe1. Ultimately, each of the four is present as an individual transcript.

A transcription initiation site and a promoter for the toxigenic element were identified as well as processing sites upstream of the toxin A and B genes. These data were used to generate an estimated size for the four individual transcripts. The estimated sizes correspond well with the measured sizes of these transcripts taken from hybridization analyses.

Transcription analyses using toxin A- and toxin B- specific probes revealed quantitative differences among strains, indicating that at least a portion of the variation in toxin production among these strains can be attributed to transcription. However, both DNA restriction fragment length polymorphism analysis and DNA sequencing of the regulatory regions in these strains showed remarkable DNA sequence conservation both within the toxin genes and in the intergenic regulatory

regions. Therefore, a trans-acting regulatory factor may be responsible for the quantitative differences in toxin-specific transcription observed among these strains.

# Curriculum Vitae

Georgia Ann Hammond

Department of Biochemistry and Anaerobic Microbiology
Virginia Polytechnic Institute and State University
Blacksburg, VA 24061-0305 USA
Telephone Number: 703-231-8682 Fax Number: 703-231-7126

Home Address: Kirkeveien 10
1430 Aas NORWAY

Birthdate: 20 January 1953

Marital Status: Married to Karl Pedersen

Educational Experience

1975 B. S. in Biology, College of William and Mary, Williamsburg, VA

1981 M. A. in Biology, College of William and Mary, Williamsburg, VA

1984-1987 Graduate studies in Biology, University of North Carolina at Chapel Hill, Chapel Hill, NC

1988-present Doctoral student in Department of Anaerobic Microbiology, Virginia Polytechnic Institute and State University, Blacksburg, VA

Work Experience

1976-1980 Laboratory Technician, Indiana University, Bloomington, IN

1981-1983 Instructor in Biology, Department of Biology, University of North Carolina at Greensboro, Greensboro, NC

1983-1984 Instructor in Biology, Program in General Biology, Washington State University, Pullman, WA

1988-1990 (summer sessions) Instructor in Biology, Department of Biology, Virginia Polytechnic Institute and State University, Blacksburg, VA