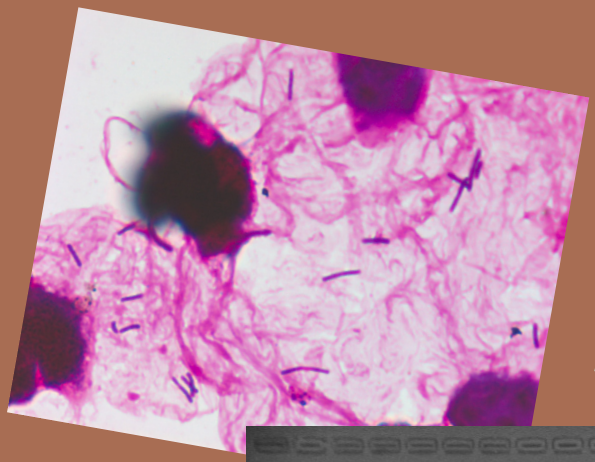
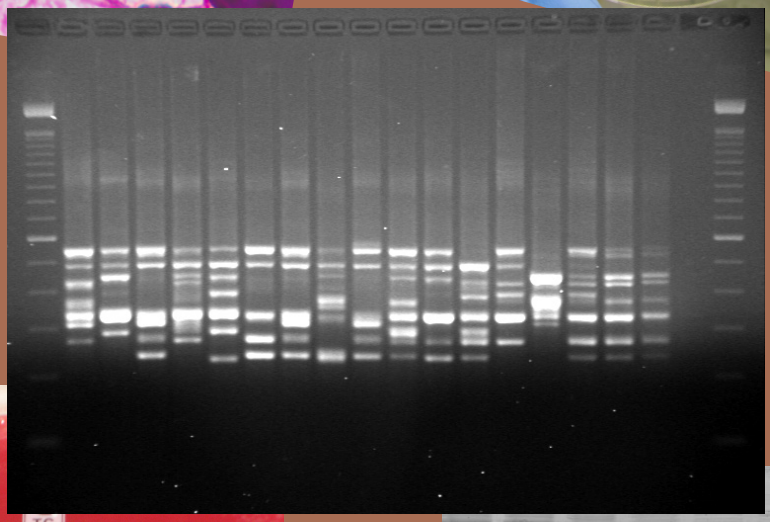
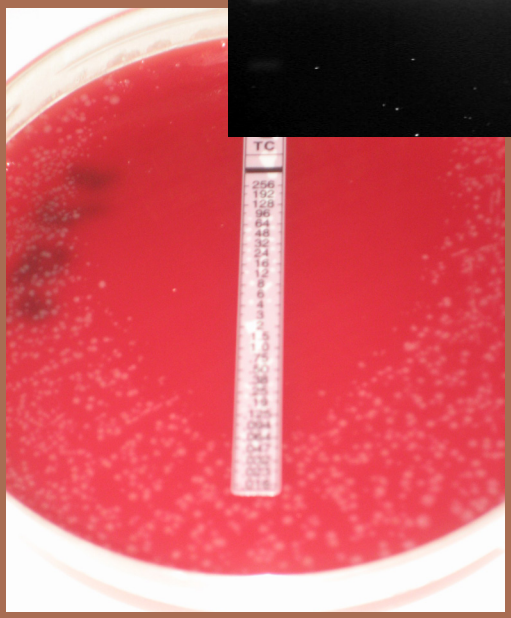
Molecular characterization of pathogenic Clostridium difficile strains

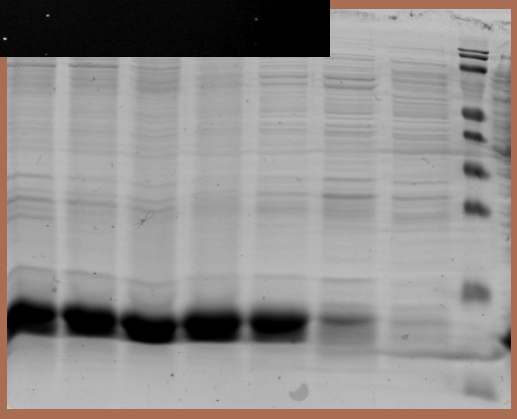
Bakker, D.

Molecular characterization of pathogenic *Clostridium difficile* strains









Dennis Bakker

Molecular characterization of pathogenic *Clostridium difficile* strains

Cover photos by Dennis Bakker
Giemsa stain of a monolayer of Caco-2 cellen with adhered *C. difficile*. Lincomycine agar plate with *C. difficile* ClosTron mutants. PCR-Ribotypering. SDS gel with whole lysate of different *C. difficile* strains. Tetracycline MIC determination for *C. difficile* strain 630.

Print & Lay-out: Mostert drukkerij
ISBN: 978-94-90858-32-2

Printing: 2014, Leiden, the Netherlands

Molecular characterization of pathogenic *Clostridium difficile* strains

Proefschrift

ter verkrijging van de graad van Doctor
aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdediging op woensdag 5 november 2014
klokke 13:45 uur

door

Dennis Bakker

Geboren 7 november 1975
te Gorinchem

Promotie commissie

Promotor:	Prof. Dr. E.J. Kuijper
Co-promotores:	Dr. ir. J. Corver Dr. W.K. Smits
Overige leden:	Prof. Dr. O.P. Kuipers (University of Groningen, The Netherlands) Prof. Dr. G. P. van Wezel Dr. A. P. Roberts (UCL Eastman Dental Institute, United Kingdom)

Liberi mei mea vita
Voor mijn kinderen, Mijn leven

Table of contents

Introduction

1

General introduction and outline of the thesis

1 General introduction

Clostridium difficile is an enteropathogenic anaerobic Gram-positive spore forming rod, which can cause a wide variety of symptoms. *Clostridium difficile* infection (CDI) is recognized as the leading cause of infectious nosocomial diarrhoea. In the last decade the incidence and severity of CDI has increased and currently *C. difficile* is also frequently recognized as an important cause of community-acquired diarrhoea in humans. Finally, *C. difficile* may be present in the intestinal tract of asymptomatic and diarrhoeal (farm) animals.

Clostridiaceae

A major part of the human gut microbiome consists of Firmicutes (1). The phylum Firmicutes was traditionally classified to include all Gram-positive bacteria, but recently it was re-classified as a bacterial core-group with low-G+C genomic DNA content (2). Many Firmicutes are able to form (endo)spores, which are highly resistant to disinfectants and extreme environmental conditions (2). The phylum of Firmicutes can be roughly divided into two main classes; the *Bacilli*, which are aerobic, and the *Clostridiaceae,* which are obligate anaerobic (3).

The genus *Clostridium* belongs to the class of *Clostridiaceae* which consists of approximately 100 species (4). *Clostridium spp.* are ubiquitously found in the environment, soil, water and in the human and mammalian gastro intestinal tract as part of the commensal microbial flora (5). The genus *Clostridium* comprises several well characterized important human pathogens such as: *Clostridium tetani*, *Clostridium botulinum*, *Clostridium perfringens*, *Clostridium sordellii* and *Clostridium difficile* (2,4,6). All diseases caused by *Clostridium spp*. are mediated by the secretion and action of the secreted (exo)toxins (2,4,6,7).

Clinical disease and epidemiology

Almost eighty years ago (1935), *Clostridium difficile* was first described by Hall and O'Toole as *Bacillus difficilis* (8,9). *Clostridium difficile* was isolated from faeces of new-born infants and described as an actively motile, heavy bodied Gram-positive rod with elongated subterminal or nearly terminal non-bulging spores (8). Injection of whole or filtrated cultures in guinea pigs and rabbits resulted in rapid death (8). Survival of the animals injected with boiled and filtrated cultures led to the conclusion that *C. difficile* was able to produce a soluble heat labile toxin (8).

More insights in the pathogenicity of *C. difficile* were obtained by investigations in germfree mice and rats (10). In addition, a prospective study by Tedesco *et al*, evaluating 200 patients, subsequently reported that clindamycin administration was associated with development of diarrhoea (21%) and pseudomembranous

colitis (PMC) (10%) (11). In retrospect this was the first study to describe *C. difficile* as a nosocomial human pathogen (11). In 1977, Larson and Price isolated a clostridial toxin from faeces of patients with antibiotic-associated colitis and PMC (12). Subsequently, Bartlett and colleagues reported that the causative agent of antibiotic-associated PMC was indeed a toxin producing *C. difficile* (13). The lack of sensitive and reliable detection and identification methods has delayed the recognition of *C. difficile* as an important nosocomial pathogen for a long time (14). However, nowadays *C. difficile* is recognized as the most frequent cause of infectious nosocomial diarrhoea worldwide (9,15,16).

Several reports indicate that up to 3% of healthy individuals and up to almost 40% of hospitalized patients are colonized with toxinogenic *C. difficile* strains (17-19). Healthy individuals are usually protected by their normal colonic microflora (17). However, disturbance of the colonic microflora by for instance antibiotic usage leads to loss of the colonization resistance, enabling *C. difficile* to colonize, overgrow and cause disease (17). Besides antibiotic usage, other risk factors for CDI are advanced age (>60 years), underlying diseases and infection pressure (20). During infection *C. difficile* can produce three toxins which affect the integrity of the colon epithelium by destruction of the tight junctions, disrupt the cytoskeleton and lead to the formation of protrusions (21-24).

The typical clinical features of CDI range from mild diarrhoea to fulminant colitis, which can be life-threatening (25-27). Endoscopic examination of patients suffering from CDI may reveal minor abnormalities to extensive pseudomembranes (25). The formation of these pseudomembranes is caused by deep haemorrhagic ulcerations of the colon epithelium and a massive host inflammatory response (28). Symptoms of pseudomembranous colitis are cramping, fever, hypo-albuminemia, leukocytosis, nausea and general malaise besides the typical lesions observed by endoscopy (25,26). Eventually, pseudomembranous colitis may evolve into a syndrome called toxic megacolon, the most serious form of disease caused by CDI (29). Toxic megacolon is a toxic dilatation of the colon causing paralysis of the peristaltic movement and systemic cytotoxicity (29) accompanied by sepsis and organ dysfunction and resulting in high (33%) mortality rates (29,30).

In the last decade the epidemiology of CDI has changed radically. North-America was the first to report numerous hospital outbreaks with an increased incidence of CDI cases with a more severe course of the disease (27,31,32). Later on reports of outbreaks in Europe followed (33-35). These epidemics were primarily due to *C. difficile* PCR Ribotype (RT) 027, also known as North America Pulse-field type (NAP) 1 and Restriction Endonuclease Analysis (REA) type B1 (027/NAP1/B1) strain. *Clostridium difficile* RT 027 strains were referred to as hypervirulent strains, due to reports of higher toxin production, higher relapse

rates, more severe course of the disease and increased mortality rates (19,27,36-39). Certain other *C. difficile* types (e.g. RT017 and 078) have also been reported to be associated with outbreaks and a more severe course of disease in a hospital setting (34,39-44). Besides the increase of CDI incidence rates in a health care setting, an increased incidence rate in a community setting was reported (45-47). In some studies *C. difficile* RT 078 was more frequently associated with community acquired CDI than other ribotypes and affected a younger population than for instance *C. difficile* RT 027 strains (48,49). The course of *C. difficile* RT 078 associated disease can be as severe as observed for the hypervirulent *C. difficile* RT 027 strain (48) and the incidence of *C. difficile* RT 078 has increased in the Netherlands and other European countries (34,50). Interestingly recent studies have demonstrated that *C. difficile* RT 078 is the predominant type in cattle and pigs, suggesting a zoonotic potential (51-54).

Diagnostic and Typing methods

Diagnosis of CDI

The emergence of *C. difficile* strains as an important nosocomial human pathogen has stimulated the development of better diagnostic, detection and typing methods. The diagnosis of CDI is primarily based on clinical symptoms in combination with laboratory assays (55). Culturing *C. difficile* and the detection of the toxin(s) with a cytotoxicity assay is regarded as the gold standard for diagnosis of CDI (28,55,56). However, in many diagnostic laboratories culturing and cell cytotoxicity assays are not performed routinely, as they are labour intensive and costly (55,57,58). Most diagnostic laboratories depend on rapid and easy-to perform enzyme immuno assays (EIAs) that are generally designed to detect *C. difficile* Toxins A and/or B (55,57). Unfortunately, EIAs have limited specificity and/or sensitivity in an endemic situation (55). The limitations of EIAs have tempted clinicians into testing multiple samples per patient and to develop two-step algorithms for CDI diagnosis (55,59). However, for prevention and management of CDI rapid and more accurate assays are essential.

Quantitative real-time PCRs (qPCRs) are potential rapid assays with better performances than EIAs (56,60-62). Additionally, molecular tests such as qPCR have high sensitivities and specificities, comparable to the cell cytotoxicity assays for diagnosis of CDI (56,60,63-66). Both commercial and in-house developed qPCRs for the detection of *C. difficile* are primarily based on the detection of the toxin B and/or toxin A genes (56,60,63-66). In some cases the detection of the toxin genes is combined with the detection of point mutations in the *tcdC* gene, or the binary toxin genes, i.e. *cdtA* and *cdtB* (56,60,63-66).

The disadvantage of molecular tests compared to conventional phenotypic tests such as the cytotoxicity assay is the lack of discrimination between asymptomatic and symptomatic carriership. Despite the improved performances (sensitivity and specificity), molecular tests have in general low positive predictive values and therefore cannot be applied as stand-alone test (64). However, due to the high negative predictive values molecular tests, such as qPCRs, are valuable as first screening method in a two-step testing algorithm for diagnosis of CDI (56,60).

Typing of *Clostridium difficile* strains

Typing methods have been developed to study the epidemiology of *C. difficile* for several reasons; for detection and investigation of outbreaks, to recognize new emerging strains, to investigate local and global transmission routes, and to study the phylogeny (67,68). One of the most crucial properties of a typing method is the capability to differentiate between strains (67-69). Other important features are reproducibility and ease of use (69,70). Typing methods can be differentiated into two groups: phenotypic assays and genome based assays. Phenotypic assays are based on gene expression products, whereas genotyping methods are based on the DNA content of the strain (67,69,70).

Until the 1980s, phenotypic assays were standard to determine relatedness of strains (28,69). The simple and rapid serotyping assay was the most commonly used phenotypic assay (69,71). Eventually the assay was able to discriminate 23 different serotypes (69). Several of these (A, G, K, S1 and S4) were associated with (severe) CDI and PMC caused by CDI (72). Other serotypes (D and Cd-5) were suggested to be associated with asymptomatic carriership (72). Other typing methods were the Radio PAGE (25 types) and immunoblotting assay (26 types) (69). Eventually, the phenotypic typing methods were abandoned due to low reproducibility and a low discriminatory power for epidemiological studies with a larger -than local- scope (69).

The emergence of the important global epidemic *C. difficile* RT 027 strain has stimulated the development of genotyping methods. Molecular typing methods rely on purified DNA of a single cultured *C. difficile* clone. Often these techniques are performed in national reference centres, since culturing of *C. difficile* is not routinely performed in most diagnostic labs (69). Molecular methods are generally more reproducible and sensitive compared to the phenotypic assays used in the 1980s (69). The most common genotyping methods to include PCR ribotyping, Multi Locus Variable number of tandem repeat Analysis (MLVA), Multi Locus Sequence Typing (MLST) and single nucleotide polymorphisms (SNP) typing (60).

PCR ribotyping is the most frequently used molecular typing method in Europe (67,73). This PCR band-size based typing method exploits the variability of the

intergenic spacer region between 16S and 23S ribosomal DNA, as well as rDNA copy numbers (73-75). In comparison to other DNA based typing assays PCR ribotyping is superior in discriminatory power, (inter laboratory) reproducibility and its low hands on time (67,70). Although PCR ribotyping is the preferred typing method for monitoring outbreaks, it is not able to differentiate between strains in an outbreak situation (70).

MLVA is a highly discriminatory (sub)typing method to discriminate between strains within one Ribotype to identify outbreak transmission routes (60,68,70,76). MLVA is based on the amplification of short tandem repeats that vary in size and are spread throughout the genome (77). Marsh *et al.*, were the first to describe an automated analysis of the number of tandem repeats per locus followed by van den Berg *et al.,* (78,79). The two independently developed MLVA typing methods contained four out of seven identical loci (78,79). Based on MLVA results, a minimum spanning tree can be constructed to determine genetic relationships among strains. Genetic relationships are based on the number of Summed Tandem Repeat Differences (STRD) (41,78). Despite the lack of variability of two loci in RT027, MLVA has been reported to be the most discriminatory subtyping method to investigate outbreaks with the epidemic *C. difficile* RT 027 strain (79,80).

The availability of multiple genome sequences has made whole genome sequencing (WGS) accessible as a typing method (77,81-83). Recently, major advances have been made with WGS-based typing of *C. difficile* strains (81,84). WGS is able to reveal the natural history, phylogeny and global spread of epidemic strains (81,82,84). In addition, WGS is a reliable sequencing method that can distinguish strains at a single nucleotide level (81,84). The identification of single nucleotide polymorphisms (SNPs) across sequenced genomes has the potential to improve the discriminatory power over the more traditional MLVA genotyping method (81,85). Although MLVA and WGS analyse different aspects of the *C. difficile* genome, a recent publication on the performance of both techniques resulted in similar discriminatory power in an outbreak situation (85). WGS is also capable of detecting mixed infections, which offers prospects of screening for mixed infections in transmission studies (86). To date, the costs of WGS are still relatively high compared to other typing methods (67,86). However, the ability of WGS to extrapolate PCR ribotyping, MLST, sequences of toxin and resistance genes and other additional data, combined with standardized computational pipelines, could balance the cost benefit towards WGS in the future (38,67,77,81-83).

Mobile elements

WGS as a typing method is based on the non-repetitive core genome (82,84,87). However, comparative genomics has demonstrated that *C. difficile* has a highly

mosaic genome which consists of a large proportion (11%) of mobile elements (77,88). These mobile elements from phage, plasmid or transposon origin constitute the so called accessory genome (83,89). Mobile DNA elements can be transferred between *C. difficile* strains and may be passed on to other bacteria (77,90). Many *C. difficile* conjugative transposons (CTn) contain genes or other functions that could potentially contribute to the fitness and or virulence of strains harbouring them (76,77,90-92).

Conjugative transposons are defined as "specific DNA elements that can integrate into one or more sites in one or more genomes" (93). In general, CTns consist of a core region (conjugation and regulation module), a recombination module and a module with accessory gene(s) (91,93,94). After excision from the genome, intermediates are formed that can be transferred through conjugation into a recipient strain, where they can integrate via specific target site recombination (91,93,94). The accessory module can be diverse, and the genes primarily present encode determinants for antibiotic resistance or metabolic functions (76,91,93-98), often accompanied by dedicated transcriptional regulators (94).

The genomic sequence of *C. difficile* strain 630 revealed the presence of six putative CTns in addition to the extensively investigated Tn*5397* element (76,77,95,99). The putative CTns are named based on their locations in the 630 genome (77). Four of the identified transposons are closely related to Tn*916* (77). The Tn*916*-like family of conjugative transposons generally encode tetracycline resistance determinants and is broadly found in Firmicutes (94). The other CTns present in *C. difficile* strain 630 are more closely related to Tn*1549*, a conjugative transposon responsible for vancomycin resistance (77,100). Transposons in other (non 630) *C. difficile* strains encode resistance to chloramphenicol, spectinomycin, streptomycin and erythromycin (81,96).

It is easy to envisage how the presence of antibiotic resistance determinants could increase the fitness of strains, but the contribution of other genes in the accessory genome to virulence remains largely unexplored. A recent study suggested that the accessory genome could be involved in the increased virulence and mortality of a sub-group of *Clostridium difficile* RT 015 strains (39). Furthermore, some virulence associated genes, including the *C. difficile* pathogenicity locus (PaLoc) have been suggested to be of phage origin (101-103). Moreover, it has been demonstrated that phages can affect toxin production levels in *C. difficile*, thereby modulating virulence (103,104).

1

Clostridium difficile virulence factors

Clostridium difficile toxins

The symptoms of CDI are mainly caused by the two main virulence factors, Toxin A (TcdA) and Toxin B (TcdB) (23,105-107). All naturally disease causing (toxinogenic) *C. difficile* strains produce Toxin B, and most of the strains also Toxin A (23,34,108,109). No naturally occurring Toxin B negative strains have been reported to date. Toxin A (308 kDa) and Toxin B (270 kDa) belong to the family of large clostridial cytotoxins (21,24). The members of this family are single chain toxins consisting of at least four domains (Figure 1): a receptor binding domain (110,111), a cysteine protease domain, a translocation domain (112,113) and a glycosyltransferase domain (21,23,24).

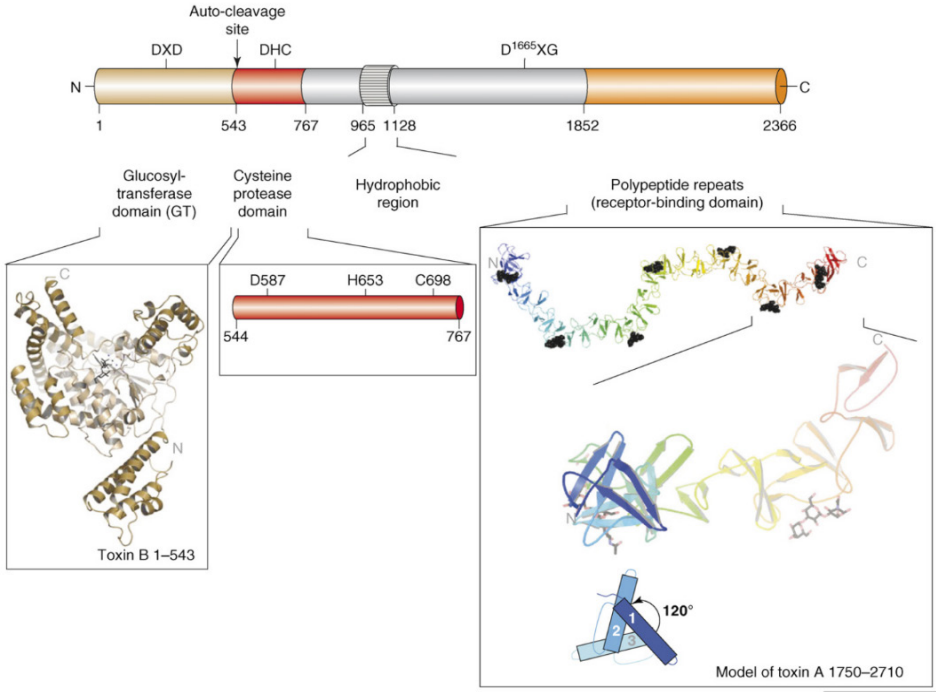


Figure 1: Schematic view of the domain structure of clostridial toxins. The catalytic domain is located at the N-terminus and the CROP region, which serves as receptor binding domain, is situated at the C-terminus. In between the N- and C-terminus the translocation domain (hydrophobic region) and the cysteine protease domain are located (113). Jank, T and Aktories K., Structure and mode of action of clostridial glycosylating toxins: the ABCD model.Trends Microbiol.16:222-229, Copyright © 2014 Elsevier Masson SAS. All rights reserved.

The receptor binding domain is located at the C-terminus and covers almost one third of the toxin, in the case of Toxin A (23,24). The C-terminus is characterized by repetitive peptide elements that are called combined repetitive oligopeptides (CROPS) (23,24,111). The crystal structure of the CROPs of Toxin A shows a solenoid like structure (110). Solenoid structures are widely present on bacterial surfaces and play a role in protein-protein or protein-carbohydrate interactions. The crystal structure of the CROP region in complex with a carbohydrate demonstrated a specific interaction (114). However, the carbohydrates used in the crystal structure are absent in sensitive cells and in human tissue (113,115) and other (human) glycan structures need to be tested for their role in toxin binding.

Despite a similar structure of the C-terminus of Toxin B, experimental data suggest that binding of Toxin A and B is mediated by different types of receptors (113,116). Interestingly, it has recently been shown that toxins can enter eukaryotic cells via a CROP independent mechanism (117). Although the biological relevance of this observation is not yet clear, it is suggestive of an alternative binding motif within the toxin (117).Binding of the toxin to the eukaryotic cell receptor results in receptor-mediated endocytosis (21,24). Internalization of the toxins occurs through the clathrin mediated pathway (118,119). After the clathrin coated vesicle is fused to an endosome (118), the low pH in the endosome results in a conformational change of the toxin characterized by an increased exposure of the hydrophobic region (21,24,115). This, in turn, allows translocation of the catalytic domain and the cysteine protease domain into the cytosol (115). Cytosolic inositol hexaphosphate (InsP6) can subsequently induce the autocleavage activity of the cysteine protease domain, resulting in the release of the catalytic domain in the cytosol (21,24,115). Interestingly, it has recently been shown that cytotoxicity of Toxin B can be independent of catalytic domain cleavage (120). The catalytic domain is a glycosyltransferase which inactivates Rho GTPases (Rho, Rac and Cdc42) in the cytosol by glycosylation (21,24,115,121). This results in the inhibition of multiple effectors that ultimately results in apoptosis of the targeted cell (21,23,24,113). A subsequent inflammatory response and further degradation of the intestinal epithelial cell layer eventually leads to the development of the clinical symptoms of CDI (23,24,122).

Transcriptional regulation of the Pathogenicity Locus

The genes encoding the major clostridial toxins, *tcdA* and *tcdB*, are located on a 19-kb genomic region called the Pathogenicity Locus (PaLoc) (23,24). In between the toxin genes the *tcdE* gene is situated, which encodes a holin-like protein (101,123,124). TcdE is a membrane associated protein (23,101) but its involvement in the release of the toxins is disputed (123,124). Besides the toxin genes and *tcdE*, the PaLoc also contains two regulatory genes, *tcdR* and *tcdC* (23,125,126).

The *tcd*R gene encodes the alternative RNA polymerase sigma (**σ**) factor, TcdR that positively regulates toxin production. It mediates binding of RNA polymerase core enzyme to the *tcdA*, *tcdB* and *tcdR* promoter regions (125-127). TcdR belongs to the sub-family of extra cytoplasmic factor (ECF) σ^{70} factors (125). In other bacteria, members of this family have also been demonstrated to positively regulate potent toxins, such as the tetanus neurotoxin of *Clostridium tetani* (128). TcdC has been reported to act like an anti-sigma factor for toxin production by destabilizing the TcdR-RNA polymerase core enzyme complex in a way that is not yet fully understood (126). Besides the PaLoc encoded regulators, toxin levels are also directly influenced by the nutritional regulators CodY and CcpA, and the sigma factors σ^{H} and σ^{D} (129-134).

In strain VPI10463, the exponential growth phase of *C. difficile* is associated with a high transcription level of the *tcdC* gene and low transcription levels of *tcdR* and the toxin genes, whereas the stationary growth phase is associated with a low transcription level of the *tcdC* gene and high transcription levels of *tcdR* and the toxin genes (135). The synthesis and secretion of the toxins is increased upon entry into the stationary growth phase (135-138). The decreasing transcription of *tcdC* correlates with diminishing TcdC protein levels in stationary growth phase (135,138,139). Interestingly, mutations in the *tcdC* gene in the hypervirulent strains RT 027 (**Δ117**) and RT 078 (C184T) have been linked to increased virulence (43,140). These mutations result in the absence of a functional TcdC caused by a frame shift mutation (Δ117 bp) or a premature stopcodon (C184T) in *tcdC* which is linked to an increased toxin production (19,140). Together, the reported inverse correlation between the transcription of *tcdC* and the toxin genes and the expression patterns of the corresponding proteins (TcdC, TcdA and TcdB), have led to the prevailing model that TcdC is an important repressor of toxin expression (125,126,135,139). However, recently some doubts have been raised about the importance of TcdC in the regulation of toxin expression. (137,138).

Some strains that are associated with an increased mortality and morbidity (e.g. RT 027 and RT 078) also produce a third toxin, called the binary toxin since it consists of two polypeptides encoded by the *cdtA* and *cdtB* genes (27,43). The binary toxin genes are located on a 6.2 kb region called the Cdt locus, or CdtLoc (141). A gene (*cdtR*) encoding for a regulatory protein is located alongside the binary toxin genes on the CdtLoc (141). It has been suggested that binary toxin may contribute to disease in hamsters (106) and *in-vitro* assays have demonstrated that the binary toxin affects adhesion of *C. difficile* to cells through induction of protrusion formation of the target cells (22,142).

Non-toxin virulence factors of *Clostridium difficile*

Clostridium difficile sporulation

Clostridium difficile is transmitted through the faecal-oral route. Spores are believed to form the primary vehicle for infection and sporulation is recognized as a crucial factor in the transmission route and as a persistence factor for *C. difficile* (143). Ingested spores germinate in the small bowel upon exposure to bile acids (144). The newly formed vegetative cells start to produce toxins and can cause disease (144).

The primary purpose for sporulation is survival of the bacteria under harsh environmental conditions for prolonged periods of time. Spores have a high tolerance against disinfectants and resistance towards antimicrobial compounds and therefore are able to persist on (hospital) surfaces and in the environment (145,146). The ability to form spores is a key factor to infect a new host, as they are able to survive the oxygenated environment outside the host as well as the harsh conditions in the host. Moreover, they confer the ability to persist whereas the commensal flora is eliminated by an antibiotic treatment (144,147).

The regulation of spore formation in *C. difficile* is still poorly characterized compared to the extensively studied spore forming *Bacillus subtilis* (148). Though the pathways downstream of the master regulator for sporulation, Spo0A, seem to a large extent conserved between *B. subtilis* and *Clostridia*, this is less so for the pathways upstream Spo0A activation (149-152). As expected, *spo0A* is required for sporulation in *C. difficile* (153). The sporulation pathway in *Bacillus subtilis* is induced under conditions of nutrition-starvation (148). Under these conditions phosphorylation of the master regulator Spo0A occurs (150,154). Phosphorylated Spo0A (Spo0A-P) is an essential activator of the sporulation signalling cascade in *Bacilli* which ultimately leads to the formation mature spores (149,150). Importantly, in other organisms Spo0A regulates many other processes than sporulation, such as stress responses, biofilm formation, competence for genetic transformation, and DNA replication but also the synthesis of virulence factors (155-164).

Clostridium difficile factors affecting colonization

Spores and toxins are established virulence factors that contribute to transmission, infection and development of disease (23,122,143,144). But an essential first step in establishment of *C. difficile* in the colon is adhesion to the mucosal surfaces and subsequent colonization of the host (Figure 2) (144). Relatively little is known about adhesion and colonization factors of *C. difficile*. The surface layer of *C. difficile* contains a whole arsenal of proteins that could potentially be involved in adhesion and colonization. Presently, a few surface proteins are identified or hypothesized to

play a role in adhesion to cells or colonization of the host. Among these proteins are fibronectin-binding protein A (165), Cwp84 (166), flagellar proteins (167) CD1581 (168), and several S-layer proteins (169).

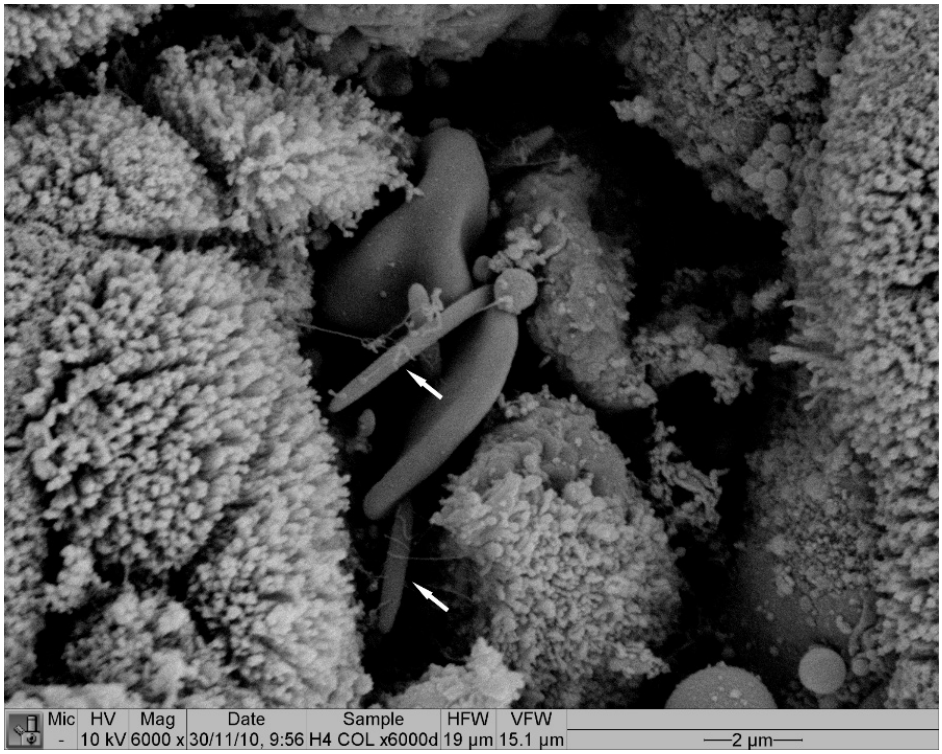


Figure 2: Visualisation (SEM picture) of an adhered vegetative *C. difficile* (arrows) to host colon epithelium (This picture was kindly provided by A.M. Buckley).

Mechanisms of stress survival

During its life cycle *C. difficile* encounters stressful environments, like antibacterial substances and proteins, elevated temperatures, extreme pH and osmotic stress. These extracellular stresses can result in the accumulation of (partially) unfolded proteins, especially in the bacterial envelope, which are non-functional or form poisonous aggregates (170). They may also affect the integrity of the bacterial cell envelope, compromising cell viability. Thus, sensing and responding to extracellular stresses is important for survival of the bacteria. Adaptation to these host-induced stresses can play a vital role in the virulence of bacteria and the establishment of an infection.

Regulated intramembrane proteolysis (RIP) is one of the mechanisms that enables bacteria to monitor and respond to extracellular stresses (Figure 3) (171).

The RIP mechanism activates regulatory pathways by releasing the sequestered sigma factors from their cognate anti-sigma factors as the result of envelope stress (171,172). In *Bacillus subtilis*, extracellular stresses are sensed by an site-1 protease, PrsW (172). Upon stress, PrsW undergoes conformational changes which allow cleavage of an anti-sigma factor at the C-terminus (172). The cleaved anti-sigma factor is recognized and further processed by extracytoplasmic proteases and a site-2 protease, RasP (171,172). The remaining part of the anti-sigma factor is further degraded in the cytoplasm by proteases like ClpX, leading to release of the sigma factor (171,172). The sigma factor then interacts with RNA polymerase to activate transcription of stress related genes (171,172).

Another well described mechanism to sense and respond to stress are two-component systems. Two component systems are composed of a sensor kinase and a response regulator (173,174). Monitoring envelope stress by the sensor kinase results in phosphorylation of the response regulator (174). This in turn leads to activation of stress related target genes (173,174). In *Bacillus subtilis* a two

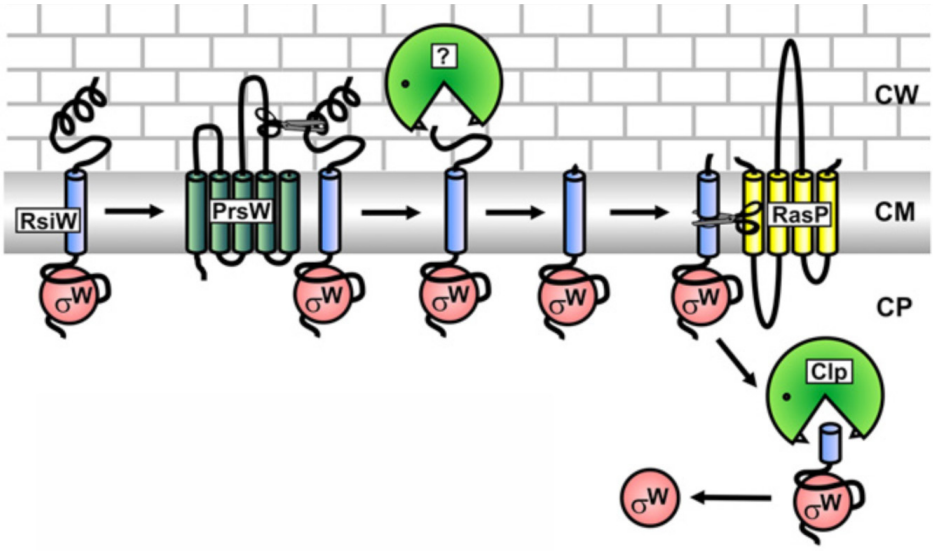


Figure 3: Schematic view of the RIP pathway in Bacillus *subtilis*. The stress induced PrsW degrades the C-terminus of the anti-sigma factor (RsiW) resulting in further processing of RsiW by extracytoplasmic proteases and RasP. The partially degraded anti-sigma factor is released into the cytoplasmic and further degraded by cytoplasmic proteases like Clp. The released sigma factor interacts with RNA polymerase which results in activation of stress related genes (171). Heinrich, J and Wiegert T. Regulated intramembrane proteolysis in the control of extracytoplasmic function sigma factors. Res. Microbiol.160:696-703

component system consisting of a sensor kinase (CssS) and a response regulator (CssR) plays a role in the detection of extracytoplasmic misfolded proteins (174). It does so primarily through the activation of expression of two homologous proteins of the High temperature requirement A (HtrA) family (174).

Bacillus subtilis HtrA and HtrB are associated with the cytoplasmic membrane and are composed of membrane spanning domain, a trypsin-like serine protease domain and one PDZ domain (174-178). Both proteins have fairly broad substrate specificity. PDZ domains are highly flexible domains and are involved in substrate recognition and/or the regulation of protease activity (177,179). Bacterial HtrA-like proteins are important for controlling protein quality and homeostasis by combining proteolytic and chaperone activities (175,176,179). Membrane anchored HtrA-like proteases are active as trimers and soluble HtrA-like proteases form larger active oligomers (176,179).

Outline of this thesis

This thesis covers a broad range of research topics varying from an applied approach to more fundamental experiments. **Chapter 1** is the current chapter where we provide a general overview.

Applied research

In **Chapter 2** we describe the evaluation of a modified in-house developed molecular assay with improved performance for diagnosis of CDI. In addition we evaluated two other in-house developed qPCRs, of which one detects the presence of both toxin genes and the other detects the frame shift mutation in the *tcdC* gene. We evaluated the qPCRs by comparing them to a commercially available molecular assay and to the gold standards. In **Chapter 3,** the relatedness between human and porcine *C. difficile* RT 078 strains is investigated using a modified and optimized MLVA. Moreover, we also investigate the susceptibility to tetracycline and determined the genetic origin of tetracycline resistance to further support the high relatedness between human and porcine RT 078 strains. In **Chapter 4** we identified a 100kb insert in a porcine RT 078 strain through whole genome sequencing and characterized this mobile element further.

Fundamental research

Chapter 5 reports that *C. difficile* Spo0A recognizes and binds to similar sequences as the well characterized *B. subtilis* Spo0A. The *in-vitro* binding assays suggest that *tcdB* may be a direct target of Spo0A. Therefore, we also characterized the toxin production of an isogenic Spo0A mutant strain. In **Chapter 6** we generated an isogenic *tcdC* mutant to investigate the possible role of TcdC in toxin production.

In **Chapter 7** we propose an alternative role for the putative anti-sigma factor TcdC. **Chapter 8** describes the identification and characterization of a *C. difficile* HtrA-like protease. The effect on the transcriptome and *in-vivo* virulence was determined by comparing the generated isogenic HtrA mutant to wild type.

Discussion

Chapter 9 provides the general discussion, including the main conclusions of this thesis and future perspectives. **Chapter 10** provides an overview of the content of this dissertation in Dutch.

Reference List

1. Ley, R. E., D. A. Peterson, and J. I. Gordon. 2006. Ecological and evolutionary forces shaping microbial diversity in the human intestine. Cell. 124:837-848.
2. Dürre P. Handbook on Clostridia. 2005.
3. Wolf, M., T. Muller, T. Dandekar, and J. D. Pollack. 2004. Phylogeny of Firmicutes with special reference to Mycoplasma (Mollicutes) as inferred from phosphoglycerate kinase amino acid sequence data. Int.J.Syst.Evol.Microbiol. 54:871-875.
4. Wells CL and Wilkins TD. Clostridia: Sporeforming Anaerobic Bacilli. 1996.
5. Cato E.P., .L.George , and S.M.Finegold. genus *Clostridium*. 2, 1141-1200. 1986. Bergey's manual of Systematic Bacteriology.
6. Hatheway, C. L. 1990. Toxigenic clostridia. Clin.Microbiol.Rev. 3:66-98.
7. Johnson, E. A. 1999. Clostridial toxins as therapeutic agents: benefits of nature's most toxic proteins. Annu.Rev.Microbiol. 53:551-75.:551-575.
8. Hall, I. C. O. 1935. Intestinal flora in new-born infants: with a description of a new pathogenic anaerobe, Bacillus difficilis. Am J Dis Child. 49:390-402.
9. Bartlett, J. G. 1994. *Clostridium difficile*: history of its role as an enteric pathogen and the current state of knowledge about the organism. Clin.Infect.Dis. 18 Suppl 4:S265-72.:S265-S272.
10. Hammarstrom, S., P. Perlmann, B. E. Gustafsson, and R. Lagercrantz. 1969. Autoantibodies to colon in germfree rats monocontaminated with *Clostridium difficile*. J.Exp.Med. 129:747-756.
11. Tedesco, F. J., R. W. Barton, and D. H. Alpers. 1974. Clindamycin-associated colitis. A prospective study. Ann.Intern.Med. 81:429-433.
12. Larson, H. E. and A. B. Price. 1977. Pseudomembranous colitis: Presence of clostridial toxin. Lancet. 2:1312-1314.
13. Bartlett JG, Chang TW, Gurwith M, Gorbach SL, and Onderdonk AB. 1978. Antibiotic-associated pseudomembranous colitis due to toxin-producing clostridia. N.Engl.J.Med.531-534.
14. Gerding, D. N. 2009. *Clostridium difficile* 30 years on: what has, or has not, changed and why? Int.J.Antimicrob.Agents. 33 Suppl 1:S2-8. doi: 10.1016/S0924-8579:-1.
15. McFarland, L. V. 1995. Epidemiology of infectious and iatrogenic nosocomial diarrhea in a cohort of general medicine patients. Am.J.Infect.Control. 23:295-305.
16. Polage, C. R., J. V. Solnick, and S. H. Cohen. 2012. Nosocomial diarrhea: evaluation and treatment of causes other than *Clostridium difficile*. Clin.Infect.Dis. 55:982-989.
17. Cartman, S. T., J. T. Heap, S. A. Kuehne, A. Cockayne, and N. P. Minton. 2010. The emergence of 'hypervirulence' in *Clostridium difficile*. Int.J.Med.Microbiol. 300:387-395.
18. McFarland, L. V. 2008. Update on the changing epidemiology of *Clostridium difficile*-associated disease. Nat.Clin.Pract.Gastroenterol.Hepatol. 5:40-48.
19. Warny, M., J. Pepin, A. Fang, G. Killgore, A. Thompson, J. Brazier, E. Frost, and L. C. McDonald. 2005. Toxin production by an emerging strain of *Clostridium difficile* associated with outbreaks of severe disease in North America and Europe. Lancet 366:1079-1084.
20. Dubberke, E. R., K. A. Reske, M. A. Olsen, K. M. McMullen, J. L. Mayfield, L. C. McDonald, and V. J. Fraser. 2007. Evaluation of *Clostridium difficile*-associated disease pressure as a risk factor for *C difficile*-associated disease. Arch.Intern.Med. 167:1092-1097.

21. Jank, T., T. Giesemann, and K. Aktories. 2007. Rho-glucosylating *Clostridium difficile* Toxins A and B: new insights into structure and function. Glycobiology 17:15R-22R.
22. Schwan, C., B. Stecher, T. Tzivelekidis, H. M. van, M. Rohde, W. D. Hardt, J. Wehland, and K. Aktories. 2009. *Clostridium difficile* toxin CDT induces formation of microtubule-based protrusions and increases adherence of bacteria. PLoS.Pathog. 5:e1000626.
23. Voth, D. E. and J. D. Ballard. 2005. *Clostridium difficile* toxins: mechanism of action and role in disease. Clin.Microbiol.Rev. 18:247-263.
24. Just, I. and R. Gerhard. 2004. Large clostridial cytotoxins. Rev.Physiol Biochem. Pharmacol. 152:23-47.
25. Bartlett, J. G. 2002. Clinical practice. Antibiotic-associated diarrhea. N.Engl.J.Med. 346:334-339.
26. Kelly, C. P., C. Pothoulakis, and J. T. LaMont. 1994. *Clostridium difficile* colitis. N.Engl.J.Med. 330:257-262.
27. Loo, V. G., L. Poirier, M. A. Miller, M. Oughton, M. D. Libman, S. Michaud, A. M. Bourgault, T. Nguyen, C. Frenette, M. Kelly, A. Vibien, P. Brassard, S. Fenn, K. Dewar, T. J. Hudson, R. Horn, P. Rene, Y. Monczak, and A. Dascal. 2005. A predominantly clonal multi-institutional outbreak of *Clostridium difficile*-associated diarrhea with high morbidity and mortality. N.Engl.J.Med. 353:2442-2449.
28. Bartlett, J. G. and D. N. Gerding. 2008. Clinical recognition and diagnosis of *Clostridium difficile* infection. Clin.Infect.Dis. 46 Suppl 1:S12-8. doi: 10.1086/521863.:S12-S18.
29. Cone, J. B. and W. Wetzel. 1982. Toxic megacolon secondary to pseudomembranous colitis. Dis.Colon Rectum. 25:478-482.
30. Dobson, G., C. Hickey, and J. Trinder. 2003. *Clostridium difficile* colitis causing toxic megacolon, severe sepsis and multiple organ dysfunction syndrome. Intensive Care Med. 29:1030.
31. Pepin, J., L. Valiquette, M. E. Alary, P. Villemure, A. Pelletier, K. Forget, K. Pepin, and D. Chouinard. 2004. *Clostridium difficile*-associated diarrhea in a region of Quebec from 1991 to 2003: a changing pattern of disease severity. CMAJ. 171:466-472.
32. Muto, C. A., M. Pokrywka, K. Shutt, A. B. Mendelsohn, K. Nouri, K. Posey, T. Roberts, K. Croyle, S. Krystofiak, S. Patel-Brown, A. W. Pasculle, D. L. Paterson, M. Saul, and L. H. Harrison. 2005. A large outbreak of *Clostridium difficile*-associated disease with an unexpected proportion of deaths and colectomies at a teaching hospital following increased fluoroquinolone use. Infect.Control Hosp.Epidemiol. 26:273-280.
33. Kuijper, E. J., B. Coignard, and P. Tull. 2006. Emergence of *Clostridium difficile*-associated disease in North America and Europe. Clin.Microbiol.Infect. 12 Suppl 6:2-18.
34. Bauer, M. P., D. W. Notermans, B. H. van Benthem, J. S. Brazier, M. H. Wilcox, M. Rupnik, D. L. Monnet, J. T. van Dissel, and E. J. Kuijper. 2011. *Clostridium difficile* infection in Europe: a hospital-based survey. Lancet. 377:63-73.
35. Hensgens, M. P., A. Goorhuis, D. W. Notermans, B. H. van Benthem, and E. J. Kuijper. 2009. Decrease of hypervirulent *Clostridium difficile* PCR Ribotype 027 in the Netherlands. Euro.Surveill. 14:19402.
36. McDonald, L. C., G. E. Killgore, A. Thompson, R. C. Owens, Jr., S. V. Kazakova, S. P. Sambol, S. Johnson, and D. N. Gerding. 2005. An epidemic, toxin gene-variant strain of *Clostridium difficile*. N.Engl.J.Med. 353:2433-2441.

37. Kuijper, E. J., F. Barbut, J. S. Brazier, N. Kleinkauf, T. Eckmanns, M. L. Lambert, D. Drudy, F. Fitzpatrick, C. Wiuff, D. J. Brown, J. E. Coia, H. Pituch, P. Reichert, J. Even, J. Mossong, A. F. Widmer, K. E. Olsen, F. Allerberger, D. W. Notermans, M. Delmee, B. Coignard, M. Wilcox, B. Patel, R. Frei, E. Nagy, E. Bouza, M. Marin, T. Akerlund, A. Virolainen-Julkunen, O. Lyytikainen, S. Kotila, A. Ingebretsen, B. Smyth, P. Rooney, I. R. Poxton, and D. L. Monnet. 2008. Update of *Clostridium difficile* infection due to PCR Ribotype 027 in Europe, 2008. Euro.Surveill 13.
38. Gurtler, V. and D. Grando. 2013. Reprint of New opportunities for improved ribotyping of C. difficile clinical isolates by exploring their genomes. J.Microbiol. Methods.10.
39. Walker, A. S., D. W. Eyre, D. H. Wyllie, K. E. Dingle, D. Griffiths, B. Shine, S. Oakley, L. O'Connor, J. Finney, A. Vaughan, D. W. Crook, M. H. Wilcox, and T. E. Peto. 2013. Relationship between bacterial strain type, host biomarkers, and mortality in *Clostridium difficile* infection. Clin.Infect.Dis. 56:1589-1600.
40. Drudy, D., S. Fanning, and L. Kyne. 2007. Toxin A-negative, toxin B-positive *Clostridium difficile*. Int.J.Infect.Dis. 11:5-10.
41. Goorhuis, A., M. C. Legaria, R. J. van den Berg, C. Harmanus, C. H. Klaassen, J. S. Brazier, G. Lumelsky, and E. J. Kuijper. 2009. Application of multiple-locus variable-number tandem-repeat analysis to determine clonal spread of toxin A-negative *Clostridium difficile* in a general hospital in Buenos Aires, Argentina. Clin.Microbiol. Infect. 15:1080-1086.
42. Dawson, L. F., E. Valiente, and B. W. Wren. 2009. *Clostridium difficile*--a continually evolving and problematic pathogen. Infect.Genet.Evol. 9:1410-1417.
43. Goorhuis, A., D. Bakker, J. Corver, S. B. Debast, C. Harmanus, D. W. Notermans, A. A. Bergwerff, F. W. Dekker, and E. J. Kuijper. 2008. Emergence of *Clostridium difficile* infection due to a new hypervirulent strain, polymerase chain reaction Ribotype 078. Clin.Infect.Dis. 47:1162-1170.
44. Goorhuis, A., S. B. Debast, L. A. van Leengoed, C. Harmanus, D. W. Notermans, A. A. Bergwerff, and E. J. Kuijper. 2008. *Clostridium difficile* PCR Ribotype 078: an emerging strain in humans and in pigs? J.Clin.Microbiol. 46:1157-07.
45. Dial, S., J. A. Delaney, A. N. Barkun, and S. Suissa. 2005. Use of gastric acid-suppressive agents and the risk of community-acquired *Clostridium difficile*-associated disease. JAMA. 294:2989-2995.
46. Freeman, J., M. P. Bauer, S. D. Baines, J. Corver, W. N. Fawley, B. Goorhuis, E. J. Kuijper, and M. H. Wilcox. 2010. The changing epidemiology of *Clostridium difficile* infections. Clin.Microbiol.Rev. 23:529-549.
47. Honda, H. and E. R. Dubberke. 2014. The changing epidemiology of *Clostridium difficile* infection. Curr.Opin.Gastroenterol. 30:54-62.
48. Goorhuis, A., D. Bakker, J. Corver, S. B. Debast, C. Harmanus, D. W. Notermans, A. A. Bergwerff, F. W. Dekker, and E. J. Kuijper. 2008. Emergence of *Clostridium difficile* infection due to a new hypervirulent strain, polymerase chain reaction Ribotype 078. Clin.Infect.Dis. 47:1162-1170.
49. Jhung, M. A., A. D. Thompson, G. E. Killgore, W. E. Zukowski, G. Songer, M. Warny, S. Johnson, D. N. Gerding, L. C. McDonald, and B. M. Limbago. 2008. Toxinotype V *Clostridium difficile* in humans and food animals. Emerg.Infect.Dis. 14:1039-1045.

50. Wilcox, M. H. Healt Protection Agency. *Clostridium difficile* Ribotyping Network for England and Northern Ireland: http://www.hpa.org.uk/web/HPAwebFile/HPAweb_C/1258560554236. 2009.
51. Keel, K., J. S. Brazier, K. W. Post, S. Weese, and J. G. Songer. 2007. Prevalence of PCR ribotypes among *Clostridium difficile* isolates from pigs, calves, and other species. J.Clin. Microbiol. 45:1963-1964.
52. Rupnik, M., A. Widmer, O. Zimmermann, C. Eckert, and F. Barbut. 2008. *Clostridium difficile* toxinotype V, Ribotype 078, in animals and humans. J.Clin.Microbiol. 46:1963-1964.
53. Hensgens, M. P., E. C. Keessen, M. M. Squire, T. V. Riley, M. G. Koene, B. E. de, L. J. Lipman, and E. J. Kuijper. 2012. *Clostridium difficile* infection in the community: a zoonotic disease? Clin.Microbiol.Infect. 18:635-645.
54. Keessen, E. C., W. Gaastra, and L. J. Lipman. 2011. *Clostridium difficile* infection in humans and animals, differences and similarities. Vet.Microbiol. 153:205-217.
55. Crobach, M. J., O. M. Dekkers, M. H. Wilcox, and E. J. Kuijper. 2009. European Society of Clinical Microbiology and Infectious Diseases (ESCMID): data review and recommendations for diagnosing *Clostridium difficile*-infection (CDI). Clin.Microbiol. Infect. 15:1053-1066.
56. Eastwood, K., P. Else, A. Charlett, and M. Wilcox. 2009. Comparison of nine commercially available *Clostridium difficile* toxin detection assays, a real-time PCR assay for *C. difficile tcdB*, and a glutamate dehydrogenase detection assay to cytotoxin testing and cytotoxigenic culture methods. J.Clin.Microbiol. 47:3211-3217.
57. Barbut, F., M. Delmee, J. S. Brazier, J. C. Petit, I. R. Poxton, M. Rupnik, V. Lalande, C. Schneider, P. Mastrantonio, R. Alonso, E. Kuipjer, and M. Tvede. 2003. A European survey of diagnostic methods and testing protocols for *Clostridium difficile*. Clin. Microbiol.Infect. 9:989-996. doi:683 (pii).
58. Kazanowski, M., S. Smolarek, F. Kinnarney, and Z. Grzebieniak. 2013. *Clostridium difficile*: epidemiology, diagnostic and therapeutic possibilities-a systematic review. Tech.Coloproctol.
59. Planche, T. D., K. A. Davies, P. G. Coen, J. M. Finney, I. M. Monahan, K. A. Morris, L. O'Connor, S. J. Oakley, C. F. Pope, M. W. Wren, N. P. Shetty, D. W. Crook, and M. H. Wilcox. 2013. Differences in outcome according to *Clostridium difficile* testing method: a prospective multicentre diagnostic validation study of *C. difficile* infection. Lancet Infect.Dis. 13:936-945.
60. Knetsch, C. W., D. Bakker, R. F. de Boer, I. Sanders, S. Hofs, A. M. Kooistra-Smid, J. Corver, K. Eastwood, M. H. Wilcox, and E. J. Kuijper. 2011. Comparison of real-time PCR techniques to cytotoxigenic culture methods for diagnosing *Clostridium difficile* infection. J.Clin.Microbiol. 49:227-231.
61. Stamper, P. D., R. Alcabasa, D. Aird, W. Babiker, J. Wehrlin, I. Ikpeama, and K. C. Carroll. 2009. Comparison of a commercial real-time PCR assay for *tcdB* detection to a cell culture cytotoxicity assay and toxigenic culture for direct detection of toxin-producing *Clostridium difficile* in clinical samples. J.Clin.Microbiol. 47:373-378.
62. van den Berg, R. J., E. J. Kuijper, L. E. van Coppenraet, and E. C. Claas. 2006. Rapid diagnosis of toxinogenic *Clostridium difficile* in faecal samples with internally controlled real-time PCR. Clin.Microbiol.Infect. 12:184-186.

63. Persson, S., M. Torpdahl, and K. E. Olsen. 2008. New multiplex PCR method for the detection of *Clostridium difficile* Toxin A (TcdA) and Toxin B (TcdB) and the binary toxin (*cdtA*/*cdtB*) genes applied to a Danish strain collection. Clin.Microbiol.Infect. 14:1057-1064.
64. Planche, T., A. Aghaizu, R. Holliman, P. Riley, J. Poloniecki, A. Breathnach, and S. Krishna. 2008. Diagnosis of *Clostridium difficile* infection by toxin detection kits: a systematic review. Lancet Infect.Dis. 8:777-784. doi:S1473-3099(08)70233-0 (pii);10.1016/S1473-3099(08)70233-0 (doi).
65. Sloan, L. M., B. J. Duresko, D. R. Gustafson, and J. E. Rosenblatt. 2008. Comparison of real-time PCR for detection of the *tcdC* gene with four toxin immunoassays and culture in diagnosis of *Clostridium difficile* infection. J.Clin.Microbiol. 46:1996-2001.
66. Wilcox, M. H. 2012. Overcoming barriers to effective recognition and diagnosis of *Clostridium difficile* infection. Clin.Microbiol.Infect. 18 Suppl 6:13-20. doi: 10.1111/1469-0691.12057.:13-20.
67. Knetsch, C. W., T. D. Lawley, M. P. Hensgens, J. Corver, M. W. Wilcox, and E. J. Kuijper. 2013. Current application and future perspectives of molecular typing methods to study *Clostridium difficile* infections. Euro.Surveill. 18:20381.
68. van, B. A. 2007. Tracing isolates of bacterial species by multilocus variable number of tandem repeat analysis (MLVA). FEMS Immunol.Med.Microbiol. 49:22-27.
69. Cohen, S. H., Y. J. Tang, and J. Silva, Jr. 2001. Molecular typing methods for the epidemiological identification of *Clostridium difficile* strains. Expert.Rev.Mol.Diagn. 1:61-70.
70. Killgore, G., A. Thompson, S. Johnson, J. Brazier, E. Kuijper, J. Pepin, E. H. Frost, P. Savelkoul, B. Nicholson, R. J. van den Berg, H. Kato, S. P. Sambol, W. Zukowski, C. Woods, B. Limbago, D. N. Gerding, and L. C. McDonald. 2008. Comparison of seven techniques for typing international epidemic strains of *Clostridium difficile*: restriction endonuclease analysis, pulsed-field gel electrophoresis, PCR-ribotyping, multilocus sequence typing, multilocus variable-number tandem-repeat analysis, amplified fragment length polymorphism, and surface layer protein A gene sequence typing. J.Clin.Microbiol. 46:431-437.
71. Delmee, M., M. Homel, and G. Wauters. 1985. Serogrouping of *Clostridium difficile* strains by slide agglutination. J.Clin.Microbiol. 21:323-327.
72. Toma, S., G. Lesiak, M. Magus, H. L. Lo, and M. Delmee. 1988. Serotyping of *Clostridium difficile*. J.Clin.Microbiol. 26:426-428.
73. Indra, A., S. Huhulescu, M. Schneeweis, P. Hasenberger, S. Kernbichler, A. Fiedler, G. Wewalka, F. Allerberger, and E. J. Kuijper. 2008. Characterization of *Clostridium difficile* isolates using capillary gel electrophoresis-based PCR ribotyping. J.Med.Microbiol. 57:1377-1382.
74. Stubbs, S. L., J. S. Brazier, G. L. O'Neill, and B. I. Duerden. 1999. PCR targeted to the 16S-23S rRNA gene intergenic spacer region of *Clostridium difficile* and construction of a library consisting of 116 different PCR ribotypes. J.Clin.Microbiol. 37:461-463.
75. Bidet, P., V. Lalande, B. Salauze, B. Burghoffer, V. Avesani, M. Delmee, A. Rossier, F. Barbut, and J. C. Petit. 2000. Comparison of PCR-ribotyping, arbitrarily primed PCR, and pulsed-field gel electrophoresis for typing *Clostridium difficile*. J.Clin.Microbiol. 38:2484-2487.

76. Bakker, D., J. Corver, C. Harmanus, A. Goorhuis, E. C. Keessen, W. N. Fawley, M. H. Wilcox, and E. J. Kuijper. 2010. Relatedness of human and animal *Clostridium difficile* PCR Ribotype 078 isolates determined on the basis of multilocus variable-number tandem-repeat analysis and tetracycline resistance. J.Clin.Microbiol. 48:3744-3749.
77. Sebaihia, M., B. W. Wren, P. Mullany, N. F. Fairweather, N. Minton, R. Stabler, N. R. Thomson, A. P. Roberts, A. M. Cerdeno-Tarraga, H. Wang, M. T. Holden, A. Wright, C. Churcher, M. A. Quail, S. Baker, N. Bason, K. Brooks, T. Chillingworth, A. Cronin, P. Davis, L. Dowd, A. Fraser, T. Feltwell, Z. Hance, S. Holroyd, K. Jagels, S. Moule, K. Mungall, C. Price, E. Rabbinowitsch, S. Sharp, M. Simmonds, K. Stevens, L. Unwin, S. Whithead, B. Dupuy, G. Dougan, B. Barrell, and J. Parkhill. 2006. The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome. Nat. Genet. 38:779-786.
78. Marsh, J. W., M. M. O'Leary, K. A. Shutt, A. W. Pasculle, S. Johnson, D. N. Gerding, C. A. Muto, and L. H. Harrison. 2006. Multilocus variable-number tandem-repeat analysis for investigation of *Clostridium difficile* transmission in Hospitals. J.Clin. Microbiol. 44:2558-2566.
79. van den Berg, R. J., I. Schaap, K. E. Templeton, C. H. Klaassen, and E. J. Kuijper. 2007. Typing and subtyping of *Clostridium difficile* isolates by using multiple-locus variable-number tandem-repeat analysis. J.Clin.Microbiol. 45:1024-1028.
80. Broukhanski, G., A. Simor, and D. R. Pillai. 2011. Defining criteria to interpret multilocus variable-number tandem repeat analysis to aid *Clostridium difficile* outbreak investigation. J.Med.Microbiol. 60:1095-1100.
81. He, M., M. Sebaihia, T. D. Lawley, R. A. Stabler, L. F. Dawson, M. J. Martin, K. E. Holt, H. M. Seth-Smith, M. A. Quail, R. Rance, K. Brooks, C. Churcher, D. Harris, S. D. Bentley, C. Burrows, L. Clark, C. Corton, V. Murray, G. Rose, S. Thurston, T. A. van, D. Walker, B. W. Wren, G. Dougan, and J. Parkhill. 2010. Evolutionary dynamics of *Clostridium difficile* over short and long time scales. Proc.Natl.Acad.Sci.U.S.A. %20 ;107:7527-7532.
82. He, M., F. Miyajima, P. Roberts, L. Ellison, D. J. Pickard, M. J. Martin, T. R. Connor, S. R. Harris, D. Fairley, K. B. Bamford, S. D'Arc, J. Brazier, D. Brown, J. E. Coia, G. Douce, D. Gerding, H. J. Kim, T. H. Koh, H. Kato, M. Senoh, T. Louie, S. Michell, E. Butt, S. J. Peacock, N. M. Brown, T. Riley, G. Songer, M. Wilcox, M. Pirmohamed, E. Kuijper, P. Hawkey, B. W. Wren, G. Dougan, J. Parkhill, and T. D. Lawley. 2013. Emergence and global spread of epidemic healthcare-associated *Clostridium difficile*. Nat.Genet. 45:109-113.
83. Stabler, R. A., M. He, L. Dawson, M. Martin, E. Valiente, C. Corton, T. D. Lawley, M. Sebaihia, M. A. Quail, G. Rose, D. N. Gerding, M. Gibert, M. R. Popoff, J. Parkhill, G. Dougan, and B. W. Wren. 2009. Comparative genome and phenotypic analysis of *Clostridium difficile* 027 strains provides insight into the evolution of a hypervirulent bacterium. Genome Biol. 10:R102.
84. Eyre, D. W., M. L. Cule, D. J. Wilson, D. Griffiths, A. Vaughan, L. O'Connor, C. L. Ip, T. Golubchik, E. M. Batty, J. M. Finney, D. H. Wyllie, X. Didelot, P. Piazza, R. Bowden, K. E. Dingle, R. M. Harding, D. W. Crook, M. H. Wilcox, T. E. Peto, and A. S. Walker. 2013. Diverse sources of C. difficile infection identified on whole-genome sequencing. N.Engl.J.Med. 369:1195-1205.

85. Eyre, D. W., W. N. Fawley, E. L. Best, D. Griffiths, N. E. Stoesser, D. W. Crook, T. Peto, A. S. Walker, and M. H. Wilcox. 2013. Comparison of multilocus variable number tandem repeat analysis and whole genome sequencing for investigation of *Clostridium difficile* transmission. J.Clin.Microbiol.
86. Eyre, D. W., M. L. Cule, D. Griffiths, D. W. Crook, T. E. Peto, A. S. Walker, and D. J. Wilson. 2013. Detection of mixed infection from bacterial whole genome sequence data allows assessment of its role in *Clostridium difficile* transmission. PLoS.Comput. Biol. 9:e1003059.
87. Walker, A. S., D. W. Eyre, D. H. Wyllie, K. E. Dingle, R. M. Harding, L. O'Connor, D. Griffiths, A. Vaughan, J. Finney, M. H. Wilcox, D. W. Crook, and T. E. Peto. 2012. Characterisation of *Clostridium difficile* hospital ward-based transmission using extensive epidemiological data and molecular typing. PLoS.Med. 9:e1001172.
88. Stabler, R. A., D. N. Gerding, J. G. Songer, D. Drudy, J. S. Brazier, H. T. Trinh, A. A. Witney, J. Hinds, and B. W. Wren. 2006. Comparative phylogenomics of *Clostridium difficile* reveals clade specificity and microevolution of hypervirulent strains. J.Bacteriol. 188:7297-7305.
89. Marsden, G. L., I. J. Davis, V. J. Wright, M. Sebaihia, E. J. Kuijper, and N. P. Minton. 2010. Array comparative hybridisation reveals a high degree of similarity between UK and European clinical isolates of hypervirulent *Clostridium difficile*. BMC.Genomics. 11:389. doi: 10.1186/1471-2164-11-389.:389-11.
90. Brouwer, M. S., P. J. Warburton, A. P. Roberts, P. Mullany, and E. Allan. 2011. Genetic organisation, mobility and predicted functions of genes on integrated, mobile genetic elements in sequenced strains of *Clostridium difficile*. PLoS.One. 6:e23014.
91. Adams, V., D. Lyras, K. A. Farrow, and J. I. Rood. 2002. The clostridial mobilisable transposons. Cell Mol.Life Sci. 59:2033-2043.
92. Brouwer, M. S., A. P. Roberts, P. Mullany, and E. Allan. 2012. In silico analysis of sequenced strains of *Clostridium difficile* reveals a related set of conjugative transposons carrying a variety of accessory genes. Mob.Genet.Elements. 2:8-12.
93. Roberts, A. P., M. Chandler, P. Courvalin, G. Guedon, P. Mullany, T. Pembroke, J. I. Rood, C. J. Smith, A. O. Summers, M. Tsuda, and D. E. Berg. 2008. Revised nomenclature for transposable genetic elements. Plasmid. 60:167-173.
94. Roberts, A. P. and P. Mullany. 2009. A modular master on the move: the Tn*916* family of mobile genetic elements. Trends Microbiol. 17:251-258.
95. Agerso, Y., A. G. Pedersen, and F. M. Aarestrup. 2006. Identification of Tn*5397*-like and Tn*916*-like transposons and diversity of the tetracycline resistance gene *tetM* in enterococci from humans, pigs and poultry. J.Antimicrob.Chemother. 57:832-839.
96. Corver, J., D. Bakker, M. S. Brouwer, C. Harmanus, M. P. Hensgens, A. P. Roberts, L. J. Lipman, E. J. Kuijper, and H. C. van Leeuwen. 2012. Analysis of a *Clostridium difficile* PCR Ribotype 078 100 kilobase island reveals the presence of a novel transposon, Tn*6164*. BMC.Microbiol. 12:130. doi: 10.1186/1471-2180-12-130.:130-12.
97. de Vries, L. E., H. Christensen, R. L. Skov, F. M. Aarestrup, and Y. Agerso. 2009. Diversity of the tetracycline resistance gene *tetM* and identification of Tn*916*- and Tn*5801*-like (Tn*6014*) transposons in *Staphylococcus aureus* from humans and animals. J.Antimicrob.Chemother. 64:490-500.
98. Spigaglia, P., F. Barbanti, and P. Mastrantonio. 2006. New variants of the *tetM* gene in *Clostridium difficile* clinical isolates harbouring Tn*916*-like elements. J.Antimicrob. Chemother. 57:1205-1209.

99. Mullany, P., M. Wilks, I. Lamb, C. Clayton, B. Wren, and S. Tabaqchali. 1990. Genetic analysis of a tetracycline resistance element from *Clostridium difficile* and its conjugal transfer to and from Bacillus *subtilis*. J.Gen.Microbiol. 136:1343-1349.
100. Garnier, F., S. Taourit, P. Glaser, P. Courvalin, and M. Galimand. 2000. Characterization of transposon Tn*1549*, conferring VanB-type resistance in *Enterococcus spp*. Microbiology. 146:1481-1489.
101. Tan, K. S., B. Y. Wee, and K. P. Song. 2001. Evidence for holin function of *tcdE* gene in the pathogenicity of *Clostridium difficile*. J.Med.Microbiol. 50:613-619.
102. Braun, V., T. Hundsberger, P. Leukel, M. Sauerborn, and C. von Eichel-Streiber. 1996. Definition of the single integration site of the pathogenicity locus in *Clostridium difficile*. Gene 181:29-38.
103. Govind, R., G. Vediyappan, R. D. Rolfe, B. Dupuy, and J. A. Fralick. 2009. Bacteriophage-mediated toxin gene regulation in *Clostridium difficile*. J.Virol. 83:12037-12045.
104. Sekulovic, O., M. Meessen-Pinard, and L. C. Fortier. 2011. Prophage-stimulated toxin production in *Clostridium difficile* NAP1/027 lysogens. J.Bacteriol. 193:2726-2734.
105. Kuehne, S. A., S. T. Cartman, J. T. Heap, M. L. Kelly, A. Cockayne, and N. P. Minton. 2010. The role of Toxin A and Toxin B in *Clostridium difficile* infection. Nature 467:711-713.
106. Kuehne, S. A., M. M. Collery, M. L. Kelly, S. T. Cartman, A. Cockayne, and N. P. Minton. 2013. Importance of Toxin A, Toxin B, and CDT in Virulence of an Epidemic *Clostridium difficile* Strain. J.Infect.Dis.
107. Lyras, D., J. R. O'Connor, P. M. Howarth, S. P. Sambol, G. P. Carter, T. Phumoonna, R. Poon, V. Adams, G. Vedantam, S. Johnson, D. N. Gerding, and J. I. Rood. 2009. Toxin B is essential for virulence of *Clostridium difficile*. Nature 458:1176-1179.
108. Kato, H., N. Kato, K. Watanabe, N. Iwai, H. Nakamura, T. Yamamoto, K. Suzuki, S. M. Kim, Y. Chong, and E. B. Wasito. 1998. Identification of Toxin A-negative, Toxin B-positive *Clostridium difficile* by PCR. J.Clin.Microbiol. 36:2178-2182.
109. van den Berg, R. J., E. C. Claas, D. H. Oyib, C. H. Klaassen, L. Dijkshoorn, J. S. Brazier, and E. J. Kuijper. 2004. Characterization of Toxin A-negative, Toxin B-positive *Clostridium difficile* isolates from outbreaks in different countries by amplified fragment length polymorphism and PCR ribotyping. J.Clin.Microbiol. 42:1035-1041.
110. Ho, J. G., A. Greco, M. Rupnik, and K. K. Ng. 2005. Crystal structure of receptor-binding C-terminal repeats from *Clostridium difficile* Toxin A. Proc.Natl.Acad.Sci.U.S.A. %20;102:18373-18378.
111. von Eichel-Streiber, C., P. Boquet, M. Sauerborn, and M. Thelestam. 1996. Large clostridial cytotoxins--a family of glycosyltransferases modifying small GTP-binding proteins. Trends Microbiol. 4:375-382.
112. Egerer, M., T. Giesemann, T. Jank, K. J. Satchell, and K. Aktories. 2007. Auto-catalytic cleavage of *Clostridium difficile* Toxins A and B depends on cysteine protease activity. J.Biol.Chem. 282:25314-25321.
113. Jank, T. and K. Aktories. 2008. Structure and mode of action of clostridial glucosylating toxins: the ABCD model. Trends Microbiol. 16:222-229.
114. Greco, A., J. G. Ho, S. J. Lin, M. M. Palcic, M. Rupnik, and K. K. Ng. 2006. Carbohydrate recognition by *Clostridium difficile* Toxin A. Nat.Struct.Mol.Biol. 13:460-461.
115. Giesemann, T., M. Egerer, T. Jank, and K. Aktories. 2008. Processing of *Clostridium difficile* toxins. J.Med.Microbiol. 57:690-696.

116. Stubbe, H., J. Berdoz, J. P. Kraehenbuhl, and B. Corthesy. 2000. Polymeric IgA is superior to monomeric IgA and IgG carrying the same variable domain in preventing *Clostridium difficile* Toxin A damaging of T84 monolayers. J.Immunol. 164:1952-1960.
117. Gerhard, R., E. Frenzel, S. Goy, and A. Olling. 2013. Cellular uptake of *Clostridium difficile* TcdA and truncated TcdA lacking the receptor binding domain. J.Med. Microbiol. 62:1414-1422.
118. Roth, M. G. 2006. Clathrin-mediated endocytosis before fluorescent proteins. Nat. Rev.Mol.Cell Biol. 7:63-68.
119. Papatheodorou, P., C. Zamboglou, S. Genisyuerek, G. Guttenberg, and K. Aktories. 2010. Clostridial glucosylating toxins enter cells via clathrin-mediated endocytosis. PLoS.One. 5:e10673.
120. Li, S., L. Shi, Z. Yang, and H. Feng. 2013. Cytotoxicity of *Clostridium difficile* Toxin B does not require cysteine protease-mediated autocleavage and release of the glucosyltransferase domain into the host cell cytosol. Pathog.Dis. 67:11-18.
121. Gerhard, R., H. Tatge, H. Genth, T. Thum, J. Borlak, G. Fritz, and I. Just. 2005. *Clostridium difficile* toxin A induces expression of the stress-induced early gene product RhoB. J.Biol.Chem. 280:1499-1505.
122. Rupnik, M., M. H. Wilcox, and D. N. Gerding. 2009. *Clostridium difficile* infection: new developments in epidemiology and pathogenesis. Nat.Rev.Microbiol. 7:526-536.
123. Govind, R. and B. Dupuy. 2012. Secretion of *Clostridium difficile* Toxins A and B requires the holin-like protein TcdE. PLoS.Pathog. 8:e1002727.
124. Olling, A., S. Seehase, N. P. Minton, H. Tatge, S. Schroter, S. Kohlscheen, A. Pich, I. Just, and R. Gerhard. 2012. Release of TcdA and TcdB from *Clostridium difficile* cdi 630 is not affected by functional inactivation of the *tcdE* gene. Microb.Pathog. 52:92-100.
125. Mani, N. and B. Dupuy. 2001. Regulation of toxin synthesis in *Clostridium difficile* by an alternative RNA polymerase sigma factor. Proc.Natl.Acad.Sci.U.S.A 98:5844-5849.
126. Matamouros, S., P. England, and B. Dupuy. 2007. *Clostridium difficile* toxin expression is inhibited by the novel regulator TcdC. Mol.Microbiol. 64:1274-1288.
127. Mani, N., D. Lyras, L. Barroso, P. Howarth, T. Wilkins, J. I. Rood, A. L. Sonenshein, and B. Dupuy. 2002. Environmental response and autoregulation of *Clostridium difficile* TxeR, a sigma factor for toxin gene expression. J.Bacteriol. 184:5971-5978.
128. Raffestin, S., B. Dupuy, J. C. Marvaud, and M. R. Popoff. 2005. BotR/A and TetR are alternative RNA polymerase sigma factors controlling the expression of the neurotoxin and associated protein genes in *Clostridium botulinum* type A and *Clostridium* tetani. Mol.Microbiol. 55:235-249.
129. Antunes, A., I. Martin-Verstraete, and B. Dupuy. 2011. CcpA-mediated repression of *Clostridium difficile* toxin gene expression. Mol.Microbiol. 79:882-899.
130. Dineen, S. S., A. C. Villapakkam, J. T. Nordman, and A. L. Sonenshein. 2007. Repression of *Clostridium difficile* toxin gene expression by CodY. Mol.Microbiol. 66:206-219.
131. Mackin, K. E., G. P. Carter, P. Howarth, J. I. Rood, and D. Lyras. 2013. Spo0A Differentially Regulates Toxin Production in Evolutionarily Diverse Strains of *Clostridium difficile*. PLoS.One. 8:e79666.
132. Underwood, S., S. Guan, V. Vijayasubhash, S. D. Baines, L. Graham, R. J. Lewis, M. H. Wilcox, and K. Stephenson. 2009. Characterization of the sporulation initiation pathway of *Clostridium difficile* and its role in toxin production. J.Bacteriol. 191:7296-7305.

133. Saujet, L., M. Monot, B. Dupuy, O. Soutourina, and I. Martin-Verstraete. 2011. The key sigma factor of transition phase, SigH, controls sporulation, metabolism and virulence factor expression in *Clostridium difficile*. J.Bacteriol.
134. McKee, R. W., M. R. Mangalea, E. B. Purcell, E. K. Borchardt, and R. Tamayo. 2013. The second messenger c-di-GMP regulates *Clostridium difficile* toxin production by controlling expression of sigD. J.Bacteriol.
135. Hundsberger, T., V. Braun, M. Weidmann, P. Leukel, M. Sauerborn, and C. von Eichel-Streiber. 1997. Transcription analysis of the genes *tcdA-E* of the pathogenicity locus of *Clostridium difficile*. Eur.J.Biochem. 244:735-742.
136. Dupuy, B. and A. L. Sonenshein. 1998. Regulated transcription of *Clostridium difficile* toxin genes. Mol.Microbiol. 27:107-120.
137. Merrigan, M., A. Venugopal, M. Mallozzi, B. Roxas, V. K. Viswanathan, S. Johnson, D. N. Gerding, and G. Vedantam. 2010. Human hypervirulent *Clostridium difficile* strains exhibit increased sporulation as well as robust toxin production. J.Bacteriol. 192:4904-4911.
138. Vohra, P. and I. R. Poxton. 2011. Comparison of toxin and spore production in clinically relevant strains of *Clostridium difficile*. Microbiology 157:1343-1353.
139. Govind, R., G. Vediyappan, R. D. Rolfe, and J. A. Fralick. 2006. Evidence that *Clostridium difficile* TcdC is a membrane-associated protein. J.Bacteriol. 188:3716-3720.
140. Curry, S. R., J. W. Marsh, C. A. Muto, M. M. O'Leary, A. W. Pasculle, and L. H. Harrison. 2007. *tcdC* genotypes associated with severe TcdC truncation in an epidemic clone and other strains of *Clostridium difficile*. J.Clin.Microbiol. 45:215-221.
141. Gerding, D. N., S. Johnson, M. Rupnik, and K. Aktories. 2013. binary toxin CDT: Mechanism, epidemiology, and potential clinical importance. Gut Microbes. 5.
142. Schwan, C., A. S. Kruppke, T. Nolke, L. Schumacher, F. Koch-Nolte, M. Kudryashev, H. Stahlberg, and K. Aktories. 2014. *Clostridium difficile* toxin CDT hijacks microtubule organization and reroutes vesicle traffic to increase pathogen adherence. Proc.Natl. Acad.Sci.U.S.A. 111:2313-2318.
143. Deakin, L. J., S. Clare, R. P. Fagan, L. F. Dawson, D. J. Pickard, M. R. West, B. W. Wren, N. F. Fairweather, G. Dougan, and T. D. Lawley. 2012. The *Clostridium difficile* spo0A gene is a persistence and transmission factor. Infect.Immun. 80:2704-2711.
144. Vedantam, G., A. Clark, M. Chu, R. McQuade, M. Mallozzi, and V. K. Viswanathan. 2012. *Clostridium difficile* infection: toxins and non-toxin virulence factors, and their contributions to disease establishment and host response. Gut Microbes. 3:121-134.
145. Fimlaid, K. A., J. P. Bond, K. C. Schutz, E. E. Putnam, J. M. Leung, T. D. Lawley, and A. Shen. 2013. Global analysis of the sporulation pathway of *Clostridium difficile*. PLoS. Genet. 9:e1003660.
146. Burns, D. A., J. T. Heap, and N. P. Minton. 2010. SleC is essential for germination of *Clostridium difficile* spores in nutrient-rich medium supplemented with the bile salt taurocholate. J.Bacteriol. 192:657-664.
147. Rupnik, M., M. H. Wilcox, and D. N. Gerding. 2009. *Clostridium difficile* infection: new developments in epidemiology and pathogenesis. Nat.Rev.Microbiol. 7:526-536.
148. Errington, J. 2003. Regulation of endospore formation in Bacillus subtilis. Nat.Rev. Microbiol. 1:117-126.
149. Molle, V., M. Fujita, S. T. Jensen, P. Eichenberger, J. E. Gonzalez-Pastor, J. S. Liu, and R. Losick. 2003. The Spo0A regulon of Bacillus subtilis. Mol.Microbiol. 50:1683-1701.

150. Paredes, C. J., K. V. Alsaker, and E. T. Papoutsakis. 2005. A comparative genomic view of clostridial sporulation and physiology. Nat.Rev.Microbiol. 3:969-978.
151. Pereira, F. C., L. Saujet, A. R. Tome, M. Serrano, M. Monot, E. Couture-Tosi, I. Martin-Verstraete, B. Dupuy, and A. O. Henriques. 2013. The Spore Differentiation Pathway in the Enteric Pathogen. PLoS.Genet. 9:e1003782.
152. Saujet, L., F. C. Pereira, M. Serrano, O. Soutourina, M. Monot, P. V. Shelyakin, M. S. Gelfand, B. Dupuy, A. O. Henriques, and I. Martin-Verstraete. 2013. Genome-Wide Analysis of Cell Type-Specific Gene Transcription during Spore Formation in. PLoS. Genet. 9:e1003756.
153. Heap, J. T., O. J. Pennington, S. T. Cartman, G. P. Carter, and N. P. Minton. 2007. The ClosTron: a universal gene knock-out system for the genus *Clostridium* . J.Microbiol. Methods 70:452-464.
154. Ferrari, F. A., K. Trach, D. LeCoq, J. Spence, E. Ferrari, and J. A. Hoch. 1985. Characterization of the spo0A locus and its deduced product. Proc.Natl.Acad. Sci.U.S.A. 82:2647-2651.
155. Hamon, M. A. and B. A. Lazazzera. 2001. The sporulation transcription factor Spo0A is required for biofilm development in Bacillus *subtilis*. Mol.Microbiol. 42:1199-1209.
156. Han, S. O., M. Inui, and H. Yukawa. 2007. Expression of Corynebacterium glutamicum glycolytic genes varies with carbon source and growth phase. Microbiology 153:2190-2202.
157. Agaisse, H. and D. Lereclus. 1994. Expression in Bacillus *subtilis* of the Bacillus *thuringiensis* cryIIIA toxin gene is not dependent on a sporulation-specific sigma factor and is increased in a spo0A mutant. J.Bacteriol. 176:4734-4741.
158. Baum, J. A. and T. Malvar. 1995. Regulation of insecticidal crystal protein production in Bacillus *thuringiensis*. Mol.Microbiol. 18:1-12.
159. Castilla-Llorente, V., D. Munoz-Espin, L. Villar, M. Salas, and W. J. Meijer. 2006. Spo0A, the key transcriptional regulator for entrance into sporulation, is an inhibitor of DNA replication. EMBO J. 25:3890-3899.
160. Chen, H. J., T. K. Tsai, S. C. Pan, J. S. Lin, C. L. Tseng, and G. C. Shaw. 2010. The master transcription factor Spo0A is required for poly(3-hydroxybutyrate) (PHB) accumulation and expression of genes involved in PHB biosynthesis in Bacillus *thuringiensis*. FEMS Microbiol.Lett. 304:74-81.
161. Poncet, S., E. Dervyn, A. Klier, and G. Rapoport. 1997. Spo0A represses transcription of the cry toxin genes in Bacillus *thuringiensis*. Microbiology. 143:2743-2751.
162. Saile, E. and T. M. Koehler. 2002. Control of anthrax toxin gene expression by the transition state regulator *abrB*. J.Bacteriol. 184:370-380.
163. Harris, L. M., N. E. Welker, and E. T. Papoutsakis. 2002. Northern, morphological, and fermentation analysis of spo0A inactivation and overexpression in *Clostridium acetobutylicum* ATCC 824. J.Bacteriol. 184:3586-3597.
164. Ravagnani, A., K. C. Jennert, E. Steiner, R. Grunberg, J. R. Jefferies, S. R. Wilkinson, D. I. Young, E. C. Tidswell, D. P. Brown, P. Youngman, J. G. Morris, and M. Young. 2000. Spo0A directly controls the switch from acid to solvent production in solvent-forming clostridia. Mol.Microbiol. 37:1172-1185.
165. Barketi-Klai, A., S. Hoys, S. Lambert-Bordes, A. Collignon, and I. Kansau. 2011. Role of fibronectin-binding protein A in *Clostridium difficile* intestinal colonization. J.Med. Microbiol. 60:1155-1161.

166. Janoir, C., S. Pechine, C. Grosdidier, and A. Collignon. 2007. Cwp84, a surface-associated protein of *Clostridium difficile*, is a cysteine protease with degrading activity on extracellular matrix proteins. J.Bacteriol. 189:7174-7180.
167. Tasteyre, A., M. C. Barc, A. Collignon, H. Boureau, and T. Karjalainen. 2001. Role of FliC and FliD flagellar proteins of *Clostridium difficile* in adherence and gut colonization. Infect.Immun. 69:7937-7940.
168. Janoir, C., C. Deneve, S. Bouttier, F. Barbut, S. Hoys, L. Caleechum, D. Chapeton-Montes, F. C. Pereira, A. O. Henriques, A. Collignon, M. Monot, and B. Dupuy. 2013. Adaptive Strategies and Pathogenesis of *Clostridium difficile* from In Vivo Transcriptomics. Infect.Immun. 81:3757-3769.
169. Calabi, E., F. Calabi, A. D. Phillips, and N. F. Fairweather. 2002. Binding of *Clostridium difficile* surface layer proteins to gastrointestinal tissues. Infect.Immun. 70:5770-5778.
170. Fang, F. C. 2004. Antimicrobial reactive oxygen and nitrogen species: concepts and controversies. Nat.Rev.Microbiol. 2:820-832.
171. Heinrich, J. and T. Wiegert. 2009. Regulated intramembrane proteolysis in the control of extracytoplasmic function sigma factors. Res.Microbiol. 160:696-703.
172. Ho, T. D. and C. D. Ellermeier. 2012. Extra cytoplasmic function sigma factor activation. Curr.Opin.Microbiol. 15:182-188.
173. Novick, R. P. and E. Geisinger. 2008. Quorum sensing in staphylococci. Annu.Rev. Genet. 42:541-64. doi: 10.1146/annurev.genet.42.110807.0*916*40.:541-564.
174. Sarvas, M., C. R. Harwood, S. Bron, and J. M. van Dijl. 2004. Post-translocational folding of secretory proteins in Gram-positive bacteria. Biochim.Biophys.Acta. 1694:311-327.
175. Hansen, G. and R. Hilgenfeld. 2013. Architecture and regulation of HtrA-family proteins involved in protein quality control and stress response. Cell Mol.Life Sci. 70:761-775.
176. Malet, H., F. Canellas, J. Sawa, J. Yan, K. Thalassinos, M. Ehrmann, T. Clausen, and H. R. Saibil. 2012. Newly folded substrates inside the molecular cage of the HtrA chaperone DegQ. Nat.Struct.Mol.Biol. 19:152-157.
177. Dalbey, R. E., P. Wang, and J. M. van Dijl. 2012. Membrane proteases in the bacterial protein secretion and quality control pathway. Microbiol.Mol.Biol.Rev. 76:311-330.
178. Meltzer, M., S. Hasenbein, N. Mamant, M. Merdanovic, S. Poepsel, P. Hauske, M. Kaiser, R. Huber, T. Krojer, T. Clausen, and M. Ehrmann. 2009. Structure, function and regulation of the conserved serine proteases DegP and DegS of Escherichia coli. Res.Microbiol. 160:660-666.
179. Clausen, T., M. Kaiser, R. Huber, and M. Ehrmann. 2011. HTRA proteases: regulated proteolysis in protein quality control. Nat.Rev.Mol.Cell Biol. 12:152-162.

Applied research

2

Comparison of real-time PCR techniques to cytotoxigenic culture methods for diagnosing *Clostridium difficile* infection

CW Knetsch[1,†], D Bakker[1,†], RF de Boer[2], I Sanders[1], S Hofs[2], AMD Kooistra-Smid[2], J Corver[1], K Eastwood[3], MH Wilcox[3] and EJ Kuijper[1]*

[1] *Leiden University Medical Center (LUMC), Department of Medical Microbiology, Albinusdreef 2, 2333 ZA, Leiden, the Netherlands*

[2] *Laboratory for Infectious Diseases (LvI), Department of Research & Development, van Ketwich Verschuurlaan 92, 9721 SW Groningen, the Netherlands*

[3] *Microbiology Department, Leeds Teaching Hospitals NHS Trust, Leeds, United Kingdom*

[†] *These two authors contributed equally to the study*

J Clin Microbiol 2011

Abstract

In the past decade, incidence of *Clostridium difficile* infections (CDI) with more severe course has increased in Europe and Northern America. Assays that are capable to rapidly diagnose CDI are essential. Two real-time PCRs (LUMC and LvI) targeting *C. difficile* toxin genes (*tcdB*; *tcdA* and *tcdB*, respectively), were compared with the BD GeneOhm PCR (targeting the *tcdB* gene), using cytotoxigenic culture as gold standard. In addition, a real-time PCR targeting the *tcdC* frame shift mutation at position 117 (Δ117 PCR) was evaluated for detecting toxigenic *C. difficile* and the presence of PCR Ribotype (RT) 027 in stool samples. In total, 526 diarrheal samples were prospectively collected and included in the study. Compared with cytotoxigenic culture, sensitivity, specificity, positive predicted value (PPV) and negative predicted value (NPV) were: for PCR LUMC 96.0%, 88.0%, 66.0% and 98.9%, for PCR LvI 100.0%, 89.4%, 69.7% and 100.0%, for PCR Δ117 98.0%, 90.7%, 71.9% and 99.5% and for PCR BD GeneOhm 88.3%, 96.9%, 86.5% and 97.4%. Compared to faeces samples cultured positive for *C. difficile* RT 027, sensitivity, specificity, PPV and NPV values of the Δ117 PCR were: 95.2%, 96.2%, 87.0% and 98.7%. We conclude that all real-time PCRs can be applied as a first screening test in an algorithm for diagnosing CDI. However, the low PPVs hinder the use of the assays as a stand-alone test. Furthermore, the Δ117 PCR may provide valuable information for minimising the spread of the epidemic *C. difficile* RT 027.

Introduction

Clostridium difficile is a major cause of nosocomial diarrhea and pseudomembranous colitis (1). Incidence of *Clostridium difficile* infection (CDI) has increased in the past decade, which is associated with the emergence of the hypervirulent PCR Ribotype (RT) 027 (2,3). The *C. difficile* enterotoxin A (TcdA) and cytotoxin B (TcdB) are considered as major virulence factors, whereas the binary toxin might play a role in virulence through the formation of microtubule-based protrusions thereby increasing the adherence of the bacteria (4-6). CDI can also be caused by strains that produce only TcdB (7), but strains producing TcdA only have not been described. Assays for the rapid diagnosis of CDI are important to prevent the spread of *C. difficile*, in particular for hypervirulent strains like the RT 027. Conventional diagnostic methods for CDI, such as cytotoxigenic culture (CYTGC) are time-consuming and not available at all routine diagnostic laboratories, whereas the performance of rapid enzyme immuno-assays to detect toxins in faeces is insufficient (4,5,11). Previously, we developed a real-time PCR that detects the presence of the *tcdB* gene (8). In this study we have improved the performance of this PCR and compared it with the commercially available BD GeneOhm PCR, and another in-house developed real-time PCR assay that detects the presence of the *tcdA* and *tcdB* gene (9). In addition, we evaluated a real-time PCR that targets the *tcdC* gene frame shift mutation at position 117 which can act as a marker for the RT 027/NAP1 strain (10,11).

Materials and methods

Cytotoxicity assay (CYT) and cytotoxigenic culture (CYTGC)

In total, 526 routine diagnostic diarrheal samples were submitted to the Department of Microbiology at Leeds Teaching Hospitals and tested prospectively by the CYT and CYTGC assays as previously described (12). Briefly, all stool samples (less then 48h old) were stored at 2 to 5°C. Twenty µL of diluted fecal sample (1:5 in PBS) was filtered and added to a monolayer of both *C. sordelli*-antitoxin protected (Prolab Diagnostics, United Kingdom) and unprotected Vero cells. A sample was considered toxin-positive when cell rounding was observed after 24 or 48 hours of incubation. In addition, cultured *C. difficile* isolates from faeces samples that were found negative by the CYT assay, were investigated for toxin production using the CYTGC assay (12). Isolates were inoculated into brain heart infusion broth (BHI). After 48 hours of incubation culture supernatants were added to a monolayer of protected and unprotected Vero cells. Cultured *C. difficile* isolates that were positive for the CYT assay were considered to be positive for the CYTGC assay.

Culture

Culture of isolates was performed as previously described (12). In short, following alcohol shock, samples were cultured on Braziers CCEY agar (Bioconnections, Wetherby, United Kingdom) supplemented with 5 mg/l lysozyme (Sigma, United Kingdom) and without egg yolk supplement. Incubation was done in an anaerobic workstation (Don Whitley, United Kingdom) for at least 48 hours. Grey-brown colonies with the characteristic horse manure odor were identified as *C. difficile*. Whenever the identification of an isolate was questionable the Microgen *C. difficile* latex agglutination kit (Microgen Bioproducts Ltd., Camberley, United Kingdom) was used to confirm the identity.

DNA extraction

Fecal samples were stored at 4°C for 1 week and then frozen at -20°C. Specimen preparation and DNA extraction for the BD GeneOhm Cdiff assay was performed according to the manufacturers' protocol. For the other real-time PCRs, fecal samples were pre-treated with stool transport and recovery (STAR) buffer (Roche, Penzberg, Germany) according to the manufacturers' protocol. For the LUMC real-time PCR, DNA was extracted on the MagNA Pure (Roche) using LC DNA Isolation Kit III (Roche, Penzberg, Germany) according to the manufacturers' protocol. In short, 100 µL supernatant of STAR buffer and chloroform pre-treated fecal sample was added to lysisbuffer (130 µL) and Prot K (20 µL). This mixture was heated at 65°C for 10 min followed by 95°C for 10 min, after which it was centrifuged for 1min at 1000 x g. 200 µL supernatant was used for automated DNA extraction. DNA was eluted in 100 µL elution buffer. The Phocine Herpes Virus (PhHV), which served as internal control, was added to the lysis buffer. For the LvI real-time PCR, DNA was extracted on the NucliSENS easyMAG (Biomérieux, Boxtel, The Netherlands) according to the specific A protocol (Biomérieux, Boxtel, The Netherlands). In short, 150 µL of STAR buffer and chloroform pre-treated faeces suspension was added to 2 *mL* lysis buffer (NucliSENS; Biomérieux, Boxtel, The Netherlands). After incubation for 10 min at room temperature, the total suspension was transferred to the sample vessel including 140 µL magnetic silica beads and used for automated DNA extraction. DNA was eluted in 110 µL elution buffer. The internal control Phocine Herpes Virus (PhHV) was added to the lysis buffer.

Real-time PCR

Amplification of part of the *tcdB* gene by the BD GeneOhm Cdiff PCR was performed on a Smartcycler (Cepheid, United Kingdom) according to the manufacturers' protocol. Primers and probes that were used for the LUMC real-

time PCR and the LvI PCR are described in Table 1. Amplification of the *tcdB* gene by the LUMC real-time PCR was performed on a CFX detection system (Biorad, Veenendaal, The Netherlands) as previously described (8) with some optimizations. The PCR amplification was performed in a 50 µL final volume, containing 25 µL Hotstar mastermix (Qiagen, Venlo, The Netherlands), forward and reverse primers at 80 nM each, 3.5 mM $MgCl_2$, 100 nM *tcdB* probe and 10 µL DNA. The PhHV primers described by Niesters (13) were used with a modified probe. Amplification protocol included an enzyme activation step for 15 min at 95°C, followed by 50 cycles of amplification; 94°C for 30 sec, 50°C for 30 sec and 72°C for 30 sec. The LvI real-time PCR was designed to target both *tcdA* and *tcdB* gene. Amplification of these genes was performed as a multiplex PCR on an AB 7500 PCR system (Applied Biosystems, Nieuwerkerk a/d IJssel, The Netherlands).

Table 1: Primers and probes used for in-house developed real-time PCRs.

Real-time PCR	Target	Primers and probes	Nucleotide sequence (5'→ 3')
LUMC	TcdB	398CLDs	GAAAGTCCAAGTTTACGCTCAAT
		399CLDas	GCTGCACCTAAACTTACACCA
		551CLD-tq-FAM	FAM-ACAGATGCAGCCAAAGTTGTTGAATT-BHQ1
LvI	TcdA	CD-*tcdA*-F	TTG TAT GGA TAG GTG GAG AAG TCA GT
		CD-*tcdA*-R	AAT ATT ATA TTC TGC ATT AAT ATC AGC CCA T
		CD-*tcdA*-MGB1	FAM-ATA TTG CTC TTG AAT ACA TAA A-NFQ-MGB
		CD-*tcdA*-MGB2	FAM-TAT TGT TCT TGA ATA CAT AAA AC-NFQ-MGB
	TcdB	CD-*tcdB*-F1	GAA ACA GGA TGG ACA CCA GGT T
		CD-*tcdB*-F2	AAG AGG ATG GAC GCC AGG TT
		CD-*tcdB*-R1	ACG GTC TAA CAG TTT TGT GCC A
		CD-*tcdB*-R2	CTG CCC TTC ATA ATG ATC TCT TAT ACG
		CD-*tcdB*-MGB	FAM-AAG AAG CTT AGA AAA TG-NFQ-MGB
Δ117 PCR	TcdC	CD-*tcdC*-F	GCA CAA AGG RTA TTG CTC TAC TGG
		CD-*tcdC*-R1	AGC TGG TGA GGA TAT ATT GCC AA
		CD-*tcdC*-R2	CAA GAT GGT GAG GAT ATA TTG CCA
		CD-*tcdC*wt-MGB	FAM-AAA CAC RCC HAA AAT AA-NFQ-MGB
		CD-*tcdC*mut-MGB	VIC-AAA CAC RCC AAA ATA A-NFQ-MGB
All	PhHV	295PhHVs	GGGCGAATCACAGATTGAATC
		296PhHVas	GCGGTTCCAAACGTACCAA
		531PhHV-tq-CY5	CY5-TTTTTATGTGTCCGCCACCATCTGGATC-BHQ2* NED-CGC CAC CAT CTG GAT-NFQ-MGB **

BHQ = Black hole quencher; NFQ = Non fluorescent quencher; MGB = Minor groove binder. *used for LUMC PCR; ** used for LVI PCR and Δ117 PCR

Each PCR reaction was performed in a 25 µL final volume, containing 1x TaqMan Universal PCR Master Mix (Applied Biosystems, Nieuwerkerk a/d IJssel, The Netherlands), forward and reverse primers at 300 nM each, TaqMan MGB probes at 100 nM each, 2.5 µg bovine serum albumin (Roche), and 5 µL DNA extract. For PhHV, the primers described by Niesters (13) were used, whereas the probe was modified to a MGB-probe. Amplification protocol included 2 min at 50°C, 10 min at 95°C followed by 40 cycles of amplification; 94°C for 15 sec, 60°C for 1 min.

The Δ117 real-time PCR was designed to target the *tcdC* gene frame shift mutation at position 117. This assay utilizes two Taqman MGB probes, a wildtype (WT) probe and a mutant (MUT Δ117) probe respectively, that both can hybridize with part of the *tcdC* gene sequence flanking the 1 bp deletion at position 117. Isolates that do not carry the 1 bp deletion, will give a stronger signal with the WT probe, while 027/NAP1 isolates will do so with the MUT Δ117 probe. Hence, the ΔCt (Ct WT – Ct MUT Δ117) for 027/NAP1 strains will be positive, whereas the ΔCt for other *C. difficile* ribotypes will be negative, which enables discrimination. The primer/probe set was used in the same multiplex setup as the LvI PCR assay described above, with the primers at 300 nM and the MGB probes (5'-FAM and 5'-VIC, 3'-NFQMGB) (Applied Biosystems) at 100 nM. Reactions were run on an ABI 7500 with the same amplification protocol as for the LvI PCR.

PCR ribotyping

PCR ribotyping was performed at the Department of Microbiology at Leeds Teaching Hospitals following the protocol from the *C. difficile* Ribotyping Network for England (CDRNE) laboratory (12).

Data analysis

Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) were determined by comparing the real-time PCRs with the CYTGC gold standard; using statistical software PASW 17.0.2. Differences in the sensitivity and specificity between the real-time PCRs were determined by using the McNemar's Test for correlated proportions. Samples with a Ct-value higher then 40 were considered as negative. In addition, samples with an internal control Ct-value that deviated more than 3 Ct-values compared to the internal control Ct-value of the negative control were considered inhibited and discarded from the analyses for the LUMC PCR. For the LVI PCR, samples were considered inhibited and discarded from analysis when the Ct value for the internal control exceeded 34.91 cycles (i.e., the mean Ct value for uninhibited specimens ± 2 standard deviations). The number of inhibited samples for the Δ117 PCR was equally determined as for the LVI PCR. Furthermore, the Δ117 PCR was compared with PCR ribotyping of CYTGC

positive isolates. The BD GeneOhm Cdiff PCR tests were interpreted according to the manufacturers' protocol. The software on the Smart cycler (Cepheid, UK) recorded the results of the PCR assay as positive, negative or unresolved.

Results

Comparing the real-time PCRs with the gold standard CYTGC

In total, we evaluated 526 diarrheal samples, of which 101 samples (19.2%) were positive in the CYTGC assay. Of 101 positive samples, 13 were derived from CYT-assay negative samples. For the evaluation of the LUMC PCR 10 samples (1.9%) were excluded from the analysis due to inhibition during the amplification step, whereas 16 samples (3.1%) and 15 samples (2.9%) were inhibited and excluded from the analysis of the LvI PCR and Δ117 PCR, respectively. Five samples (1%) were recorded as unresolved by the BD GeneOhm PCR assay and excluded from the analysis. Sensitivity, specificity, PPV and NPV for all PCR methods against CYTGC are shown in Table 2. All stool samples that tested positive for *C. difficile* by the CYTGC assay were also detected by the LvI PCR. Comparable sensitivity was achieved by the LUMC PCR (96%) and Δ117 PCR (98%), while the sensitivity of the commercial BD GeneOhm PCR was lower (88.3%). The difference in sensitivity between the BD GeneOhm PCR and the three in-house developed PCRs was significant with *p-values* of 0.00, 0.01 and 0.04 for the LvI PCR, the Δ117 PCR and the LUMC PCR, respectively. In contrast, the BD GeneOhm PCR showed higher specificity (96.9%) compared to the LvI PCR (89.4%), the Δ117 PCR(90.7%) and the LUMC PCR (88.0%). The difference in specificity between the BD GeneOhm PCR and all three in house developed PCRs was significant with a p-value off 0.00. Compared to CYTGC, all PCRs had similar NPV ranging from 97.4 to 100%. The BD GeneOhm PCR had the highest PPV (86.5%) compared to the LvI PCR (69.7%), the Δ117 PCR (71.9%) and the LUMC PCR (66.0%).

Discrepancy analysis

Analysis of false positive results showed an overlap of the numbers of false positives detected by the different PCR methods (Figure 1A). In total, 13 false positive results (14% of total amount of positives) were found by the BD GeneOhm PCR compared to CYTGC. Of these false positives, 54% (n = 7) was also detected as such by all other PCR methods. Compared to CYTGC, the most false positives (34%) were detected by the LUMC PCR, whereas the LvI PCR and the Δ117 PCR had30% and 28% false positive samples, respectively.

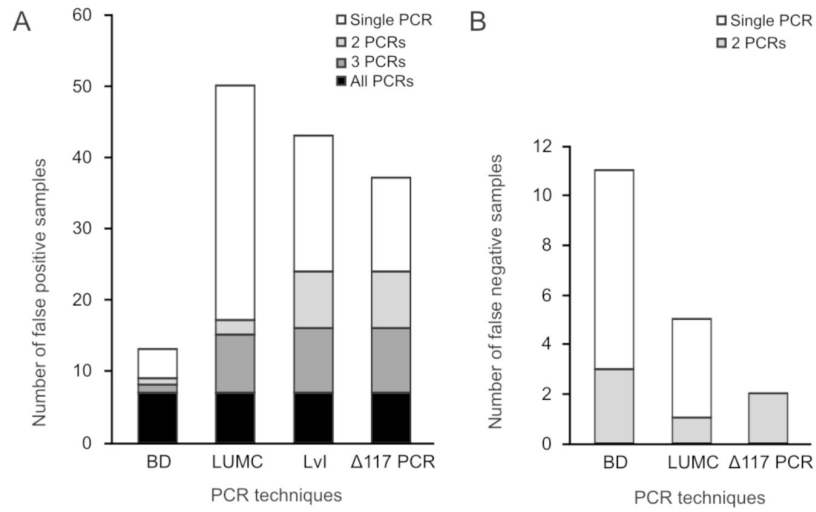


Figure 1: (A) False positive results and **(B)** false negative results detected by real-time PCRs with overlapping samples**.** All false positive and false negative results from each PCR method compared to the CYTGC assay were analyzed for resemblances. Resemblances in false positive and false negative results were ordered by PCR method. No false negative results were found by the LvI real-time PCR.

Table 2. Comparison of four real-time PCR methods with gold standard CYTGC.

Assay	Samples included	Samples inhibited	Result	CYTGC Assay (N)		Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	PPV (%)	NPV (%)
				+	-				
LUMC PCR	526	10	+	97	50	96.0	88.0	66.0	98.9
			-	4	365	(90.3-98.5)	(84.5-90.7)		
LvI PCR	522[b]	16	+	99	43	100.0	89.4	69.7	100.0
			-	0	364	(96.3-100)	(86.1-92.1)		
Δ117 PCR[a]	522[b]	15	+	97	38	98.0	90.7	71.9	99.5
			-	2	370	(92.9-99.4)	(87.5-93.1)		
BD GeneOhm	512[c]	N/A	+	83	13	88.3	96.9	86.5	97.4
			-	11	405	(80.3-93.3)	(94.8-98.2)		

For each PCR method the number of samples included and inhibited is shown. Sensitivity, specificity, negative predictive value (NPV) and positive predictive value (PPV) of the PCR methods are given as percentages, and the corresponding 95% confidence interval (95% CI) is shown in parentheses.[a]. The lowest Ct value belonging to either the *tcdC* wildtype or mutant probe was used for the evaluation of the Δ117 PCR as a screening assay. [b] 4 samples were not present in the collection. [c] data for 14 samples were not available.

Fourteen percent of the false positives detected by the LUMC PCR were also detected as such by all other PCR methods, whereas for the LvI PCR and the Δ117 PCR 16% and 19% of the false positives had a similar test outcome by all other PCR methods. Figure 1B shows the number of false negative results detected by the real-time PCRs. No false negative results were found by the LvI real-time PCR with the CYTGC as standard. Compared to CYTGC, 11 false negative results (2.6% of the total amount of negatives) were detected by the BD GeneOhm PCR, whereas the LUMC PCR and the Δ117 PCR had 1.1% and 0.5% false negative samples, respectively. None of the false negative samples were detected as such by all three PCR methods; only overlapping results between two PCR methods were found.

Comparing the Δ117 PCR with PCR ribotyped CYTGC positive samples

Of the 99 CYTGC positive samples, a total of 21 samples were typed as RT 027 by PCR-Ribotyping (Table 3). The Δ117 PCR was able to confirm 20 of these samples (95%) by detection of the 1 bp deletion at position 117 in the *tcdC* gene, with a ΔCt (Ct WT – Ct MUT Δ117) = + 2.9 cycles difference on average. Compared with CYTGC positive, RT 027 samples, sensitivity, specificity, PPV and NPV values were for this assay: 95.2%, 96.2%, 87.0% and 98.7%. The Δ117 PCR detected 3 samples carrying the Δ117 mutation, which were ribotyped as RT 005, RT 106 and an unknown RT, not RT 027.

Table 3: Comparison of the LvI Δ117 PCR with PCR ribotyping.

Assay	Samples included[a]	Result	Ribotyping of cultured isolates		Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
			Ribotype 027	Ribotype non-027				
Δ117 PCR	99	+	20	3[b]	95.2	96.2	87.0	98.7
		-	1	73				

All CYTGC positive samples (N=99) were analyzed by the Δ117 PCR. Sensitivity, specificity, negative predictive value (NPV) and positive predictive value (PPV) of the PCR method are given as percentages.[a] Only CYTGC positive samples were included.[b] Three CYTGC positive samples were typed as PCR Ribotype 005, Ribotype 106, and rare Ribotype (not 027).

Discussion

Rapid and accurate diagnosis of CDI is essential for patient management and prevention of nosocomial transmission. A main issue in diagnosing CDI is that most conventional tests do not have sufficient performance for applying it as a standalone test. Molecular tests are increasingly applied for diagnosing CDI and are also mentioned in a recent published guideline as potentially rapid assays with better performances (4). This study was performed to compare the diagnostic value of three in-house developed real-time PCRs and a commercially available BD GeneOhm Cdiff assay, using the appropriate gold standard on 526 prospectively collected stool samples. The sensitivity of the in-house developed PCRs was better than the BD GeneOhm test, in contrast to the specificity. Subsequently, NPVs were similar whereas the PPV was the highest for the BD GeneOhm test (86.5%). Peterson *et al.* (14) evaluated a real-time PCR that targeted the *tcdB* gene of *C. difficile* and reported a sensitivity of 93.3% and a specificity of 97.4% This is in line with what has been reported by Sloan *et al.* (15) on the performance of a real-time PCR, which was directed against the *tcdC* gene. They reported a sensitivity of 86% and specificity of 97%. The sensitivity and specificity reported by both studies are comparable to what has been found in this study for the BD GeneOhm Cdiff PCR.

Our study had a prevalence of toxigenic *C. difficile* positive samples of approximately 20%, repeated samples from positive patients were excluded. The prevalence was high, due to a selection of faeces samples with high suspicion of CDI. The PPV is dependent on the prevalence of the disease in the tested population. Several studies report that in the hospital 5%-10% of the antibiotic-associated diarrhea samples contain *C. difficile* (16-18). In the community the prevalence of CDI is close to 2% (19,20). Therefore, we calculated PPVs for all PCR methods at prevalence's of 2%, 5% and 10%. The calculated PPVs at 10% and 5% prevalence decreased to 76% and 60% for the BD GeneOhm PCR, 51% and 33% for the LvI PCR, 54% and 36% for the Δ117 PCR and 47% and 30% for the LUMC PCR. At 2% prevalence PPVs are 37% (BD GeneOhm), 16% (LvI), 18% (Δ117) and 14% (LUMC).

The performance of the LUMC real-time PCR reported in this evaluation was better than reported previously by van den Berg *et al.* (8). This difference can be explained by the difference in prevalence of CDI positive samples used in this study (20%) and the previous study (6%). Furthermore, we optimized PCR conditions (PCR-mix, modified probe), and reached a detection limit of 10^3 CFU/gram stool samples (data not shown) which is improved compared to 10^5 CFU/gram as reported by van den Berg *et al.* (14).

When the false positive PCR results were analyzed, 54% of the false positives detected by the BD GeneOhm PCR were also detected as such by the other PCRs. This suggests that 54% of the false positives contained *C. difficile* specific *tcdB* DNA, since the samples were detected by three PCRs targeting the *tcdB* gene using different primer sets. In addition, these samples were also found positive by the Δ117 PCR using the *tcdC* gene as a target. It can not be excluded that the cultures were false negative due to previous antibiotical treatment but discrepancies with CYTGC assay still remain present. We consider this finding as an important lack of the currently available gold standard and think that future clinical studies are necessary to interpret the findings more precisely.

In comparison to CYTGC positive samples with *C. difficile* RT 027, the Δ117 PCR targeting the *tcdC* gene Δ117 1 bp deletion has a high concordance of 95.2%. This high concordance makes the utility of the Δ117 PCR for direct detection of the epidemic strain promising, although further research is needed to determine if other *C. difficile* RTs contain the *tcdC* point mutation at position 117 and, consequently, are detected by this PCR. Furthermore, the performance of this PCR for detection of toxigenic *C. difficile* indicates that this assay has the potential to diagnose CDI, without pre-screening for the toxin genes A and B. The *tcdC* gene has been recognized as a putative negative regulator of *tcdA* and *tcdB* and thereby is indicative of the presence of the pathogenicity locus (21).

A difference between the three in-house developed real-time PCRs was the percentage of inhibited samples. In total, 3.1% (n = 16), and 2.9% (n=15) of all samples (n = 522) tested by the LvI PCR and Δ117 PCR were inhibited respectively, whereas 1.9% (n = 10) of all samples (526) tested by the LUMC PCR were inhibited. The PCRs used different DNA extraction methods and different platforms which might contribute to the observed differences.

Most rapid diagnostic tests do not have sufficient performance for applying it as a standalone test. Recently, Planche *et al.* (17) defined that a test is applicable as a standalone test when a sensitivity of at least 90% and a specificity of at least 97% is reached. The three in-house developed PCRs lack specificity, whereas the BD GeneOhm PCR lacks sensitivity resulting in too low PPVs of all PCRs ranging from 66% to 86.5% at 20% CDI-prevalence. These PPVs decrease substantially when calculating PPVs for CDI-prevalence's that are more common for a clinical setting (10%) or observed in the community (2%). None of our evaluated real-time PCR methods fulfilled the criteria defined by Planche *et al.* (17). Therefore, it was concluded that they cannot be applied as a standalone test. This finding is in line with what has been found for toxin detection assays and other molecular based assays by other studies (12,16,17). However, due to their high NPVs all four evaluated PCR methods can be applied as a first negative screening test for CDI

in a two-step algorithm. In this algorithm the PCR assay is followed by a second confirmation step to confirm the first positive test result.

Reference List

1. Bartlett, J. G. 2002. Clinical practice. Antibiotic-associated diarrhea. N.Engl.J.Med. 346:334-339.
2. McDonald, L. C., G. E. Killgore, A. Thompson, R. C. Owens, Jr., S. V. Kazakova, S. P. Sambol, S. Johnson, and D. N. Gerding. 2005. An epidemic, toxin gene-variant strain of *Clostridium difficile*. N.Engl.J.Med. 353:2433-2441.
3. Kuijper, E. J., B. Coignard, and P. Tull. 2006. Emergence of *Clostridium difficile*-associated disease in North America and Europe. Clin.Microbiol.Infect. 12 Suppl 6:2-18.
4. Lyras, D., J. R. O'Connor, P. M. Howarth, S. P. Sambol, G. P. Carter, T. Phumoonna, R. Poon, V. Adams, G. Vedantam, S. Johnson, D. N. Gerding, and J. I. Rood. 2009. Toxin B is essential for virulence of *Clostridium difficile*. Nature 458:1176-1179.
5. Voth, D. E. and J. D. Ballard. 2005. *Clostridium difficile* toxins: mechanism of action and role in disease. Clin.Microbiol.Rev. 18:247-263.
6. Schwan, C., B. Stecher, T. Tzivelekidis, H. M. van, M. Rohde, W. D. Hardt, J. Wehland, and K. Aktories. 2009. *Clostridium difficile* toxin CDT induces formation of microtubule-based protrusions and increases adherence of bacteria. PLoS.Pathog. 5:e1000626.
7. van den Berg, R. J., E. C. Claas, D. H. Oyib, C. H. Klaassen, L. Dijkshoorn, J. S. Brazier, and E. J. Kuijper. 2004. Characterization of toxin A-negative, toxin B-positive *Clostridium difficile* isolates from outbreaks in different countries by amplified fragment length polymorphism and PCR ribotyping. J.Clin.Microbiol. 42:1035-1041.
8. van den Berg, R. J., N. Vaessen, H. P. Endtz, T. Schulin, d. van, V, and E. J. Kuijper. 2007. Evaluation of real-time PCR and conventional diagnostic methods for the detection of *Clostridium difficile*-associated diarrhoea in a prospective multicentre study. J.Med.Microbiol. 56:36-42.
9. de Boer, R. F., J. J. Wijma, T. Schuurman, J. Moedt, B. G. Dijk-Alberts, A. Ott, A. M. Kooistra-Smid, and Y. T. van Duynhoven. 2010. Evaluation of a rapid molecular screening approach for the detection of toxigenic *Clostridium difficile* in general and subsequent identification of the *tcdC* Delta117 mutation in human stools. J.Microbiol. Methods. doi:S0167-7012(10)00248-4 (pii);10.1016/j.mimet.2010.07.017 (doi).
10. MacCannell, D. R., T. J. Louie, D. B. Gregson, M. Laverdiere, A. C. Labbe, F. Laing, and S. Henwick. 2006. Molecular analysis of *Clostridium difficile* PCR Ribotype 027 isolates from Eastern and Western Canada. J.Clin.Microbiol. 44:2147-2152.
11. Curry, S. R., J. W. Marsh, C. A. Muto, M. M. O'Leary, A. W. Pasculle, and L. H. Harrison. 2007. *tcdC* genotypes associated with severe TcdC truncation in an epidemic clone and other strains of *Clostridium difficile*. J.Clin.Microbiol. 45:215-221.
12. Eastwood, K., P. Else, A. Charlett, and M. Wilcox. 2009. Comparison of nine commercially available *Clostridium difficile* toxin detection assays, a real-time PCR assay for *C. difficile tcdB,* and a glutamate dehydrogenase detection assay to cytotoxin testing and cytotoxigenic culture methods. J.Clin.Microbiol. 47:3211-3217.
13. Niesters, H. G. 2002. Clinical virology in real time. J.Clin.Virol. 25 Suppl 3:S3-12.
14. Peterson, L. R., R. U. Manson, S. M. Paule, D. M. Hacek, A. Robicsek, R. B. Thomson, Jr., and K. L. Kaul. 2007. Detection of toxigenic *Clostridium difficile* in stool samples by real-time polymerase chain reaction for the diagnosis of C. difficile-associated diarrhea. Clin.Infect.Dis. 45:1152-1160.

15. Sloan, L. M., B. J. Duresko, D. R. Gustafson, and J. E. Rosenblatt. 2008. Comparison of real-time PCR for detection of the *tcdC* gene with four toxin immunoassays and culture in diagnosis of *Clostridium difficile* infection. J.Clin.Microbiol. 46:1996-2001.
16. Crobach, M. J., O. M. Dekkers, M. H. Wilcox, and E. J. Kuijper. 2009. European Society of Clinical Microbiology and Infectious Diseases (ESCMID): data review and recommendations for diagnosing *Clostridium difficile*-infection (CDI). Clin.Microbiol. Infect. 15:1053-1066.
17. Planche, T., A. Aghaizu, R. Holliman, P. Riley, J. Poloniecki, A. Breathnach, and S. Krishna. 2008. Diagnosis of *Clostridium difficile* infection by toxin detection kits: a systematic review. Lancet Infect.Dis. 8:777-784.
18. Barbut, F., M. Delmee, J. S. Brazier, J. C. Petit, I. R. Poxton, M. Rupnik, V. Lalande, C. Schneider, P. Mastrantonio, R. Alonso, E. Kuipjer, and M. Tvede. 2003. A European survey of diagnostic methods and testing protocols for *Clostridium difficile*. Clin. Microbiol.Infect. 9:989-996.
19. Bauer, M. P., D. Veenendaal, L. Verhoef, P. Bloembergen, J. T. van Dissel, and E. J. Kuijper. 2009. Clinical and microbiological characteristics of community-onset *Clostridium difficile* infection in The Netherlands. Clin.Microbiol.Infect. 15:1087-1092.
20. Wilcox, M. H., L. Mooney, R. Bendall, C. D. Settle, and W. N. Fawley. 2008. A case-control study of community-associated *Clostridium difficile* infection. J.Antimicrob. Chemother. 62:388-396.
21. Spigaglia, P. and P. Mastrantonio. 2002. Molecular analysis of the pathogenicity locus and polymorphism in the putative negative regulator of toxin production (TcdC) among *Clostridium difficile* clinical isolates. J.Clin.Microbiol. 40:3470-3475.

3

Relatedness of human and animal *Clostridium difficile* PCR Ribotype 078 isolate determined on the basis of Multilocus Variable-Number tandem repeat Analysis and tetracycline resistance

D. Bakker[1], J. Corver[1], C. Harmanus[1], A. Goorhuis[1], E.C. Keessen[3], W.N. Fawley[2], M.H. Wilcox[2] and E.J. Kuijper[1]

[1] *Department of Medical Microbiology, Center of Infectious Diseases, Leiden University Medical Center, the Netherlands.*
[2] *Department of Microbiology, Leeds Teaching Hospitals & University of Leeds, Leeds, UK.*
[3] *Institute for Risk Assessment Sciences, division of VPH University of Utrecht, The Netherlands.*

J. Clin. Microbiol. 2010

Abstract

One hundred and two human and 56 porcine *Clostridium difficile* PCR Ribotype (RT) 078 strains from four European countries were investigated by an optimized Multiple Locus Variable number tandem repeat Analysis (MLVA) and for tetracycline susceptibility. Eighty-five percent of all isolates were genetically related, irrespective of human or porcine origin. Human strains were significantly more resistant to tetracycline than porcine strains. All tetracycline resistant strains contained the Tn*916*-like transposon, harboring the *tet*(M) gene. We conclude that strains from human and porcine origin are genetically related, irrespective of country of origin. Further studies are needed to clarify the genetic relatedness of *C. difficile* RT 078 strains, including whether this is a consequence of less natural variability in this ribotype versus other ribotypes.

Introduction

Recently, we reported that *Clostridium difficile* PCR Ribotype (RT) 078 is an increasing cause of *Clostridium difficile* Infections (CDI) in humans in the Netherlands with similar disease severity as the hypervirulent RT 027 (1). Also the incidence of CDI in England caused by RT 078 has increased (2). In addition, recent studies have demonstrated that RT 078 is the predominant type in cattle and pigs (3,4). In the Netherlands, we have noticed an overlap in the occurrence of human CDI cases caused by RT 078 and the distribution of pig farms in the eastern part of the country (1). This suggests a possible link between human and porcine RT 078 strains.

To investigate the relatedness between human and porcine RT 078 strains, we applied a Multiple locus variable number tandem repeat analysis (MLVA) developed for *C. difficile* on a collection of RT 078 isolates (1,5,6). This MLVA has been proven to be more discriminatory than other genotyping methods (7,8). Since it has been suggested that the wide dissemination of *Staphylococcus aureus* ST398 in pigs and humans is associated with the frequent usage of tetracycline in pig farms, we also investigated the susceptibility to tetracycline and the genetic origin of tetracycline resistance (9-11).

Clostridium difficile strains

A total of 102 human and 56 porcine RT 078 strains were available for this study. Table 1 depicts the location of isolation and the year of isolation of each strain. All human RT 078 strains were recovered from diarrhoeal patients. The "Leeds collection" (n=67) consisted of 44 strains originating from an outbreak in Northern Ireland, 20 strains from other parts of the UK and 3 strains originating from Ireland. The "Leiden collection" consisted of 34 strains of endemic cases in the year 2006-2007. The 56 porcine strains were collected from 11 Dutch pig farms in the years 2006, 2007 and 2009. All pig farms had persistent problems of neonatal diarrhoea. Forty-seven (84%) isolates were recovered from diarrhoeal piglets.

Modification of MLVA

MLVA was adjusted for RT 078 due to the lack of specific PCR products for loci $A6_{Cd}$, $B7_{Cd}$, $C6_{Cd}$ and $G8_{Cd}$. Sequence analysis of 7 human and 8 porcine RT 078 strains revealed multiple mismatches in the primer annealing sites for loci $B7_{Cd}$, $C6_{Cd}$ and $G8_{Cd}$ and the absence of locus $A6_{cd}$. We adjusted the magnesium chloride concentration (4 mM) and annealing temperatures for loci $B7_{Cd}$, $G8_{Cd}$ (47°C) and locus $C6_{Cd}$ (46°C). All MLVA PCRs were performed in a singleplex format. MLVA

PCRs for the other loci and the analysis of the MLVA data were performed with the previously described conditions (5). The calculated Variable Number of Tandem Repeats (VNTR) of the adjusted MLVA was in complete concordance with the manually measured VNTR. The absence of locus $A6_{Cd}$ could theoretically result in less discriminative power of the MLVA. Therefore, we reanalyzed the MLVA on previously typed RT 027 (n=57) and 017 (n=71) strains (5,6). This reanalysis based on 6 loci resulted in equal numbers of Genetically related Clusters (GCs) and Clonal Complexes (CCs). Subsequently, we concluded that the optimized MLVA for RT 078, based on 6 loci is capable to discriminate between strains from various countries and origins.

Tetracycline susceptibility

All strains were tested for their susceptibility to tetracycline. The breakpoint for tetracycline was defined as MIC ≥ 8 mg/l (12). Seventy-five of the 102 human and 15 of the 56 porcine strains were resistant to tetracycline. Tetracycline resistance was not found among randomly selected isolates of the 2 most common human types 001 (n=10) and 014 (n=10). There were significantly more human strains resistant to tetracycline compared to porcine strains ($p<0.005$ measured by Chi-squared method). This difference could be explained by the fact that we included 44 outbreak strains. Thirty-eight (86%) of the outbreaks strains were resistant to tetracycline versus 15 (65%) resistant strains originating from other parts of the UK and Ireland.

The origin of tetracycline resistance was investigated by detection of mobile elements Tn*5397*-like and Tn*916*-like transposons as previously described (13). All tetracycline resistant strains contained the Tn*916*-like transposon, harboring the *tet*(M) gene. Filtermating experiments demonstrated the transfer of the Tn*916*-like transposon from a donor strain to a recipient strain. We did not detect either of the transposons in tetracycline susceptible strains. This observation suggests a high degree of relatedness of human and porcine isolates and is in agreement with recently published findings (13). Although, we cannot exclude horizontal transfer of the Tn*916*-like transposon since this transposon is widely distributed in Gram-positive bacteria and additional data are required from tetracycline resistant non-078 RTs (14-16). Recent publications show that tetracycline resistance is predominantly present in RTs 012, 017, 046 and 078 (12,17,18). A screening of randomly selected strains of these types demonstrated that tetracycline resistance in RT 012 and 046 is encoded by the Tn*5397*-like transposon, whereas the origin of tetracycline resistance in RT 017 and 078 is encoded by the Tn*916*-like transposon.

Application of optimized MLVA

Table 1 depicts the results of the MLVA of each strain per locus. A minimal spanning tree (MST) was constructed to determine the genetic relationships among strains as previously described (Figure 1) (1). In total, 116 strains belonged to one of the GCs, defined as a Summed Tandem Repeat difference (STRD) ≤ 10. The largest GC (green cluster) contained 103 strains, encompassing 47 porcine strains, 41 Leeds collection strains and 15 Leiden collection strains. Fifty-five strains were

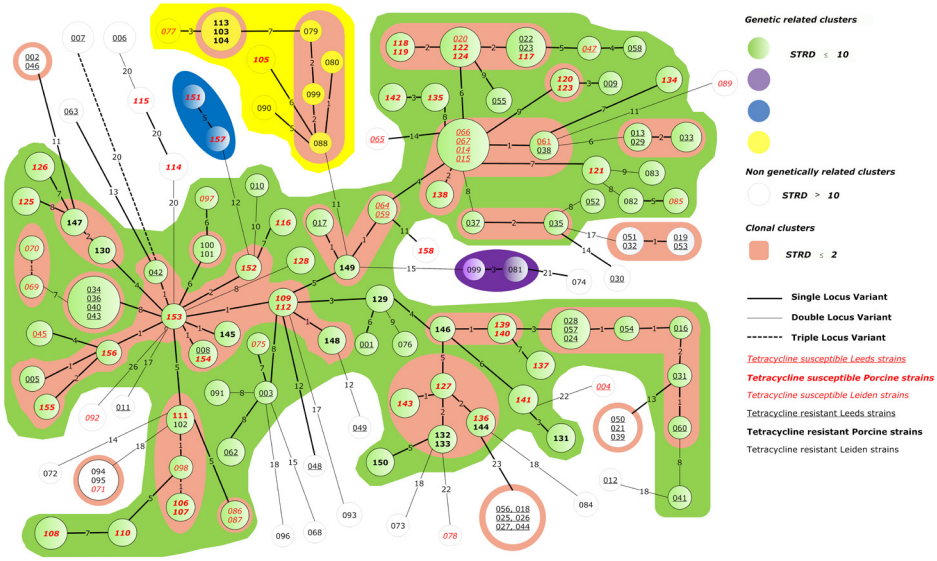


Figure 1: Minimum spanning tree analysis of 158 *C. difficile* RT 078 isolates by MLVA: Each circle represents either a unique isolate or more isolates that have identical MLVA types. The numbers between the circles represent the summed tandem repeat difference (STRD) between MLVA Types. Thick lines represent single locus variants, thin lines represent double locus variants and the interrupted lines represent triple locus variance between MLVA types. Clonal clusters are defined by a STRD ≤ 2 and genetically related clusters are defined by a STRD ≤ 10. Porcine (n=56) isolates susceptible are printed in ***red italic bold font***, human isolates susceptible to tetracycline from the Leeds collection are printed in *red italic underlined font* and human isolates from the Leiden collection susceptible to tetracycline are printed in *red italic normal font*. Isolates resistant to tetracycline from the Leeds collection are printed in black underlined font. Porcine isolates resistant to tetracycline are printed in **bold font** and human isolates from the Leiden collection are printed in normal font. The numbers represent isolates from different geographical locations: The Leeds collection: Northern Ireland 1-44, Yorkshire and Humber 45-53, South West England 54-57, North West England 58-62, East England 63-64, Ireland 65-67. The Leiden collection: Noord Holland 68-76, Zuid Holland 77-79, Utrecht 80-87, Gelderland 88-91, Brabant 92-93, Groningen 94-95, Friesland 96, Flevoland 97, Limburg 98, Overijssel 99-101 and Belgium 102. Porcine strains 103-158.

susceptible and 48 strains were resistant to tetracycline. The yellow GC contained only strains which originated in the Netherlands, encompassing both tetracycline susceptible and resistant strains. The blue GC contained only porcine strains susceptible to tetracycline which originated from one pig farm. The last GC (purple) contained only tetracycline resistant strains from the Leiden collection. Sixteen of the 23 recognized CCs (defined as a STRD ≤2) belonged to the largest GC, whereas the other 7 CCs were either single or double locus variants of the largest GC. Five CCs contained only porcine strains and 13 CCs contained only human strains, of which 8 CCs derived from the Leeds collection and 5 CCs derived from the Leiden collection. The remaining 5 CCs contained both porcine and human strains, irrespective of origin of country. Two CCs contained human strains isolated in different locations from a specific region. Nine of the 23 CCs contained outbreak strains and strains from distinct settings. Interestingly, 12 CCs contained only strains resistant to tetracycline and 8 CCs contained both tetracycline susceptible and resistant strains. The remaining 3 CCs contained only tetracycline susceptible strains. In total, 4 MLVA profiles could be recognized that contained both human and porcine strains. Overall, the MST could not differentiate between geographical origin or tetracycline phenotype, irrespective of human or porcine origin.

Table 1: Depicts of each strain the results of the optimized MLVA for each of the 7 loci ($A6_{Cd}$, $B7_{Cd}$, $C6_{Cd}$, $G8_{Cd}$ $E7_{Cd}$, $F3_{Cd}$ and $H9_{Cd}$).

MST no.	location of isolation (anonymised)	year of isolation	$A6_{Cd}$	$B7_{Cd}$	$C6_{Cd}$	$E7_{Cd}$	$F3_{Cd}$	$G8_{Cd}$	$H9_{Cd}$
001	A	2008	n/a	19	28	8	4	4	2
002	B	2008	n/a	18	26	8	4	4	2
003	E	2008	n/a	24	34	8	4	4	2
004	E	2008	n/a	13	18	6	4	4	2
005	B	2008	n/a	15	33	8	4	3	2
006	H	2008	n/a	26	27	12	4	4	2
007	GP	2008	n/a	17	49	7	4	1	2
008	I	2008	n/a	16	32	8	4	4	2
009	J	2008	n/a	22	30	8	4	6	2
010	A	2008	n/a	16	26	5	4	6	2
011	M	2008	n/a	25	25	8	4	4	2
012	T	2008	n/a	18	29	5	4	0	2
013	U	2008	n/a	21	39	5	4	6	2
014	GP	2008	n/a	21	38	8	4	3	2
015	W	2008	n/a	21	38	8	4	3	2

MST no.	location of isolation (anonymised)	year of isolation	$A6_{Cd}$	$B7_{Cd}$	$C6_{Cd}$	$E7_{Cd}$	$F3_{Cd}$	$G8_{Cd}$	$H9_{Cd}$
016	X	not supplied	n/a	19	39	5	4	4	2
017	GP	2008	n/a	21	34	8	4	5	2
018	M	2008	n/a	37	36	8	4	4	2
019	AA	not supplied	n/a	6	42	6	4	5	2
020	T	not supplied	n/a	21	32	8	4	3	2
021	E	2008	n/a	19	24	5	4	4	2
022	GP	2008	n/a	19	32	8	4	3	2
023	AD	2008	n/a	19	32	8	4	3	2
024	E	2008	n/a	19	38	5	4	5	2
025	E	2008	n/a	37	36	8	4	4	2
026	GP	2008	n/a	37	36	8	4	4	2
027	GP	2008	n/a	37	36	8	4	4	2
028	M	2008	n/a	19	38	5	4	5	2
029	AA	not supplied	n/a	21	39	5	4	6	2
030	AH	2008	n/a	35	43	5	4	5	2
031	M	2008	n/a	19	37	5	4	4	2
032	M	2008	n/a	6	43	6	4	5	2
033	GP	2008	n/a	21	41	5	4	6	2
034	AI	2008	n/a	24	33	8	4	4	2
035	AJ	2008	n/a	21	43	5	4	5	2
036	E	2008	n/a	24	33	8	4	4	2
037	AJ	2008	n/a	21	43	5	4	3	2
038	AI	2008	n/a	21	39	8	4	3	2
039	AK	2008	n/a	19	24	5	4	4	2
040	AI	2008	n/a	24	33	8	4	4	2
041	AK	2008	n/a	20	45	5	4	0	2
042	M	2008	n/a	17	33	8	4	4	2
043	M	2008	n/a	24	33	8	4	4	2
044	M	2008	n/a	37	36	8	4	4	2
045	K	2008	n/a	15	23	8	4	5	2
046	D	not supplied	n/a	18	26	8	4	4	2
047	G	2008	n/a	19	32	8	4	8	2
048	G	2008	n/a	28	34	8	4	4	2
049	C	2007	n/a	20	33	8	4	3	2
050	G	2008	n/a	19	24	5	4	4	2
051	AB	2007	n/a	6	43	6	4	5	2
052	AC	2007	n/a	15	43	7	4	5	2

MST no.	location of isolation (anonymised)	year of isolation	$A6_{Cd}$	$B7_{Cd}$	$C6_{Cd}$	$E7_{Cd}$	$F3_{Cd}$	$G8_{Cd}$	$H9_{Cd}$
053	D	2007	n/a	6	42	6	4	5	2
054	F	2008	n/a	19	39	5	4	5	2
055	Y	2007	n/a	29	32	8	4	0	2
056	Y	2007	n/a	37	36	8	4	4	2
057	AF	2008	n/a	19	38	5	4	5	2
058	L	2007	n/a	15	32	8	4	8	2
059	V	2007	n/a	21	34	8	4	3	2
060	Z	2007	n/a	20	37	5	4	4	2
061	AE	1953	n/a	21	38	8	4	3	2
062	AL	not supplied	n/a	24	26	8	4	4	2
063	O	2007	n/a	17	43	5	4	4	2
064	S	2008	n/a	21	34	8	4	3	2
065	P	not supplied	n/a	35	38	8	4	3	2
066	Q	2008	n/a	21	38	8	4	3	2
067	R	2008	n/a	21	38	8	4	3	2
068	AL	2006	n/a	26	38	12	2	4	8
069	AM	2007	n/a	26	33	10	2	4	9
070	AM	2007	n/a	26	33	11	2	4	9
071	AL	2007	n/a	29	40	9	2	4	8
072	AN	2007	n/a	24	41	9	2	4	8
073	AO	2007	n/a	25	35	13	2	4	8
074	AP	2007	n/a	20	20	12	2	4	8
075	AZ	2006	n/a	26	34	11	2	4	8
076	AZ	2006	n/a	21	34	12	2	4	9
077	AV	2007	n/a	21	37	12	2	4	8
078	BB	2007	n/a	31	35	11	2	4	8
080	AQ	2006	n/a	23	30	9	2	4	8
081	BE	2006	n/a	20	41	12	2	4	8
082	BF	2006	n/a	22	31	10	2	4	8
083	BG	2006	n/a	24	31	11	2	4	8
084	BE	2007	n/a	31	36	7	2	4	8
085	BG	2006	n/a	27	31	10	2	4	8
086	BG	2006	n/a	23	33	9	2	4	8
087	BG	2006	n/a	23	33	9	2	4	8
088	AR	2006	n/a	23	30	11	2	4	8
089	AU	2007	n/a	25	39	12	2	4	8
090	BA	2006	n/a	23	25	11	2	4	8

MST no.	location of isolation (anonymised)	year of isolation	$A6_{Cd}$	$B7_{Cd}$	$C6_{Cd}$	$E7_{Cd}$	$F3_{Cd}$	$G8_{Cd}$	$H9_{Cd}$
091	BJ	2006	n/a	26	32	10	2	4	8
092	AT	2007	n/a	18	52	11	2	4	8
093	BI	2006	n/a	27	34	12	2	4	8
094	AY	2007	n/a	29	40	9	2	4	8
095	AX	2007	n/a	29	40	9	2	4	8
096	AW	2007	n/a	26	43	13	2	4	8
097	BC	2006	n/a	19	28	10	2	4	8
098	BD	2007	n/a	18	32	9	2	4	8
099	BP	2006	n/a	23	41	12	2	4	8
100	BK	2007	n/a	18	33	10	2	4	8
101	BK	2007	n/a	18	33	10	2	4	8
102	AS	2005	n/a	18	33	9	2	4	8
103	BL	2006	n/a	21	37	9	2	4	8
104	BL	2006	n/a	21	37	9	2	4	8
105	BM	2007	n/a	23	36	11	2	4	8
106	BN	2007	n/a	19	32	9	2	4	8
107	BN	2007	n/a	19	32	9	2	4	8
108	BN	2007	n/a	13	25	9	2	4	8
109	BN	2007	n/a	18	34	4	2	4	8
110	BN	2007	n/a	13	32	9	2	4	8
111	BN	2007	n/a	18	33	9	2	4	8
112	BN	2007	n/a	18	34	4	2	4	8
113	BO	2007	n/a	21	37	9	2	4	8
114	BR	2009	n/a	8	43	8	4	4	2
115	BR	2009	n/a	28	43	8	4	4	2
116	BS	2009	n/a	25	33	8	4	6	2
117	BS	2009	n/a	21	32	8	4	3	2
118	BS	2009	n/a	25	32	8	4	3	2
119	BS	2009	n/a	25	32	8	4	3	2
121	BS	2009	n/a	23	31	8	4	3	2
122	BS	2009	n/a	23	32	8	4	3	2
123	BS	2009	n/a	24	30	8	4	3	2
124	BS	2009	n/a	23	32	8	4	3	2
125	BT	2009	n/a	20	45	8	4	4	2
126	BT	2009	n/a	20	44	8	4	4	2
127	BT	2009	n/a	16	38	8	4	4	2
128	BT	2009	n/a	18	40	8	4	5	2

MST no.	location of isolation (anonymised)	year of isolation	$A6_{Cd}$	$B7_{Cd}$	$C6_{Cd}$	$E7_{Cd}$	$F3_{Cd}$	$G8_{Cd}$	$H9_{Cd}$
129	BU	2009	n/a	21	34	8	4	4	2
130	BU	2009	n/a	18	37	8	4	4	2
131	BU	2009	n/a	15	35	8	4	4	2
132	BU	2009	n/a	16	35	8	4	4	2
133	BU	2009	n/a	16	35	8	4	4	2
134	BV	2009	n/a	16	39	8	4	3	2
135	BV	2009	n/a	15	38	8	4	3	2
136	BV	2009	n/a	16	36	8	4	4	2
137	BV	2009	n/a	28	38	8	4	5	2
138	BV	2009	n/a	21	38	8	4	3	2
139	BV	2009	n/a	21	38	8	4	5	2
140	BV	2009	n/a	21	38	8	4	5	2
141	BV	2009	n/a	15	38	8	4	4	2
142	BV	2009	n/a	15	35	8	4	3	2
143	BV	2009	n/a	16	38	8	4	5	2
144	BW	2009	n/a	16	36	8	4	4	2
145	BW	2009	n/a	17	33	8	4	4	2
146	BW	2009	n/a	21	38	8	4	4	2
147	BW	2009	n/a	20	37	8	4	4	2
148	BW	2009	n/a	17	34	8	4	4	2
149	BW	2009	n/a	23	34	8	4	4	2
150	BW	2009	n/a	11	35	8	4	4	2
151	BX	2009	n/a	8	35	8	4	1	2
152	BX	2009	n/a	18	33	8	4	6	2
153	BX	2009	n/a	18	33	8	4	4	2
154	BX	2009	n/a	18	32	8	4	4	2
155	BX	2009	n/a	20	33	8	4	3	2
156	BX	2009	n/a	18	33	8	4	3	2
157	BX	2009	n/a	8	35	8	4	6	2
158	BX	2009	n/a	12	34	8	4	3	2

All strains have a MST number which corresponds to the number used in the MST (Figure 1); geographical location: Northern Ireland 1-44, Yorkshire and Humber 45-53, South West England 54-57, North West England 58-62, East England 63-64, Ireland 65-67. Noord Holland 68-76, Zuid Holland 77-79, Utrecht 80-87, Gelderland 88-91, Brabant 92-93, Groningen 94-95, Friesland 96, Flevoland 97, Limburg 98, Overijssel 99-101 and Belgium 102. Porcine strains 103-158. For each strain an anonymised location (institute or pigfarm) of isolation, year of isolation and the MLVA results for each locus is described.

Conclusions

The suggested high relatedness between human and porcine RT 078 strains is in concordance with earlier publications based on MLVA, whole-genome analysis and Multi-Locus Sequence typing (19-22). The relatedness between human and porcine RT 078 strains in this study could be an indication of a common source as suggested in previous publications (1,19,20). However, the geographical locations of some related isolates are very distinct and can not logically be explained by any direct epidemiological link. A possible common source of RT 078 could be further supported by the observation that all tetracycline resistant strains contain the mobile element Tn*916*-like transposon which has also been described for tetracycline resistant enterococci from human and porcine origin (13). Interspecies transmission or transmission through meat is suggested as sources of infection but these are not yet established (23-25). However, data of a direct epidemiological link between human and porcine strains in this study is lacking. We also need to consider that the high relatedness between human and porcine RT 078 strains could be the consequence of less natural variability in RT 078 than in other types. A limitation of this study is the inclusion of porcine strains from only one country, while human strains derived from 4 European countries. Further studies are needed to investigate the possible transmission routes between humans and animals.

Reference List

1. Goorhuis, A., D. Bakker, J. Corver, S. B. Debast, C. Harmanus, D. W. Notermans, A. A. Bergwerff, F. W. Dekker, and E. J. Kuijper. 2008. Emergence of *Clostridium difficile* infection due to a new hypervirulent strain, polymerase chain reaction Ribotype 078. Clin.Infect.Dis. 47:1162-1170.
2. Wilcox, M. H. Healt Protection Agency. *Clostridium difficile* Ribotyping Network for England and Northern Ireland: http://www.hpa.org.uk/web/HPAwebFile/HPAweb_C/1258560554236. 2009.
3. Keel, K., J. S. Brazier, K. W. Post, S. Weese, and J. G. Songer. 2007. Prevalence of PCR ribotypes among *Clostridium difficile* isolates from pigs, calves, and other species. J.Clin. Microbiol. 45:1963-1964.
4. Rupnik, M., A. Widmer, O. Zimmermann, C. Eckert, and F. Barbut. 2008. *Clostridium difficile* toxinotype V, Ribotype 078, in animals and humans. J.Clin.Microbiol. 46:1963-1964.
5. van den Berg, R. J., I. Schaap, K. E. Templeton, C. H. Klaassen, and E. J. Kuijper. 2007. Typing and subtyping of *Clostridium difficile* isolates by using multiple-locus variable-number tandem-repeat analysis. J.Clin.Microbiol. 45:1024-1028.
6. Goorhuis, A., M. C. Legaria, R. J. van den Berg, C. Harmanus, C. H. Klaassen, J. S. Brazier, G. Lumelsky, and E. J. Kuijper. 2009. Application of multiple-locus variable-number tandem-repeat analysis to determine clonal spread of toxin A-negative *Clostridium difficile* in a general hospital in Buenos Aires, Argentina. Clin.Microbiol. Infect. 15:1080-1086.
7. Killgore, G., A. Thompson, S. Johnson, J. Brazier, E. Kuijper, J. Pepin, E. H. Frost, P. Savelkoul, B. Nicholson, R. J. van den Berg, H. Kato, S. P. Sambol, W. Zukowski, C. Woods, B. Limbago, D. N. Gerding, and L. C. McDonald. 2008. Comparison of seven techniques for typing international epidemic strains of *Clostridium difficile*: restriction endonuclease analysis, pulsed-field gel electrophoresis, PCR-ribotyping, multilocus sequence typing, multilocus variable-number tandem-repeat analysis, amplified fragment length polymorphism, and surface layer protein A gene sequence typing. J.Clin.Microbiol. 46:431-437.
8. Fawley, W. N., J. Freeman, C. Smith, C. Harmanus, R. J. van den Berg, E. J. Kuijper, and M. H. Wilcox. 2008. Use of highly discriminatory fingerprinting to analyze clusters of *Clostridium difficile* infection cases due to epidemic Ribotype 027 strains. J.Clin.Microbiol. 46:954-960.
9. de Neeling, A. J., M. J. van den Broek, E. C. Spalburg, M. G. van Santen-Verheuvel, W. D. Dam-Deisz, H. C. Boshuizen, A. W. van de Giessen, D. E. van, and X. W. Huijsdens. 2007. High prevalence of methicillin resistant *Staphylococcus aureus* in pigs. Vet.Microbiol. 122:366-372.
10. Wulf, M., N. A. van, A. Eikelenboom-Boskamp, V. J. de, W. Melchers, C. Klaassen, and A. Voss. 2006. Methicillin-resistant Staphylococcus aureus in veterinary doctors and students, the Netherlands. Emerg.Infect.Dis. 12:1939-1941.
11. Mevius D.J., Wit B., and van Pelt W. Monitoring of Antimicrobial Resistance and Antibiotic Usage in Animals in the Netherlands in 2006/2007 available at: http://www.cvi.wur.nl/NR/rdonlyres/DDA15856-1179-4CAB-BAC628C4728ACA03/83791/MARAN_2007_def2.pdf. 2007.

12. Barbut, F., P. Mastrantonio, M. Delmee, J. Brazier, E. Kuijper, and I. Poxton. 2007. Prospective study of *Clostridium difficile* infections in Europe with phenotypic and genotypic characterisation of the isolates. Clin.Microbiol.Infect. 13:1048-1057.
13. Agerso, Y., A. G. Pedersen, and F. M. Aarestrup. 2006. Identification of Tn*5397*-like and Tn*916*-like transposons and diversity of the tetracycline resistance gene tet(M) in enterococci from humans, pigs and poultry. J.Antimicrob.Chemother. 57:832-839.
14. Roberts, M. C. 2005. Update on acquired tetracycline resistance genes. FEMS Microbiol. Lett. 245:195-203.
15. Storrs, M. J., C. Poyart-Salmeron, P. Trieu-Cuot, and P. Courvalin. 1991. Conjugative transposition of Tn*916* requires the excisive and integrative activities of the transposon-encoded integrase. J.Bacteriol. 173:4347-4352.
16. Rice, L. B. 1998. Tn*916* family conjugative transposons and dissemination of antimicrobial resistance determinants. Antimicrob.Agents Chemother. 42:1871-1877.
17. Huang, H., H. Fang, A. Weintraub, and C. E. Nord. 2009. Distinct ribotypes and rates of antimicrobial drug resistance in *Clostridium difficile* from Shanghai and Stockholm. Clin.Microbiol.Infect. 15:1170-1173.
18. Noren, T., I. Alriksson, T. Akerlund, L. G. Burman, and M. Unemo. 2009. In vitro susceptibility to 17 antimicrobials among clinical *Clostridium difficile* isolates collected 1. Clin.Microbiol.Infect.
19. Debast, S. B., L. A. van Leengoed, A. Goorhuis, C. Harmanus, E. J. Kuijper, and A. A. Bergwerff. 2009. *Clostridium difficile* PCR Ribotype 078 toxinotype V found in diarrhoeal pigs identical to isolates from affected humans. Environ.Microbiol. 11:505-511.
20. Stabler, R. A., D. N. Gerding, J. G. Songer, D. Drudy, J. S. Brazier, H. T. Trinh, A. A. Witney, J. Hinds, and B. W. Wren. 2006. Comparative phylogenomics of *Clostridium difficile* reveals clade specificity and microevolution of hypervirulent strains. J.Bacteriol. 188:7297-7305.
21. Griffiths, D., W. Fawley, M. Kachrimanidou, R. Bowden, D. W. Crook, R. Fung, T. Golubchik, R. M. Harding, K. J. Jeffery, K. A. Jolley, R. Kirton, T. E. Peto, G. Rees, N. Stoesser, A. Vaughan, A. S. Walker, B. C. Young, M. Wilcox, and K. E. Dingle. 2009. Multilocus Sequence Typing of *Clostridium difficile*. J.Clin.Microbiol.
22. Marsh, J. W., M. M. O'Leary, K. A. Shutt, S. P. Sambol, S. Johnson, D. N. Gerding, and L. H. Harrison. 2009. Multilocus variable number tandem repeat analysis and multilocus sequence typing reveal genetic relationships among *Clostridium difficile* isolates genotyped by restriction endonuclease analysis. J.Clin.Microbiol.
23. Jhung, M. A., A. D. Thompson, G. E. Killgore, W. E. Zukowski, G. Songer, M. Warny, S. Johnson, D. N. Gerding, L. C. McDonald, and B. M. Limbago. 2008. Toxinotype V *Clostridium difficile* in humans and food animals. Emerg.Infect.Dis. 14:1039-1045.
24. Songer, J. G., H. T. Trinh, G. E. Killgore, A. D. Thompson, L. C. McDonald, and B. M. Limbago. 2009. *Clostridium difficile* in retail meat products, USA, 2007. Emerg.Infect.Dis. 15:819-821.
25. Weese, J. S., B. P. Avery, J. Rousseau, and R. J. Reid-Smith. 2009. Detection and enumeration of *Clostridium difficile* spores in retail beef and pork. Appl.Environ. Microbiol. 75:5009-5011.

4

Analysis of a *Clostridium difficile* PCR Ribotype 078 100 kilobase island reveals the presence of a novel transposon, Tn*6164*

J. Corver[1], D. Bakker[1], M. Brouwer[2], C. Harmanus[1], M. Hengens[1], A. Roberts[2], L. Lipman[3],E. Kuijper[1] and H. Van Leeuwen[1].

[1] *Department of Medical Microbiology, Section Experimental Microbiology, Center of Infectious Diseases, Leiden University Medical Center, Leiden, The Netherlands.*

[2] *Division of Microbial Diseases, UCL Eastman Dental Institute, University College London, London, UK*

[3] *Division of Veterinary Public Health, Faculty of Veterinary Medicine, Institute of Risk Assessment Sciences.*

BMC Microbiol. 2012

Abstract

Clostridium difficile is the main cause of antibiotic associated diarrhea. In the past decade, the number of *C. difficile* patients has increased dramatically, coinciding with the emergence of two PCR Ribotypes, 027 and 078. PCR Ribotype 078 is also frequently found during *C. difficile* outbreaks in pigfarms. Previously, the genome of the PCR Ribotype 078 strain M120, a human isolate, was described to contain a unique insert of 100 kilobases.

Analysis of this insert revealed over 90 open reading frames, encoding proteins originating from transposons, phages and plasmids. The insert was shown to be a transposon (Tn*6164*), as evidenced by the presence of an excised and circularised molecule, containing the ligated 5'and 3'ends of the insert. Transfer of the element could not be shown through filter-mating experiments. Whole genome sequencing of PCR Ribotype 078 strain 31618, isolated from a diarrheic piglet, showed that Tn*6164* was not present in this strain. To test the prevalence of Tn*6164*, a collection of 231 *Clostridium difficile* PCR Ribotype 078 isolates from human (n = 173) and porcine (n = 58) origin was tested for the presence of this element by PCR. The transposon was present in 9 human, tetracycline resistant isolates, originating from various countries in Europe, and none of the pig strains. Nine other strains, also tetracycline resistant human isolates, contained half of the transposon, suggesting multiple insertion steps yielding the full Tn*6164*. Other PCR ribotypes (n = 66) were all negative for the presence of the transposon. Multi locus variable tandem repeat analysis revealed genetic relatedness among transposon containing isolates. Although the element contained several potential antibiotic resistance genes, it did not yield a readily distinguishable phenotype.

Tn*6164* is a newly described transposon, occurring sporadically in *C. difficile* PCR Ribotype 078 strains. Although no transfer of the element could be shown, we hypothesize that the element could serve as a reservoir of antibiotic resistance genes for other bacteria. Further research is needed to investigate the transfer capabilities of the element and to substantiate the possible role of Tn*6164* as a source of antibiotic resistance genes for other gut pathogens.

Introduction

Over the past decade, *Clostridium difficile* has emerged as an important gut pathogen, causing hospital- and community-acquired diarrhea. The number of patients and the severity of disease have increased dramatically, due to the emergence of two hypervirulent PCR ribotypes (RT), 027 (1) and 078 (2,3). Traditionally, RT 027 has been linked to nosocomial outbreaks. In contrast, RT 078 has been detected frequently in farming animals, especially pigs (2,4), and is found more during community acquired infection. The increase in *C. difficile* infections (CDI) of humans has boosted interest in *C. difficile* biology, diagnostics and pathogenesis.

In the past few years, multiple genome sequences of several RTs have been determined (5-8). The analyses of the genomes, aided by comparative genomics of DNA-DNA microarrays (9,10) has shown that the genomes of *C. difficile* are highly variable with inserts of mobile DNA from phage, plasmid or transposon origin. These mobile DNA elements are actively moving within *C. difficile* genomes and are frequently passed on to neighboring bacteria, harboring mosaic genomes (7,11). It is unclear what role the mobile elements play in the virulence of *C. difficile*. Some virulence linked genes, for example the holin-like *tcdE*, have a phage origin (12). In fact, it has been suggested that the whole Pathogenicity Locus (PaLoc), encoding the major *C. difficile* virulence factors TcdA and TcdB, is of phage origin (13,14). Recently, phages have been shown to upregulate toxin production in *C. difficile*, thereby increasing the virulence (15). *C. difficile* transposons have been shown to contain antibiotic resistance genes (5,7,16,17), and therefore acquiring such an element could increase the virulence and/or colonization potential of a particular strain.

Mobile elements play an important role in the diversification of bacterial genomes. One important group of mobile genetic elements is the Tn*916* family of conjugative transposons (also known as integrative and conjugative elements (ICEs)) (18). These conjugative transposons usually code for tetracycline resistance and are found primarily in the *Firmicutes*. Numerous transposons have been described to be present in *C. difficile* genomes (5,7,11,17,19). Several elements closely related to Tn*916* are present in diverse *C. difficile* strains, including Tn*5397* which confers tetracycline resistance (20,21). Other transposons have been described to confer resistance to chlooramphenicol and erythromycin (5).

Recently, the first full length genome of a RT 078 strain was published (5). This M120 strain has been isolated from an Irish diarrheic patient. It was shown that RT 078 is highly divergent from RTs 027, 001, 017 and 012. In addition, this RT 078 strain was described to contain a unique 100 kb insert that showed 80% similarity to sequences of *Thermoanaerobacter* species and *Streptococcus pneumoniae* (5). In this

paper we show that the 100 kb insert is a mobile element that is only sporadically present in RT 078 strains. Furthermore, we show that the 100 kb consists of at least two independent mobile elements that were fused during evolution.

Material and Methods

Bacterial isolates and culture conditions

RT 078 *C. difficile* strain 31618 was obtained from a pig farm in the eastern part of the Netherlands where neonatal diarrhea was present. Culturing of the feces yielded *C. difficile*, as determined by an in-house PCR for the presence of the *gluD* gene encoding the glutamate dehydrogenase specific for *C. difficile* (39). PCR Ribotype was determined as previously described (40).

The other RT 078 strains used in this study were obtained from a previously described RT 078 strain collection (16), consisting of strains isolated from humans and pigs, supplemented with human RT 078 strains from the European *Clostridium difficile* Infection Survey (ECDIS) study in 2010 (32). In addition, recently isolated RT 078 strains from Dutch diarrheic piglets (2007–2010) and human (2006–2010) strains collected by the Dutch *C. difficile* Reference Laboratory (CDRL) were used. The 58 Pig strains were collected on 27 pig farms in the Netherlands.

RT 126 strains used in this study originate from the ECDIS study, isolated in 2010, from several countries in Europe (32). PCR Ribotype reference strains (n = 68) were obtained from the CDRL. The nontoxinogenic strain CD37 (41,42) was used as a recipient in filter mating experiments as this has previously been shown to be a good recipient for mobile genetic elements from other *C. difficile* strains (11). *C. difficile* strain M120 was kindly provided by Dr. Trevor Lawley (Sanger Institute). Standard culturing of *C. difficile* isolates was carried out on blood agar plates at 37 °C and anaerobic conditions.

DNA Sequencing, reference assembly and annotation

DNA was isolated from one colony of the 31618 strain by standard techniques (43). The isolate was sequenced using the Illumina platform (Solexa) at the Leiden Genome Technology Center (LGTC) at the LUMC, using the manufacturers' protocols. Single end reads were generated and submitted to the NCBI sequence read archive (http://www.ncbi.nlm.nih.gov/sra) under accession number SRX030155. A reference assembly of the reads was carried out against strain *C. difficile* RT 078 strain M120 (GenBank accession no. FN665653), using CLC genomics workbench (CLCbio, Aarhus, Denmark). Number of reads used was 5267302, of which 2968638 reads could be mapped to the M120 genome sequence. The unique 100 kb

insert present in M120 was readily identified with the CLC genomics workbench. The ORFs present in the insert were identified by CLC genomics workbench and annotation was carried out manually, using BLAST and SMART. ORFs identified as "protein of unknown function" were further analyzed by profile-profile searches through HHpred (http://toolkit. tuebingen.mpg.de/hhpred).

Bioinformatic comparison of the mixed origin of Tn*6164*

The genome of strain M120 was compared to the genomes of *C. difficile* 630 (Genbank accession no. AM180355), *Thermoanaerobacter* sp. (GenBank accession no. CP002210), *S. pneumonia* (Genbank accession no. CP002121) and *C. fetus* (Genbank accession no. FN594949) using the Artemis Comparison Tool (44).

Circularization of the transposon

In order to investigate if the putative element could excise itself from the genome, PCR analysis was performed to amplify the joint region of a circular molecule using primers at the ends of the element, facing outward (primers 14 and 15 in Table 1). PCR amplifications were carried out using the NEB *Taq* Polymerase kit (New England Biolabs, Herts, UK) according to the manufacturer's instructions with 10 mM dNTPs (NEB). The primers that were used are listed in Table 1 (Sigma-Genosys, UK).

Filter-matings assays

Filter-matings were carried out as described previously (45). *C. difficile* strains M120 and CD37 were cultured on Brain heart infusion (BHI) (Oxoid Ltd.) agar supplemented with 5% Horse blood (E&O laboratories). *C. difficile* strain CD37 was used as recipient. Transconjugants were selected for on BHI plates supplemented with 25 μg/mL rifampicin (Sigma Aldrich) and 10 μg/mL tetracycline (Sigma Aldrich). Transconjugants were examined using PCR with primer pair Lok1/Lok3 to confirm identity of the recipient strain and primer pairs Tn*6164* accessory region Fw + Rev and Tn*916* Fw + Rev to confirm the transfer of Tn*6164* or Tn*6190*.

Inverse PCR

C. difficile genomic DNA was digested with PstI or EcoRI. After purification, the genomic DNA fragments were self-ligated to create circular DNAs. Subsequently, the DNA was precipitated and dissolved in H_2O. PCR was carried out on the DNA, using primers 4-rev and 5-rev or 14 and 15 (annealing at 58°C, 35 cycles). PCR products were visualized by gel electrophoresis and sequences were determined through direct sequencing on the purified PCR amplicons or through cloning into pCR2.1/TOPO (Invitrogen) and subsequent sequencing with the plasmid-located

primers T7 and M13reverse. Oligonucleotides used in this study are shown in Table 1.

PCR

PCRs were carried out using Gotaq polymerase (Promega, Leiden, the Netherlands). Reactions contained 0.4 mM dNTPs, 0.4 uM oligonucleotides. Annealing temperature of the PCR was set at 50°C and PCRs were standardized at 30 cycles.

Antibiotic resitance

The MIC for tetracycline was determined using E tests (BioMérieux, Boxtel, the Netherlands) on blood plates under anaerobic conditions at 37°C. Breakpoint for

Table 1: Oligonucleotides used in this study.

Name	Sequence (5'- 3')	Purpose
1	GAGATATGGTTATGAGATTAGG	Presence/absence of insert
2	CCCACCTTTATAGCATCATATAG	Absence of insert
3	CTAACCTATCAACTCAACCCC	Presence of insert
4	AGGATAAGACCGCAGCAGAA	Presence 5'half of insert
5	AAAAACGACGGTTTTCCTGTG	Presence 5'half of insert
6	GGGCAAATAGAAAGTCAAAACG	Presence 3'half of insert
7	AAGTGGTGTTTTCTTTGGAGGA	Presence 3'half of insert
8	CCACAGGGATACCTTTCTCGTGC	Presence of *tet*(44) gene
9	TTCCATATCCTCGGGTTTTTGCAT	Presence of *tet*(44) gene
10	CAGGTGTTGAAATAGATATTGAG	Detect 3' end half insert
11	CAGAAGTCGATCCTTTCTGGG	Detect 3' end half insert
12	GGTGGCTGAACTCGTTAATC	Detect 3' end half insert
13	CTCCACATGGCTCGAGTTG	Detect 3' end half insert
14	GAGGAATTTAACAGAACAGTATTT	Excision studies
15	TCTATCCTGCCTTCTCAACC	Excision studies
16	CGAATCGCTGAAATGACTGA	Excision studies
17	GCGAATGATTTCATGGAAGG	Excision studies
18	CGACTGCATTACCAGTTCCA	Excision studies
Lok1 (13)	AAAATATACTGCACATCTGTATAC	Transconjugant screening
Lok3 (13)	TTTACCAGAAAAAGTAGCTTTAA	Transconjugant screening
19	CAGCTGCAGTTTTTCCATGA	Transconjugant screening
20	GCAGCTAACGGTGATGACAA	Transconjugant screening
Tn*916* Fw (30)	GACGGAAGATACTTATACA	Transconjugant screening
Tn*916* Rev (30)	GCCTTTGGATTCATTCCTGC	Transconjugant screening

tetracycline was 8 μg/mL. Spectinomycin resistance was determined by an agar dilution method of *C. difficile* colonies on BHI agar plates, supplemented with increasing amounts of spectinomycin. Streptomycin resistance was tested by disk diffusion method, using Sensi-Neotabs (Rosco, Denmark) (Streptomycin 500 ug disks) on blood plates under anaerobic conditions at 37°C.

Statistical analyses

Patients samples with the full 100 kb insert were compared to patients' samples with a part of the insert or no insert. The Chi-square test and *t*-test were used to calculate the p-value. Analyses were performed using the SPSS for Windows software package, version 17.0.

MLVA

Sixty eight strains were subjected to MLVA, of which 39 were previously characterized (16). MLVA and construction of the minimal spanning tree based on the MLVA results were carried out as described previously (16).

Results

Previously, an insert, unique for *C. difficile*, was described in the genome of strain M120, a PCR Ribotype078 strain, isolated from an Irish diarrheic patient (5). We analyzed the open reading frames (ORFs) present in the insert to investigate their nature and origin (see Figure 1 and Table 2).

The 100 kb insert has a modular composition

Bioinformatic analysis revealed that the insert has a modular composition. The 3' end of the insert (module E) is homologous to Tn*1806* of *S. pneumoniae* which confers erythromycin resistance. Although this element has not been shown to transfer via conjugation, transfer via transformation was shown (22). In *C. difficile* strain M120 this element appears to be the backbone into which several other elements have been inserted (see Figure 1 top panel). The first 7.3 kb on the 5' end of the insert (module A) has only moderate homology (60–70% maximum sequence identity) to known sequences. Interestingly, this part of the insert contains 2 putative modification DNA methylases and a putative endonuclease, possibly enabling a form of molecular vaccination as described by Kobayashi *et al.* (23). During this process methylation protects the incoming element from host endonucleases and, following integration, will protect the host chromosome from endonucleases present on other mobile genetic elements. This sequence is followed by a complete prophage of approximately 39.5 kb (module B), which shows 92% sequence identity to a *Thermoanaerobacter sp.* prophage (Genbank

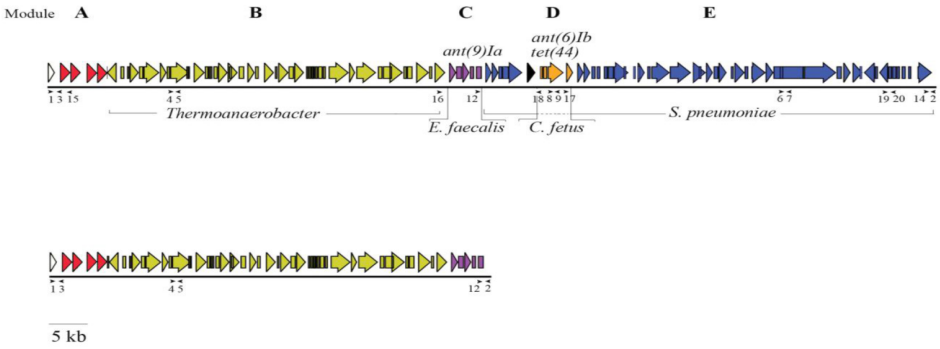


Figure 1: Schematic view of full Tn*6164* (top panel) and half the element (bottom panel) and its open reading frames, flanked by *C. difficile* regions. Various parts of the insert are colored according to their homology. White, *C. difficile*; Red, Module A; Yellow, Module B; Purple, Module C; Orange, Module D; Blue, Module E; black, unknown. Location of the oligonucleotides used for the data in Table 3 is indicated by arrowheads.

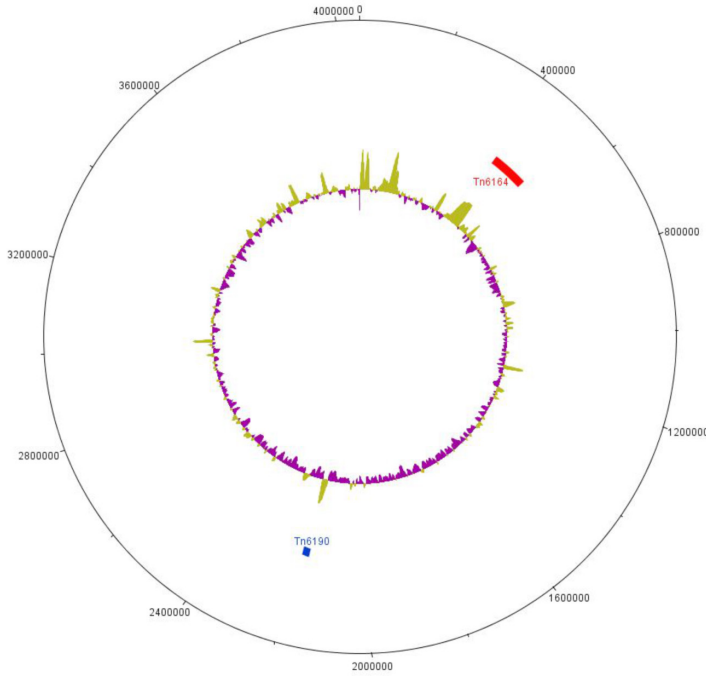


Figure 2: Circular representation of the genome of *C. difficile* strain M120. The two concentric circles represent the genome (outer circle) and the G + C content (inner circle; window size 10,000; Step size 200). Green represents values higher than average (29%), purple below average. In between the two circles, the presence of the two transposable elements is indicated in red (Tn*6164*) and blue (Tn*6190*). Figure was created using DNA plotter (46)

accession no. CP002210). The next 4.5 kb stretch (module C) is 99% identical to part of the *Enterococcus faecalis* plasmid pEF418 containing, amongst others, a putative methyltransferase and a putative spectinomycin adenyltransferase (*ant(9) Ia*) (24). It is also described to be part of a pathogenicity island in *Streptococcus suis* (25). Finally, an insertion of approximately 4.5 kb (module D) with 90% sequence identity to the transferable pathogenicity island of *Campylobacter fetus subsp fetus* (26) is present within the sequence of Tn*1806*. This sequence contains, amongst others, putative *tet* (44) and *ant(6)-Ib* genes, which could respectively confer tetracycline and streptomycin resistance.

The G + C content of the entire insert (34%) was significantly higher than that of the entire genome (29%), clearly indicating that the insert was of foreign origin (Figure 2). In addition, within the insert the different modules could be distinguished by their G + C contents. The G + C content of module A, B, C, D and E was 31%, 41%, 35%, 28% and 31%, respectively.

The 100 kb insert is a transposon

Based on the bioinformatic comparison of the insert described above, the possible excision of 3 (independent) elements was predicted. Primers were designed (primers 14–20, see Table 1) to amplify the circular intermediates of the complete insert (primers 14 and 15), the putative *Thermoanaerobacter* sp. phage (module B, primers 15 and 16) and the *C. fetus* pathogenicity island (module D, primers 17 and 18) of the element. PCR confirmed only the excision and circularisation of the entire insert (results not shown). It is expected that the serine recombinase at the 3' end of the element is responsible for excision (see Table 2). Sequencing of the circular intermediate was used to determine the precise ends of the element, showing the element is flanked by a TG dinucleotide; serine recombinases prefer a 2 bp crossover site identical in the target site and joint of the circular intermediate (27). *In silico* extraction of this sequence from the genome confirms that the element is present in the homologous target site of CTn*2* in strain 630 (7). The precise size of the element is 106,711 bp and it runs from bp 418,525-525,236 (including the TG dinucleotide at both ends) in the M120 genomic sequence (GenBank accession no. FN665653). Upon our request, the transposon number Tn*6164* was provided by the Transposon registry (28) (http://www.ucl.ac.uk/eastman/tn/index.php). To test the conjugative transfer of the element, filter mating assays were performed, selecting for the possible tetracycline resistance by means of the *tet*(44) gene. However, M120 contains also a copy of *tet*(M) present on a conjugative transposon with 97% sequence identity to Tn*916* (16), which we have designated Tn*6190*. This element has inserted intragenically in the homologue of *C. difficile* strain 630 ORF CD2015. Tn*6190* contains homologues to all Tn*916* ORFs except *orf*12 which is involved in regulation of *tet*(M) through transcriptional attenuation (29).

Table 2: Open reading frames encoded by Tn*6164*.

Gene	Position on Tn*6164*	Module	Sequence identity to	Annotation	Gene	Position on Tn*6164*	Module	Sequence identity to	Annotation
Orf1	650-1930	A	-	putative modification methylase	*Orf47*	50511-51038	C	*E. faecalis* pEF418	putative adenine phosphoribosyltransferase
Orf2	1915-3186	A	-	putative modification methylase	*Orf48*	51251-51979	C	*E. faecalis* pEF418	putative spectinomycin/ streptomycin adenyltransferase
Orf3	3252-3962	A	-	hypothetical protein	*Orf49*	52403-53176	E	*S. pneumoniae*	phage protein/replication initiator
Orf4	3952-5031	A	-	ATPase associated with various cellular activities	*Orf50*	53176-54000	E	*S. pneumoniae*	DNA replication protein
Orf5	5047-6312	A	-	LlaJI restriction endonuclease	*Orf51*	53993-54478	E	*S. pneumoniae*	DUF 3801
Orf6	C 7557-6361	A	-	Protein with unknown function, contains a C-terminal CGNR Zinc finger motif	*Orf52*	54475-55209	E	*S. pneumoniae*	phage antirepressor protein
Orf7	8000-8494	B	*Thermoanaerobacter sp.*	ECF RNA polymerase sigma-24 factor	*Orf53*	55202-56890	E	*S. pneumoniae*	TraG/TraD family protein
Orf8	8809-9126	B	*Thermoanaerobacter sp.*	rRNA biogenesis protein rrp5, putative	*Orf54*	57454-58486	E	-	DUF 318 Predicted Permease (HHPred)
Orf9	9123-10250	B	*Thermoanaerobacter sp.*	Phage associated protein	*Orf55*	59048-59398	D	*C. fetus*	glyoxalase family protein
Orf10	10256-10816	B	*Thermoanaerobacter sp.*	phage-associated protein	*Orf56*	59411-59938	D	*C. fetus*	transcriptional regulator
Orf11	10813-12747	B	*Thermoanaerobacter sp.*	DNA-directed DNA polymerase	*Orf57*	59988-61910	D	*C. fetus*	tetracycline resistance protein

Gene	Position on Tn*6164*	Module	Sequence identity to	Annotation	Gene	Position on Tn*6164*	Module	Sequence identity to	Annotation
Orf12	12795-13625	B	*Thermoanaerobacter sp.*	Prophage antirepressor	*Orf58*	62225-63082	D	*C. fetus*	aminoglycoside 6-adenylyltransferase (AAD(6)
Orf13	13629-14048	B	*Thermoanaerobacter sp.*	DUF 4406 (HHPred)	*Orf59*	63575-64348	E	*S. pneumoniae*	replication initiator/phage
Orf15	16910-18259	B	*Thermoanaerobacter sp.*	SNF2-related protein	*Orf61*	65314-65814	E	*S. pneumoniae*	TnpX site-specific recombinase family protein
Orf16	18264-18722	B	*Thermoanaerobacter sp.*	phage-associated protein	*Orf62*	65938-66399	E	*S. pneumoniae*	Flavodoxin
Orf17	18842-19201	B	*Thermoanaerobacter sp.*	HNH endonuclease	*Orf63*	66817-67302	E	*S. pneumoniae*	putative conjugative transposon protein
Orf18	19314-19865	B	*Thermoanaerobacter sp.*	Phage terminase, small subunit	*Orf64*	67299-68033	E	*S. pneumoniae*	phage antirepressor protein
Orf19	19883-21058	B	*Thermoanaerobacter sp.*	S-adenosylmethionine synthetase	*Orf65*	68026-69816	E	*S. pneumoniae*	TraG/TraD family protein/putative conjugal transfer protein
Orf20	21039-22283	B	*Thermoanaerobacter sp.*	DNA methylase N-4/N-6 domain-containing protein	*Orf66*	70395-70706	E	*S. pneumoniae*	putative single-strand binding protein
Orf21	22384-23076	B	*Thermoanaerobacter sp.*	hypothetical/virulence-related protein	*Orf67*	70934-71797	E	*S. pneumoniae*	conjugative transposon membrane protein
Orf22	23445-24344	B	*Thermoanaerobacter sp.*	Putative amidoligase enzyme	*Orf68*	72099-72509	E	*S. pneumoniae*	conjugative transposon membrane protein
Orf23	24382-24843	B	*Thermoanaerobacter sp.*	AIG2/GGCT-like protein	*Orf69*	72580-74823	E	*S. pneumoniae*	type IV conjugative transfer system protein

Gene	Position on Tn*6164*	Module	Sequence identity to	Annotation	Gene	Position on Tn*6164*	Module	Sequence identity to	Annotation
Orf24	25462-26685	B	*Thermoanaerobacter sp.*	phage terminase	*Orf70*	74831-77410	E	*S. pneumoniae*	conjugative transposon cell wall hydrolase/NlpC/P60 family
Orf25	26793-27122	B	-	conserved hypothetical protein	*Orf71*	77648-79216	E	*S. pneumoniae*	putative surface protein
Orf26	27189-28451	B	*Thermoanaerobacter sp.*	HK97 family phage portal protein	*Orf72*	79231-80088	E	*S. pneumoniae*	putative bacteriocin
Orf27	28448-29128	B	*Thermoanaerobacter sp.*	Peptidase S14, ClpP	*Orf73*	80162-80773	E	*S. pneumoniae*	Predicted transcriptional regulator
Orf28	29140-30339	B	*Thermoanaerobacter sp.*	HK97 family phage major capsid protein	*Orf74*	80766-81749	E	*S. pneumoniae*	Protein with unknown function
Orf29	30585-30899	B	*Thermoanaerobacter sp.*	uncharacterized phage protein	*Orf75*	82268-82621	E	*S. pneumoniae*	transcriptional regulator, ArsR family
Orf30	30903-31238	B	*Thermoanaerobacter sp.*	phage head-tail adaptor, putative	*Orf76*	82696-83940	E	*S. pneumoniae*	major facilitator superfamily MFS_1
Orf31	31252-31662	B	*Thermoanaerobacter sp.*	HK97 family phage protein	*Orf77*	83927-84403	E	*S. pneumoniae*	toxin-antitoxin system, toxin component, GNAT domain protein
Orf32	31659-32012	B	*Thermoanaerobacter sp.*	Protein of unknown function (DUF806);	*Orf78*	84758-86491	E	*S. pneumoniae*	DNA topoisomerase III
Orf33	32016-32618	B	*Thermoanaerobacter sp.*	DUF3647 Phage protein (HHPred)	*Orf79*	86484-87449	E	*S. pneumoniae*	possible DNA (cytosine-5-)-methyltransferase
Orf34	33330-35786	B	*Thermoanaerobacter sp.*	Phage tape measure protein	*Orf80*	87436-95079	E	*S. pneumoniae*	superfamily II DNA and RNA helicase
Orf35	35800-36573	B	*Thermoanaerobacter sp.*	phage putative tail component	*Orf81*	95123-95779	E	*S. pneumoniae*	putative single-stranded DNA binding protein

Gene	Position on Tn*6164*	Module	Sequence identity to	Annotation	Gene	Position on Tn*6164*	Module	Sequence identity to	Annotation
Orf36	36692-39100	B	*Thermoanaerobacter sp.*	phage minor structural protein	*Orf82*	95939-96841	E	*S. pneumoniae*	transcriptional regulator, XRE family
Orf37	39320-39901	B	*Thermoanaerobacter sp.*	Putative Sipho Phage tail protein (HHPred)	*Orf83*	97071-98282	E	*S. pneumoniae*	transporter, major facilitator family/ multidrug resistance protein 2
Orf38	39928-42369	B	*Thermoanaerobacter sp.*	glycosyl hydrolase-like protein	*Orf84*	C 99739-98462	E	*S. pneumoniae*	relaxase/type IV secretory pathway protein VirD2
Orf39	42430-42855	B	*Thermoanaerobacter sp.*	toxin secretion/phage lysis holin	*Orf85*	C 101169-99795	E	*S. pneumoniae*	conjugal transfer relaxosome component TraJ
Orf40	42855-43556	B	*Thermoanaerobacter sp.*	N-acetylmuramoyl-L-alanine amidase	*Orf86*	C 101403-100321	E	*S. pneumoniae*	toxin-antitoxin system, toxin component, Fic family
Orf41	43975-45540	B	*Thermoanaerobacter sp.*	phage integrase family site-specific recombinase / resolvase	*Orf87*	C 101878-101396	E	*S. pneumoniae*	putative membrane protein
Orf42	45541-45954	B	*Thermoanaerobacter sp.*	recombinase/integrase	*Orf88*	C 102435-101887	E	*S. pneumoniae*	putative toxin-antitoxin system, toxin component
Orf43	46222-47529	B	*Thermoanaerobacter sp.*	phage integrase family site-specific recombinase	*Orf89*	C 102845-102444	E	*S. pneumoniae*	regulator/toxin-antitoxin system, antitoxin component
Orf44	47987-48856	C	*E. faecalis* pEF418	Nucleotidyl transferase	*Orf90*	103034-103555	E	*S. pneumoniae*	conserved hypothetical protein
Orf45	48837-49571	C	*E. faecalis* pEF418	methyltransferase	*Orf91*	103825-104235	E	*S. pneumoniae*	sigma-70, region 4
Orf46	49604-50467	C	*E. faecalis* pEF418	putative aminoglycoside 6-adenylyltansferase	*Orf92*	104966-106712	E	*S. pneumoniae*	site-specific recombinase, resolvase family

During filter mating experiments with M120 as a donor strain and CD37 as a recipient, all putative transconjugants were identified as the recipient strain. In total 70 transconjugants were tested by PCR, using primers Lok1, Lok3 (13), 19, 20, Tn*916* Fw, and Tn*916* Rev (30). However, none contained Tn*6164,* all contained only Tn*6190* (results not shown).

Tn*6164* is sporadically present in PCR Ribotype078

Simultaneously with the publication of the M120 sequence, we obtained Illumina sequence reads of the *C. difficile* strain 31618, which was isolated from a diarrheic piglet from a pig farm in the Netherlands (16). Comparative genomic analysis of 31618 to M120 revealed an almost complete overlap of the two genomes. However, reference assembly of the 31618 reads to M120 showed that Tn*6164* was not present in 31618 (results not shown). This prompted us to investigate the prevalence of Tn*6164* in RT 078 strains. We designed a PCR to show presence (primers 1 and 3) or absence (primers 1 and 2) of Tn*6164* in RT 078 genomic DNA (see Figure 1 top panel). In addition, in view of the heterogeneous origin of Tn*6164* and to investigate the presence of both the *Thermoanaerobacter* prophage and *Streptococcus* DNA (Modules B and E, respectively), we designed two more PCRs (primers 4–5 and 6–7). Finally, we designed a PCR to detect the presence of the *tet*(44) gene present on Tn*6164* (Module D, primers 8 and 9). Besides the sequenced 31618 strain, 173 human RT 078 strains and 58 porcine RT 078 strains (from 27 pig farms) were tested for the presence of these elements.

A minority of the isolates tested did contain a DNA insert at the indicated location in the genome; 18 of the 231 isolates (7.8%) were positive in the 1–3 PCR (Table 3). Remarkably, all 18 strains were tetracycline resistant human isolates. None of the porcine strains contained an insert at the position tested. Strains positive in the 1–3 PCR were negative in the 1–2 PCR, and *vice versa*, showing complete complementarity of the two PCRs in RT 078 strains.

Evidence for multiple insertions in Tn*6164*

All the strains that contained an insert (based on the 1–3 PCR) were further analyzed for the presence of Module B and E present in Tn*6164*, using primer pairs 4–5 and 6–7 (see Figure 1 top panel and Table 1). Only nine of 18 strains positive for PCR 1–3 were positive for PCRs 4–5 and 6–7, suggesting the presence of the complete element as described for M120. The other 9 strains were only positive for Module B (PCR 4–5), showing the existence of alternative (shorter) elements (see Table 3), as predicted by the bioinformatic analysis. The strains that were positive for Module E (PCR 6–7) were also positive for Module D (PCR 8–9, see Table 3). In contrast, strains containing Module B, but not Module E, thus containing only half the

element, also lacked Module D. This indicates that the 3'end of half the element was situated upstream of Module D. Of the isolates that were only positive for the PCR 4–5, the exact 3'end of the insert was determined by sequencing the PCR product obtained with primers 12 and 2 (see Figure 1 bottom panel), which yielded a 350 bp product. The border of the 3'end was between the 3' end of Module C and the 5'end of Module E. A similar sequence was found at the homologous site when the full element was present, but also at the 3' end of the full element, the 5' end of the element, the joint of the circular intermediate and the predicted target site as based on the 630 sequence (see Table 4). This indicates that Tn*6164* was created by two elements integrating in the same target site (next to each other) and fusing, with a second copy of the target site still present between the two original elements within Tn*6164*.

Table 3: Detection of specific regions of Tn*6164* in RT 078 strains.

Strain	PCR 1-2*	PCR1-3§	PCR 4-5#	PCR 6-7¶	PCR 8-9†	PCR 12-2‡
56/69	-	+	+	-	-	+
26222	-	+	+	-	-	+
26114	-	+	+	-	-	+
26247	-	+	+	-	-	+
26235	-	+	+	-	-	+
6065935	-	+	+	-	n.t.	n.t.
ES1203	-	+	+	-	-	n.t.
7047337	-	+	+	-	n.t.	n.t.
8088158	-	+	+	-	n.t.	n.t.
50/19	-	+	+	+	+	-
GR0106	-	+	+	+	+	n.t.
DE1210	-	+	+	+	+	n.t.
BG1209	-	+	+	+	+	n.t.
NO1311	-	+	+	+	+	n.t.
NO1307	-	+	+	+	+	n.t.
IE1102	-	+	+	+	+	n.t.
GR0301	-	+	+	+	+	n.t.
10053737	-	+	+	+	n.t.	n.t.

*PCR only positive when no insert is present, §PCR only positive when insert is present #PCR detects Module B, ¶PCR detects module E, †PCR detects module D. ‡ PCR only positive in strains containing half of the element. Location of the oligonucleotides used is indicated in Figure 1. +, PCR positive; -, PCR negative; n.t., not tested

The sequences at the 3' end of the element in strains that contain half the insert or the full insert are identical. These are related to the sequence at the 5' end of the element and the middle section of the full element and also to the joint of the circular intermediate of Tn*6164* and the empty target site, compared to the empty target site of CTn*2* from strain 630. Sequence shown in underlined bold is the dinucleotide which is predicted to be recognised by the serine recombinase.

Table 4: Sequences of the joints between the genome and Tn*6164* and the joint of the circular form.

CGCAT**TG**CG-AGACTATAG	3'ends of half insert
CGCAT**TG**CG-AGACTATAG	3'ends of full insert
CTCA-**TG**TGGAGTGCGTGG	5'end of full insert
GCCA-**TG**TGGAGACTATAG	middle section of full element
CACA-**TG**CGTTGTCTTGTG	Joint of circular intermediate Tn*6164*
CACAT**TG**TG-AGACTGTAG	CTn*2* target site in strain 630

Absence of Tn*6164* sequences in other PCR ribotypes

Since RT 126 has been shown to be very closely related to RT 078, with an almost indistinguishable PCR Ribotype banding pattern, we also tested a small collection of RT 126 strains with the 1–2 and 1–3 PCRs. In none of the 10 RT 126 strains tested could we demonstrate the presence of an insert at the site in which Tn*6164* was inserted in M120 (results not shown).

In addition, a collection of 66 other PCR ribotypes was tested as well. This collection consisted of the 25 most frequently found PCR ribotypes in Europe, supplemented with the Leeds-Leiden collection (31). None of the other PCR ribotypes, was positive for PCR 1–3, 4–5 or 6–7.

No antibiotic resistance phenotype linked to presence of Tn*6164*

Since several putative antibiotic resistance genes were found to be present on the element (see Figure 1 and Table 2), strains containing full Tn*6164*, only half of the element, or no element at all were tested for antibiotics resistance. Resistance to tetracycline, spectinomycin and streptomycin was tested using several methods (see materials and methods). Surprisingly, no correlation was found between the presence of *tet*(44), *ant*(*6*)*Ib* or *ant*(*9*)*Ia* and resistance to tetracycline, spectinomycin or streptomycin (see Table 5).

Table 5: Antibiotic sensitivity of RT 078 strains with(out) Tn*6164*.

Genes present (transposon)	Strain	MIC Tet (µg/mL)	MIC Spec (µg/mL)	Strep
	56/69	24	> 750	N.D.
	26222	16	N.D.	R
ant(*9*)*Ia* (Tn*6164*)	26114	32	N.D.	R
tet(M) (Tn*6190*)	26247	16	> 750	R
	26235	48	N.D.	N.D.
	06065935	8	N.D.	R
	50/19	48	>750	S
	GR0106	12	>750	R
ant(*9*)*Ia* (Tn*6164*)	DE1210	8	>750	R
ant(*6*) (Tn*6164*)	BG1209	8	>750	R
tet(44) (Tn*6164*)	NO1311	12	>750	R
tet(M) (Tn*6190*)	NO1307	8	>750	R
	IE1102	12	>750	R
	GR0301	8	>750	R
	10053737	N.D	N.D	R
tet(M) (Tn*6190*)	45/22	8	>750	N.D.
	29/74	<8	>750	N.D.
	31618	N.D.	<250	N.D.
None	07053152	<8	N.D.	R
	R20291(027)	N.D.	<250	N.D.

R, resistant (no halo around diffusion disk); S, sensitive (15 mm halo).

Strains containing full Tn*6164* are all genetically related

Since we could not find many isolates containing Tn*6164*, we reasoned that the element could be relatively recently acquired and that the isolates thus might be genetically closely related. Therefore, we applied MLVA (3,16) on all the isolates containing Tn*6164*, or only half of it, supplemented with a number of isolates without the element, to investigate the genetic relatedness of the strains. In Figure 3, a minimal spanning tree of all the isolates containing an element is shown, with control strains. Based on the MLVA, all the isolates containing full Tn*6164* (n = 9) are genetically related (STRD < 10) and four of them are in one clonal complex. Six isolates containing half of the element are also in this genetically related cluster, whereas the other three isolates containing half the element are not (STRD > 10).

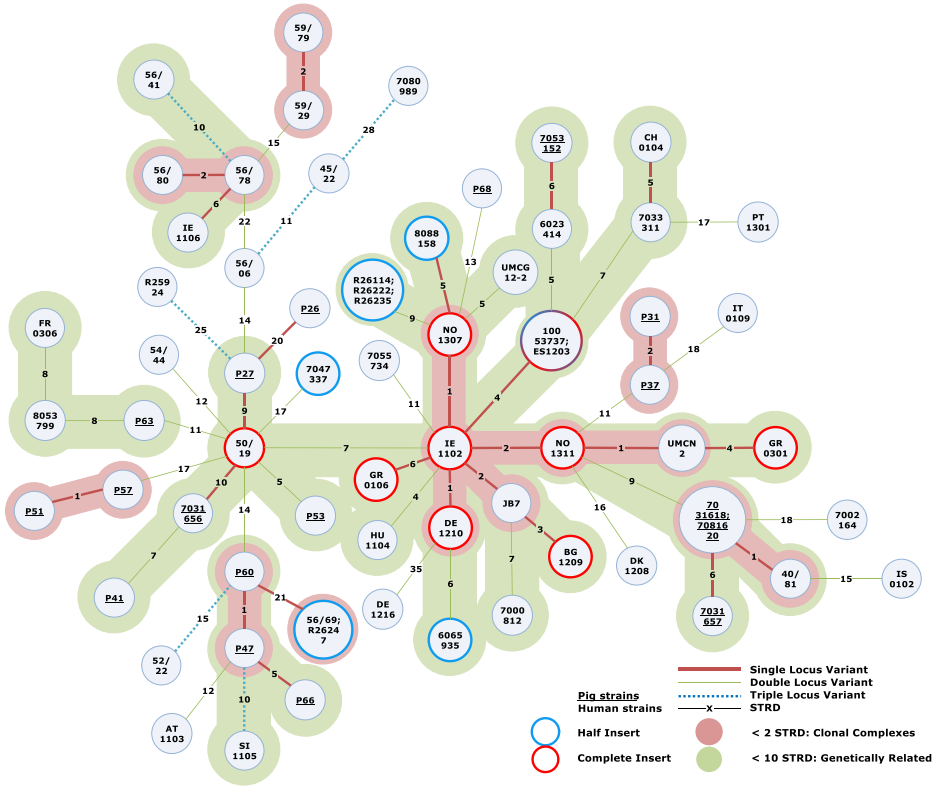


Figure 3: Minimum spanning tree of all the RT 078 isolates that contained an insert (50 or 100 kb), supplemented with strains not containing the element. Each circle represents either one unique isolate or more isolates that have identical MLVA types. Red circles indicate strains with full Tn*6164* and blue circles indicate strains with half the element. The numbers between the circles represent the summed tandem-repeat differences (STRD) between MLVA types. Underlined numbers represent porcine strains and normal numbers represent human isolates. Thick red lines represent single-locus variants; thin green lines represent double-locus variants and dotted blue lines represent triple locus variants between MLVA types. Clonal clusters are defined by an STRD of <2 (pink area), and genetically related clusters are defined by an STRD of <10 (green area).

Suggestive link between the 100 kb insert and increased virulence

To investigate a possible increased virulence of strains containing the element, clinical parameters of patients with a *C. difficile* infection due to a strain that contained Tn*6164* were compared to parameters of patients that suffered from a strain that did not contain the full element. Patients with Tn*6164* resembled patients without the element concerning demographic characteristics. Clinical characteristics were only known for patients from the ECDIS study (32) and patients registered in the CDRL (n = 84). Patients with and without the element suffered from severe diarrhea in similar proportions. Mortality due to CDI was more common in patients infected with *C. difficile*::Tn*6164* (29% *vs* 3%). This suggests that Tn*6164* might convert PCR Ribotype078 strains to a more virulent strain. However, since the number of patients infected with a Tn*6164*-positive strain, and for which the clinical data was available, was very low (n = 7), no multivariate analysis could be performed, which means that a bias cannot be ruled out. Further research is needed to confirm a possible link between increased virulence and the presence of Tn*6164*.

Discussion

PCR Ribotype 078 has recently emerged as a hypervirulent *C. difficile* strain (2,3). Previously published MLVA studies have shown that all PCR Ribotype 078 strains are closely related (3), irrespective of human or porcine origin (16), fostering the notion that PCR Ribotype 078 infection could be a zoonosis. Recently, the full genome sequence of a *C. difficile* PCR Ribotype 078 strain was published (5). This M120 strain was shown to contain a unique insert of approximately 100 kilobases. In this paper we show that this insert is a transposable element, Tn*6164*. It is not representative for all PCR Ribotype 078 strains. On the contrary, we found that the majority of the PCR Ribotype 078 strains do not contain the element. Moreover, some strains contain only half of the element. So, three different kinds of PCR Ribotype 078 can now be distinguished: Those with a full length element, those with half the element, and those with no element at all. Tn*6164* was exclusively found in tetracycline resistant PCR Ribotype 078 strains, isolated from humans. We tested a collection of other PCR ribotypes, of which none contained the element. Since we only tested 1 strain per PCR ribotype, we cannot rule out the possibility that Tn*6164* is present in other PCR ribotypes. We covered the whole genomic spectrum of *C. difficile* since we tested multiple samples of each genetic clade previously identified (10,33-35). In addition, Tn*6164* has not been found in any other *C. difficile* genome that has been published so far than M120. Although

Tn*6164* contained a *tet*(44) gene, we could not demonstrate increased tetracycline resistance of strains containing the element.

Previously, it has been shown that this gene, present on a homologues resistance island, is active in *C. fetus* (26). In *C. difficile*, the copresence of the *tet*(44) gene on Tn*6164* and the *tet*(M) gene on the Tn*6190* in one bacterium does not result in an increased resistant phenotype. Also the spectinomycin and streptomycin resistance genes did not result in a phenotype, despite the presence of two potential aminoglycoside resistance genes (*ant*(*9*)*Ia*) and *ant*(*6*)) on Tn*6164* (see Figure 1 and Table 2). We do not know if the resistance genes are expressed in M120. However, since we show the presence of the circular intermediate transposon DNA, some activity of transposon related genes is expected.

Since we have only found Tn*6164* in strains also containing Tn*6190*, it is possible that Tn*6164* transfer is dependent on Tn*6190*. Further research is needed to investigate the possibility of Tn*6190* dependent transfer of Tn*6164*. In addition, remarkably, Tn*6164* (the whole or half the element) was significantly ($p = 0.01$) more found in strains isolated from humans than in strains isolated from pigs. Although the same strains circulate in humans and pigs (16), and also Tn*6190* circulates in pig strains (16), we did not find any porcine strain that contained the element. We have no explanation for this difference. None of the transconjugants tested showed the presence of Tn*6164*, but all contained Tn*6190*. These results indicate that Tn*6164* has a (much) lower transfer frequency than Tn*6190*. Nevertheless, a complete set of proteins, required for transfer, is present on Tn*6164*. Loss of Tn*6190* or introduction of another selection marker in Tn*6164* (11) could prove to be a strategy to further study the capability of conjugative transfer of this element.

Tn*6164* has integrated intergenically between homologs of the 630 ORFs CD0406 and CD0437, a tRNA methyltransferase and a hypothetical protein respectively. In strain 630, this target site is occupied by the conjugative transposon CTn*2* (7,11). There is no significant homology between Tn*6164* and CTn*2*. The empty target site is present in many sequenced strains of *C. difficile*. However, no other mobile genetic elements have been reported to integrate at this site.

It was impossible to phenotypically distinguish strains containing Tn*6164* from strains without the element. Although we have no transcriptional data available of the genes that are located on Tn*6164* it is clear that it could provide an advantage under certain circumstances. In this respect it is interesting to note that the patients suffering from an element-containing strain are suggested to undergo a more severe illness than patients with a strain not containing Tn*6164*. However, because of the low number of strains containing the insert no multivariate analysis could be carried out. Therefore, we cannot rule out that these data are biased. Further research is needed to confirm this observation.

Isolates containing the full element originated from all over Europe, including Ireland, England, Norway, Germany, Bulgaria, Greece and the Netherlands, whereas isolates containing half the element were only found in the United Kingdom, Spain and the Netherlands. MLVA showed genetic relatedness between most of the strains, although no epidemiologic link between the strains from different countries could be found. It has recently been shown that RT 078 strains show a lot less heterogeneity in MLVA than for instance RT 027 or RT 017 (36-38). This could indicate a higher level of relatedness, or it could mean that the mechanism behind the MLVA variability is different in RT 078 strains than in other RTs (16).

Altogether, we show the presence of a 100 kb transposon in some *C. difficile* RT 078 strains. Although we could not show any evolutionary benefits of the transposon, it could very well serve as a reservoir of antibiotic resistance (26), for commensal bacteria in the human gut.

Conclusion

Tn*6164* is a novel transposon of approximately 100 kb, found sporadically in *Clostridium difficile* RT 078 strains, isolated from humans. Tn*6164* has a modular composition and is the product of multiple insertions of separate elements from various origins, as evidenced by the existence of strains containing only half the element. Strains containing Tn*6164* were all genetically related. We were not able to find a readily distinguishable phenotype for strains containing the element, although several potential antibiotic resistance genes were present on Tn*6164*. Tn*6164* may act as a source of antibiotic resistance genes in the human gut. Further research is needed to investigate if Tn*6164* plays a role in the virulence of RT 078 *Clostridium difficile* strains.

Reference List

1. Pepin J, Valiquette L, Cossette B: Mortality attributable to nosocomial *Clostridium difficile*-associated disease during an epidemic caused by a hypervirulent strain in Quebec. *CMAJ* 2005, 173:1037–1042.
2. Goorhuis A, Debast SB, van Leengoed LA, Harmanus C, Notermans DW, Bergwerff AA, *et al*: *Clostridium difficile* PCR Ribotype 078: an emerging strain in humans and in pigs? *J Clin Microbiol* 2008, 46:1157.
3. Goorhuis A, Bakker D, Corver J, Debast SB, Harmanus C, Notermans DW, *et al*: Emergence of *Clostridium difficile* infection due to a new hypervirulent strain, polymerase chain reaction Ribotype 078. *Clin Infect Dis* 2008, 47:1162–1170.
4. Debast SB, van Leengoed LA, Goorhuis A, Harmanus C, Kuijper EJ, Bergwerff AA: *Clostridium difficile* PCR Ribotype 078 toxinotype V found in diarrhoeal pigs identical to isolates from affected humans. *Environ Microbiol* 2009, 11:505–511.
5. He M, Sebaihia M, Lawley TD, Stabler RA, Dawson LF, Martin MJ, *et al*: Evolutionary dynamics of *Clostridium difficile* over short and long time scales. *Proc Natl Acad Sci USA* 2010, 107:7527–7532.
6. Stabler RA, He M, Dawson L, Martin M, Valiente E, Corton C, *et al*: Comparative genome and phenotypic analysis of *Clostridium difficile* 027 strains provides insight into the evolution of a hypervirulent bacterium. *Genome Biol* 2009, 10:R102.
7. Sebaihia M, Wren BW, Mullany P, Fairweather NF, Minton N, Stabler R, *et al*: The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome. *Nat Genet* 2006, 38:779–786.
8. Forgetta V, Oughton MT, Marquis P, Brukner I, Blanchette R, Haub K, *et al*: Fourteen-Genome Comparison Identifies DNA Markers for Severe-Disease-Associated Strains of *Clostridium difficile*. *J Clin Microbiol* 2011, 49:2230–2238.
9. Marsden GL, Davis IJ, Wright VJ, Sebaihia M, Kuijper EJ, Minton NP: Array comparative hybridisation reveals a high degree of similarity between UK and European clinical isolates of hypervirulent *Clostridium difficile*. *BMC Genomics* 2010, 11:389.
10. Stabler RA, Gerding DN, Songer JG, Drudy D, Brazier JS, Trinh HT, *et al*: Comparative phylogenomics of *Clostridium difficile* reveals clade specificity and microevolution of hypervirulent strains. *J Bacteriol* 2006, 188:7297–7305.
11. Brouwer MSM, Warburton PJ, Roberts AP, Mullany P, Allan E: Genetic Organisation, Mobility and Predicted Functions of Genes on Integrated, Mobile Genetic Elements in Sequenced Strains of *Clostridium difficile*. *PLoS One* 2011, 6:e23014.
12. Tan KS, Wee BY, Song KP: Evidence for holin function of *tcdE* gene in the pathogenicity of *Clostridium difficile*. *J Med Microbiol* 2001, 50:613–619.
13. Braun V, Hundsberger T, Leukel P, Sauerborn M, von Eichel-Streiber C: Definition of the single integration site of the pathogenicity locus in *Clostridium difficile*. *Gene* 1996, 181:29–38.
14. Govind R, Vediyappan G, Rolfe RD, Dupuy B, Fralick JA: Bacteriophage-mediated toxin gene regulation in *Clostridium difficile*. *J Virol* 2009, 83:12037–12045.
15. Sekulovic O, Meessen-Pinard M, Fortier LC: Prophage-Stimulated Toxin Production in *Clostridium difficile* NAP1/027 Lysogens. *J Bacteriol* 2011, 193:2726–2734.

16. Bakker D, Corver J, Harmanus C, Goorhuis A, Keessen EC, Fawley WN, *et al*: Relatedness of human and animal *Clostridium difficile* PCR Ribotype 078 isolates determined on the basis of multilocus variable-number tandem-repeat analysis and tetracycline resistance. *J Clin Microbiol* 2010, 48:3744–3749.
17. Adams V, Lyras D, Farrow KA, Rood JI: The clostridial mobilisable transposons. *Cell Mol Life Sci* 2002, 59:2033–2043.
18. Roberts AP, Mullany P: A modular master on the move: the Tn*916* family of mobile genetic elements. *Trends Microbiol* 2009, 17:251–258.
19. Brouwer MSM, Roberts AP, Mullany P, Allan E: In silico analysis of sequenced strains of *Clostridium difficile* reveals a related set of conjugative transposons carrying a variety of accessory genes. *Mobile Genetic Elements* 2012, 2. http://dx.doi.org/10.4161/mge.2.1.19297.
20. Mullany P, Wilks M, Lamb I, Clayton C, Wren B, Tabaqchali S: Genetic analysis of a tetracycline resistance element from *Clostridium difficile* and its conjugal transfer to and from Bacillus subtilis. *J Gen Microbiol* 1990, 136:1343–1349.
21. Wang H, Roberts AP, Lyras D, Rood JI, Wilks M, Mullany P: Characterization of the ends and target sites of the novel conjugative transposon Tn*5397* from *Clostridium difficile*: excision and circularization is mediated by the large resolvase, TndX. *J Bacteriol* 2000, 182:3775–3783.
22. Camilli R, Del GM, Iannelli F, Pantosti A: New genetic element carrying the erythromycin resistance determinant erm(TR) in Streptococcus pneumoniae. *Antimicrob Agents Chemother* 2008, 52:619–625.
23. Kobayashi I: Behavior of restriction-modification systems as selfish mobile elements and their impact on genome evolution. *Nucleic Acids Res* 2001, 29:3742–3756.
24. Murphy E: Nucleotide sequence of a spectinomycin adenyltransferase AAD(9) determinant from Staphylococcus aureus and its relationship to AAD(3") (9). *Mol Gen Genet* 1985, 200:33–39.
25. Chen C, Tang J, Dong W, Wang C, Feng Y, Wang J, *et al*: A glimpse of streptococcal toxic shock syndrome from comparative genomics of S. suis 2 Chinese isolates. *PLoS One* 2007, 2:e315.
26. Abril C, Brodard I, Perreten V: Two novel antibiotic resistance genes, *tet*(44) and *ant(6)-Ib*, are located within a transferable pathogenicity island in Campylobacter fetus subsp. fetus. *Antimicrob Agents Chemother* 2010, 54:3052–3055.
27. Smith MC, Thorpe HM: Diversity in the serine recombinases. *Mol Microbiol* 2002, 44:299–307.
28. Roberts AP, Chandler M, Courvalin P, Guedon G, Mullany P, Pembroke T, *et al*: Revised nomenclature for transposable genetic elements. *Plasmid* 2008, 60:167–173.
29. Su YA, He P, Clewell DB: Characterization of the tet(M) determinant of Tn*916*: evidence for regulation by transcription attenuation. *Antimicrob Agents Chemother* 1992, 36:769–778.
30. Ciric L, Mullany P, Roberts AP: Antibiotic and antiseptic resistance genes are linked on a novel mobile genetic element: Tn*6087*. *J Antimicrob Chemother* 2011, 66:2235–2239.
31. Knetsch CW, Hensgens MPM, Harmanus C, van der Bijl MW, Savelkoul PH, Kuijper EJ, *et al*: Genetic markers for *Clostridium difficile* lineages linked to hypervirulence. *Microbiology* 2011, 157:3113–3123.
32. Bauer MP, Notermans DW, van Benthem BH, Brazier JS, Wilcox MH, Rupnik M, *et al*: *Clostridium difficile* infection in Europe: a hospital-based survey. *Lancet* 2011, 377:63-73.

33. Griffiths D, Fawley W, Kachrimanidou M, Bowden R, Crook DW, Fung R, *et al*: Multilocus sequence typing of *Clostridium difficile*. *J Clin Microbiol* 2010, 48:770–778.
34. Stabler RA, Dawson LF, Valiente E, Cairns MD, Martin MJ, Donahue EH, *et al*: Macro and Micro Diversity of *Clostridium difficile* Isolates from Diverse Sources and Geographical Locations. *PLoS One* 2012, 7:e31559.
35. Dingle KE, Griffiths D, Didelot X, Evans J, Vaughan A, Kachrimanidou M, *et al*: Clinical *Clostridium difficile*: clonality and pathogenicity locus diversity. *PLoS One* 2011, 6:e19993.
36. Fawley WN, Freeman J, Smith C, Harmanus C, van den Berg RJ, Kuijper EJ, *et al*: Use of highly discriminatory fingerprinting to analyze clusters of *Clostridium difficile* infection cases due to epidemic Ribotype 027 strains. *J Clin Microbiol* 2008, 46:954–960.
37. van den Berg RJ, Schaap I, Templeton KE, Klaassen CH, Kuijper EJ: Typing and subtyping of *Clostridium difficile* isolates by using multiple-locus variable-number tandem-repeat analysis. *J Clin Microbiol* 2007, 45:1024–1028.
38. Goorhuis A, Legaria MC, van den Berg RJ, Harmanus C, Klaassen CH, Brazier JS, *et al*: Application of multiple-locus variable-number tandem-repeat analysis to determine clonal spread of toxin A-negative *Clostridium difficile* in a general hospital in Buenos Aires, Argentina. *Clin Microbiol Infect* 2009, 15:1080-1086.
39. Paltansing S, van den Berg RJ, Guseinova RA, Visser CE, van der Vorm ER, Kuijper EJ: Characteristics and incidence of *Clostridium difficile*-associated disease, The Netherlands, 2005. *Clin Microbiol Infect* 2007, 13:1058–1064.
40. Bidet P, Barbut F, Lalande V, Burghoffer B, Petit JC: Development of a new PCR-ribotyping method for *Clostridium difficile* based on ribosomal RNA gene sequencing. *FEMS Microbiol Lett* 1999, 175:261–266.
41. Hachler H, Kayser FH, Berger-Bachi B: Homology of a transferable tetracycline resistance determinant of *Clostridium difficile* with Streptococcus (Enterococcus) faecalis transposon Tn*916*. *Antimicrob Agents Chemother* 1987, 31:1033–1038.
42. Brouwer MS, Allan E, Mullany P, Roberts AP: Draft Genome Sequence of the Nontoxigenic *Clostridium difficile* Strain CD37. *J Bacteriol* 2012, 194:2125–2126.
43. van den Berg RJ, Claas EC, Oyib DH, Klaassen CH, Dijkshoorn L, Brazier JS, *et al*: Characterization of toxin A-negative, toxin B-positive *Clostridium difficile* isolates from outbreaks in different countries by amplified fragment length polymorphism and PCR ribotyping. *J Clin Microbiol* 2004, 42:1035–1041.
44. Carver T, Berriman M, Tivey A, Patel C, Bohme U, Barrell BG, *et al*: Artemis and ACT: viewing, annotating and comparing sequences stored in a relational database. *Bioinformatics* 2008, 24:2672–2676.
45. Hussain HA, Roberts AP, Mullany P: Generation of an erythromycin-sensitive derivative of *Clostridium difficile* strain 630 (630Deltaerm) and demonstration that the conjugative transposon Tn*916*DeltaE enters the genome of this strain at multiple sites. *J Med Microbiol* 2005, 54:137–141.
46. Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J: DNAPlotter: circular and linear interactive genome visualization. *Bioinformatics* 2009, 25:119–120.

Fundamental research

5

Clostridium difficile 630Δ*erm* Spo0A regulates sporulation, but does not contribute to toxin production, by direct high-affinity binding to target DNA

K. Rosenbusch, D. Bakker, E.J. Kuijper, W. Smits

Leiden University Medical Center (LUMC), Department of Medical Microbiology, Albinusdreef 2, 2333 ZA, Leiden, the Netherlands

PLoS One 2012

Abstract

Clostridium difficile is a Gram positive, anaerobic bacterium that can form highly resistant endospores. The bacterium is the causative agent of *C. difficile* infection (CDI), for which the symptoms can range from a mild diarrhea to potentially fatal pseudomembranous colitis and toxic megacolon.

Endospore formation in Firmicutes, including *C. difficile*, is governed by the key regulator for sporulation, Spo0A. In *Bacillus subtilis*, this transcription factor is also directly or indirectly involved in various other cellular processes.

Here, we report that *C. difficile* Spo0A shows a high degree of similarity to the well characterized *B. subtilis* protein and recognizes a similar binding sequence. We find that the laboratory strain *C. difficile* 630Δ*erm* contains an 18bp-duplication near the DNA-binding domain compared to its ancestral strain 630. *In vitro* binding assays using purified C-terminal DNA binding domain of the *C. difficile* Spo0A protein demonstrate direct binding to DNA upstream of *spo0A* and *sigH*, early sporulation genes and several other putative targets. *In vitro* binding assays suggest that the gene encoding the major clostridial toxin TcdB may be a direct target of Spo0A, but supernatant derived from a *spo0A* negative strain was no less toxic towards Vero cells than that obtained from a wild type strain, in contrast to previous reports.

These results identify for the first time direct (putative) targets of the Spo0A protein in *C. difficile* and make a positive effect of Spo0A on production of the large clostridial toxins unlikely.

Introduction

Sporulation is an adaptive strategy that enables bacteria to survive harsh environmental conditions for prolonged periods of time, and is an integral part of the transmission of sporulating pathogens and their tolerance and resistance towards antimicrobial compounds.

Spo0A is the key regulator for sporulation (1,2). Most of our knowledge about the protein is based on work in *Bacilli*. Spo0A is a response regulator that demonstrates phosphorylation dependent binding to DNA (3–5). Phosphorylation occurs through the concerted action of several proteins that together form a so called phosphorelay (6). The signaling cascade allows for the integration of environmental signals into the regulation of Spo0A dependent processes, including sporulation. The two functional domains, the N-terminal phosphorylation and dimerization domain (receiver domain), and the C-terminal DNA binding (effector) domain are separated by a hinge region that is relatively poorly conserved (7). Phosphorylation is believed to result in a structural rearrangement that facilitates dimerization (8,9), resulting in the disruption of transcription-inhibitory contacts between the receiver and effector domains. The isolated DNA binding domain can bind legitimate targets of the Spo0A protein due to the absence of the transcription inhibitory contacts, thereby bypassing the need for phosphorylation (10). Extensive characterization of Spo0A targets has revealed a motif that represents a high affinity Spo0A binding site, the 0A box (10,11). The crystal structure of the DNA binding domain confirms specific and non-specific contacts between the protein and the consensus sequence (12,13). It is noteworthy that Spo0A regulates many other processes than sporulation, such as competence for genetic transformation, DNA replication, and biofilm formation in *B. subtilis* (14–16), virulence factors and stress responses in for instance *B. anthracis* and *B. thuringiensis* (17–21), and solvent production in *Clostridium acetobutylicum* (22,23).

C. difficile is a Gram positive, anaerobic bacterium that is the causative agent of *C. difficile* infection (CDI) (for recent reviews see (24,25)). Though many people are asymptomatically colonized by *C. difficile*, the bacterium can cause serious health problems, such as pseudomembranous colitis and toxic megacolon, under the influence of risk factors such as age and antibiotic use. As a result, CDI was long regarded a nosocomial infection. Recently, however, an increase in the cases of community acquired CDI can be observed (26). Outbreaks of CDI have been linked to so called hypervirulent strains, such as PCR ribotypes (RT) 027 (BI/NAP1) and 078 (27,28). Its main virulence factors are the major clostridial Toxins A and B (29,30). In addition, certain strains of *C. difficile*, including RTs 027 and 078, additionally encode a binary toxin (31,32). *C. difficile* is transmitted via the

fecal-oral route. It is believed that spores are crucial to successfully infect new hosts, as they are able to withstand the harsh environment of the stomach, and survive antibiotic treatments that alter the endogenous flora, after which *C. difficile* can overgrow (24,25).

There is limited knowledge about the regulation of sporulation in *C. difficile*. It has been reported that *spo0A,* as expected, is required for the formation of spores (33) and the gene is required for persistence and transmission in mice (34). Though the pathways downstream of Spo0A seem to a large extent conserved between *B. subtilis* and *Clostridia*, this is less so for the pathways leading to activation of Spo0A (2). It has been suggested that the orphan histidine kinase CD2492 is involved in the activation of Spo0A (35). Similarly, it was reported that multiple orphan histidine kinases can phosphorylate Spo0A in *C. acetobutylicum* (36). Recently, it was reported that *spo0A* can be transcribed from a SigH-dependent promoter (37). It is unknown which genes are regulated by direct binding of Spo0A to their upstream regions.

Here, we establish an *in vitro* binding assay for *C. difficile* Spo0A and demonstrate for the first time direct binding of this transcription factor to DNA upstream of several putative target genes.

Material and Methods

Bacterial strains and media

Escherichia coli strains were routinely grown in Luria-Bertani broth or plates, supplemented with appropriate antibiotics. Chloramphenicol was used at a final concentration of 20 μg/mL for agar plates and 10 μg/mL for liquid cultures. Ampicillin was used at a final concentration of 100 μg/mL. Kanamycin was used at a final concentration of 20 μg/mL. Cloning was carried out using *E. coli* DH5*a*, overexpression was performed in E. coli Rosetta(DE3) pLysS (Novagen). C. difficile strains were grown in a glucose-free trypton-yeast based medium (TTY; 3% w/v bacto-trypton (BD), 2% yeast extract (Fluka), 0.1% w/v thioglycollate (Sigma) pH7.4), supplemented with 20 μg/mL of lincomycin when appropriate, or on CLO or TSS plates (Biomerieux).

Plasmid construction

All plasmids are listed in Table 1. Primers (obtained from Sigma Aldrich) are listed in Table 2 and specific cycling conditions are available on request. Unless noted otherwise, PCR reactions were carried out using Pfu polymerase (Fermentas) according to the instructions of the manufacturer.

Plasmid pWKS1251, for the overproduction of Spo0A-DNA Binding Somain (DBD) carrying a C-terminal 6xHis-tag, was constructed as follows. A sequence corresponding to the DBD of Spo0A was amplified using primers oWKS-1123a and oWKS-1124 using chromosomal DNA from *C. difficile* strain 630∠*erm* as a template. The resulting fragment was cloned into pCR2.1-TOPO (Invitrogen), yielding pWKS1247. This plasmid was digested with NdeI and XhoI, separated on a 1% agarose/0.5x TAE (20 mM Tris Acetate, 0.5 mM EDTA) gel, the fragment corresponding to the DNA binding domain was recovered by gel-isolation (using a GeneJET Gel Extraction kit, Fermentas) and cloned into similarly digested pMF14 (10) that had been gel-isolated in the same manner. The construct was verified by PCR, restriction analyses and DNA sequencing using primers oWKS-135 and oWKS-136.

Plasmid pWKS1245, for the production of full length Spo0A carrying a C-terminal 6xHis-tag, was constructed in a similar manner using chromosomal DNA from *C. difficile* 630∠*erm* as a template, but using the PCR product of primers oWKS-1122 and oWKS-1123a. Plasmids used as PCR templates for generating EMSA probes were constructed by cloning the PCR products into pCR2.1-TOPO. The inserts, and in the case of the mutated P*abrB* promoters the presence of the desired point mutations in the consensus 0A box, were verified by DNA sequencing using primers oWKS-24 and oWKS-25.

DNA sequencing

Sequence grade plasmids were isolated using a Nucleospin Plasmid QuickPure kit (Macherey Nagel) according to the manufacturer's instructions, except that two lysis reactions were combined onto a single filter and eluted with 65°C prewarmed AE buffer. All constructs were sequenced using BigDye Terminator chemistry (Invitrogen) on an ABI3130 sequencer (Perkin Elmer), according to the instructions of the manufacturers. In short, ~200 ng of plasmid was mixed with 3.2 pmol of primer, 1 μL Terminator Ready Reaction Mix (Invitrogen) in a final volume of 20 μL. After thermocycling, DNA was precipitated and washed with 65% isopropanol, and dissolved in 12 μL HiDi formamid (Invitrogen) at 96°C for 2 mins and stored in the dark at 4°C until the sequencing run. Sequence analyses were performed in CloneManager Professional Suite 7 (SciEd) and Geneious version 5.6.2 (Biomatters Ltd).

Table 1: Plasmids used in this study.

Name	Description *	Reference
pMF14	pET21b-*spo0A*-DBD-6xHis (Bsu)	(10)
pWKS1245	pET21b-*spo0A*-DBD-6xHis (Cdi)	This study
pWKS1251	pET21b-*spo0A*-6xHis (Cdi)	This study
pWKS1303	pCR2.1TOPO-P*abrB* (Bsu)	This study
pWKS1304	pCR2.1TOPO-P*abrB* (C4T)	This study
pWKS1305	pCR2.1TOPO-P*abrB* (C4T/G5T)	This study
pWKS1306	pCR2.1TOPO-P*abrB* (G2T)	This study
pWKS1307	pCR2.1TOPO-P*abrB* (G2T/C4T)	This study
pWKS1308	pCR2.1TOPO-P*abrB* (G5T)	This study
pWKS1309	pCR2.1TOPO-P*citG* (Bsu)	This study
pWKS1316	pCR2.1TOPO-P*spo0A* (Cdi)	This study
pWKS1317	pCR2.1TOPO-P*spoIIAA* (Cdi)	This study
pWKS1318	pCR2.1TOPO-P*spoVG* (Cdi)	This study
pWKS1322	pCR2.1TOPO-P*sigH* (Cdi)	This study
pWKS1323	pCR2.1TOPO-P*spoIIGA* (Cdi)	This study
pWKS1324	pCR2.1TOPO-P*spoIIE* (Cdi)	This study
pWKS1326	pCR2.1TOPO-P*tcdB* (Cdi)	This study
pWKS1329	pCR2.1TOPO-P*tcdA* (Cdi)	This study
pWKS1332	pCR2.1TOPO-P*lplA* (Cdi)	This study
pWKS1340	pCR2.1TOPO-P*tcd*R (Cdi)	This study
pWKS1341	pCR2.1TOPO-P*ssuA* (Cdi)	This study

* Bsu = DNA derived from *B. subtilis* JH642. Cdi = DNA derived from *C. difficile* 630Δ*erm*.

Protein purification

Plasmids pWKS1245 and pWKS1251 were transformed into *E. coli* Rosetta(DE3) pLysS (Novagen). Transformants were used to inoculate 25 mL of LB with appropriate antibiotics. After overnight incubation, the cells were 1:100 diluted in 500 mL fresh medium containing appropriate antibiotics. Protein production was induced with 1 mM IPTG at an OD_{600} of 0.7 and growth was continued for another three hours before harvesting. Cells were washed with ice cold PBS and stored at -80°C for later use. Purification of the proteins was essentially done as described (10). In short, cells were disrupted in 4 mL lysis buffer (2mM PMSF, 10 mM imidazole, 5 mM beta-mercaptoethanol, 300 mM NaCl, 50 mM NaH_2PO_4, pH7.9). Cleared cell lysates we incubated with 2 mL pre-equilibrated 50% TALON slurry (Clontech) in a final volume of 15 mL lysis buffer for 1hr. The resin was allowed to settle on a Poly-Prep column (BioRad) and washed with 2 mL wash

Table 2: Oligonucleotides used in this study.

Primer	Sequence (5'- 3')	Fragment
oWKS-24	GTGAGCGGATAACAATTTCACACAGG	pCR2.1TOPO insert
oWKS-25	GGTTTTCCCAGTCACGACGTTGTAA	pCR2.1TOPO insert
oWKS-135	GCGAAATTAATACGACTCACTATAGG	pET21b(+) insert
oWKS-136	CAGCCAACTCAGCTTCCTTTC	pET21b(+) insert
oWKS-1223a	TGCTCGAGTTTAACCATACTATGTTCTAGT	C. difficile spo0A-DBD
oWKS-1124	CTCATATGACTAGAAGTGATTTTGTTAAG	C. difficile spo0A and spo0A-DBD
oWKS-1122	TTTCATATGGGGGGATTTTTAGTGG	C. difficile spo0A
oWKS-1193	AGGATTTTTAGTAGGATAATAGC	*B. subtilis Pab*rB*
oWKS-1194	CTCCCAAGAGATACTTATTTG	*B. subtilis PabrB
oWKS-1223	CTCCCAAGAGATACTTATTTGTTTAAATTATATTT TTCTTCGTCATTATTCAACAAAATC	*B. subtilis PabrB -C4A
oWKS-1224	CTCCCAAGAGATACTTATTTGTTTAAATTATATTT TTCTTCGTCATTATTAAACAAAATC	*B. subtilis PabrB -C4A-G5A
oWKS-1225	CTCCCAAGAGATACTTATTTGTTTAAATTATATTT TTCTTCGTCATTATTCGAAAAAATC	*B. subtilis PabrB -G2A
oWKS-1226	CTCCCAAGAGATACTTATTTGTTTAAATTATATTT TTCTTCGTCATTATTCAAAAAAATC	*B. subtilis PabrB -G2A-C4A
oWKS-1230	CTCCCAAGAGATACTTATTTGTTTAAATTATATTT TTCTTCGTCATTATTAGACAAAATC	*B. subtilis PabrB -G5A
oWKS-1199	TGAAAGCGGAGGATACGAAG	*B. subtilis PcitG
oWKS-1200	CAATTCTGTATTCCATTTATGTATC	*B. subtilis PcitG
oWKS-1166	ACACAGGAGGTATCGTACAG	*C. difficile Pspo0A
oWKS-1167	CCCCCCATTAAAAAACATCTTC	*C. difficile Pspo0A
oWKS-1227	GGTGCTACTAGAATGTCTGC	*C. difficile PsigH
oWKS-1228	GACAATTGTCCACTAACTCATAAC	*C. difficile PsigH
oWKS-1231	GCCTCTACCCTTCTATTCAC	*C. difficile PplA
oWKS-1232	TAACATTAAAGTAGTTCCTCCATAG	*C. difficile PplA
oWKS-1235	GCTGATGGACGAACCATTTG	*C. difficile PssuA
oWKS-1236	CAATTAAAACTCCCCCAATTCCC	*C. difficile PssuA
oWKS-1166	ACACAGGAGGTATCGTACAG	*C. difficile Pspo0A
oWKS-1167	CCCCCCATTAAAAAACATCTTC	*C. difficile Pspo0A
oWKS-1277	GGTGCTACTAGAATGTCTGC	*C. difficile PsigH
oWKS-1228	GACAATTGTCCACTAACTCATAAC	*C. difficile PsigH
oWKS-1231	GCCTCTACCCTTCTATTCAC	*C. difficile PplA
oWKS-1232	TAACATTAAAGTAGTTCCTCCATAG	*C. difficile PplA
oWKS-1235	GCTGATGGACGAACCATTTG	*C. difficile PssuA
oWKS-1236	CAATTAAAACTCCCCCAATTCCC	*C. difficile PssuA
oWKS-1170	GGTTGGTGGAAGAAACATGG	*C. difficile PspoIIAA
oWKS-1171	AATCCCTCCTTCAATAGTTTTG	*C. difficile PspoIIAA
oWKS-1172	AAACGTGCGGATATTATAACTG	*C. difficile PspoIIE
oWKS-1173	ACAACTGCTACACTTCTTTG	*C. difficile PspoIIE
oWKS-1174	ACTTCTACCTAGATATATTGGATTC	*C. difficile PspoIIGA
oWKS-1175	ACTCAATATACACCATGTATCC	*C. difficile PspoIIGA
oWKS-1168	GGATATCCCTAGTTGTTCATAG	*C. difficile PspoVG

buffer (20 mM imidazole, 300 mM NaCl, 50 mM NaH_2PO_4, pH7.9). The protein was stepwise eluted in 1 mL fractions after applying 2 mL elution buffer to the column (identical to wash buffer but with 50, 100, 250 or 500 mM imidazole). The whole procedure was carried out at 4°C. Fractions were assayed for purity and yield and suitable fractions were dialysed against 2x 1L dialysis buffer (50 mM Tris-HCl pH8, 1 mM EDTA, 0.5 mM DTT) using Slide-A-Lyzer cassettes with a molecular weight cut-off of 3.5kDa (Pierce).

Proteins were stored at -80°C in storage buffer (identical to dialysis buffer but containing 20% glycerol). Protein concentrations were determined using Bradford reagent (BioRad), according to the manufacturer's instructions.

Electrophoretic mobility shift assays

DNA fragments for use in EMSA experiment were generated by PCR using GoTaq polymerase (Promega) and chromosomal DNA from *B. subtilis* JH642 (Bacillus Genetic Stock Center 1A96; http://www.bgsc.org), plasmids listed in Table 1, or chromosomal DNA from *C. difficile* 630Δ*erm* (38) as a template. Primers for generation of the EMSA probes are listed in Table 2. DNA fragments of the expected size were isolated from a 1xTAE / 8% native polyacrylamide gel using diffusion buffer (0.5 M ammonium acetate, 10 mM magnesium acetate, 1 mM EDTA pH 8, 0.1% SDS) and a QIAExII kit (Qiagen), according to the manufacturer's instructions. Recovered DNA was end-labeled with 32P-γ-ATP using FR buffer and T4 kinase (Invitrogen) according to the instructions of the manufacturer. Specific activity was determined on a LS6000 scintillation counter (Beckman).

EMSA conditions were based on previous studies (10). In short, binding reactions were carried out in binding buffer (10 mM Tris-HCl pH7.6, 1 mM EDTA, 50 mM NaCl, 1 mM DTT, 5% glycerol) in the presence of 200 μg/mL bovine serum albumin (NEB) and 200 cpm/μL radiolabeled DNA fragment. Reactions were incubated for 20 minutes at 30 °C prior to loading on a 1xTAE / 8% non-denaturing polyacrylamide gel that was prerun for 20 minutes at 50V in 1x TAE buffer. Electrophoresis was carried out for 120 min at 85V. After vacuum drying the gels onto filter paper, they were imaged after overnight exposure on Phosphorimager screens on a Typhoon instrument (GE Healthcare).

Cytotoxicity assay

The toxic effects of *C. difficile* culture supernatants on Vero cells (39) were determined as follows. Supernatant from a bacterial culture was harvested by centrifuging cells for 3 minutes at 14000x g and filtered on a 0.45 μM cellulose acetate filter using a syringe. Supernatants were 2-fold serially diluted in cell culture

medium (Dulbecco modified Eagle medium (Lonza) supplemented with 100 μg/mL penicillin, 100 U/mL streptomycin, 10% fetal calf serum), before applying them to a monolayer of Vero cells, and incubation was continued for another hour. As a positive control, 50 μL 1:10 diluted purified toxin (Techlab) was added to the cells. To determine if observed cytotoxic effects were specific for the large clostridial toxins, commercially available anti-toxin against TcdA and TcdB (Techlab) was added to 10-fold diluted bacterial supernatant for 60 min prior to incubation on the Vero cells. Toxin end-point titres were defined as the lowest dilution at which no cytopathological effects (cell rounding) were observed. Statistical significance was evaluated with an independent sample t-test.

Generation of antibodies against Spo0A and immunoblot detection

Immunization of mice with full length *C. difficile* Spo0A-6xHis was kindly performed at the Welcome Trust Sanger Institute (Hinxton, UK). Cells from 1 mL of C. difficile culture were collected by centrifugation for 1 min at 14000 rpm in a table top centrifuge and resuspended in 200 μL resuspension buffer (10 mM Tris HCl pH8, 10 mM EDTA, 0.5 mg/mL lysozyme, 1 mM Pefabloc SC (Roche)). After incubation for 30 mins at 37 °C, 50 μL of 5x SDS sample buffer (0.1 M DTT, 2% SDS, 50 mM Tris HCl pH 6.8, 10% glycerol, 0.0025% BPB) was added, and samples were heated to 96°C for 5 mins. Total cell lysates (amounts corrected for OD_{600}) were separated on a 12% SDS-PAGE gel prior to semi-dry blotting for 1h at 10V to a polyvinylidene fluoride (PVDF) membrane. Membranes were blocked in PBST buffer (phosphate buffered saline with 0.1% v/v Tween-20) containing 5% membrane blocking reagent (Amersham Biosciences). To visualize Spo0A protein cleared polyclonal serum from a single mouse at a 1:3000 dilution was used, followed by either a goat-anti-mouse HRP-conjugated secondary antibody followed by ECL+ detection (Amersham Bioscience), or a goat-anti-mouse-biotin-conjugated secondary antibody (Dako) followed by a tertiary mouse-anti-biotin Cy3-conjugated antibody (Jackson). Detection was done using on a Typhoon instrument (GE Healthcare). Background corrected peak volumes were quantified using ImageQuant TL (Amersham Biosciences).

Bioinformatics and software

Alignments of *B. subtilis* and *C. difficile spo0A* were made using ClustalW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2/) on the basis of the published genome sequences, Genbank accession numbers AL009126 and AM180355, respectively, and the 630Δ*erm spo0A* sequence as determined in this study. The sequence for *spo0A* of *C. difficile* strain 630Δ*erm* was deposited in Genbank (accession no *JX050222*). Consensus Spo0A boxes were identified using a Single string Search

command in Genome2D (40), allowing 0 mismatches. The box positions were linked to up- and downstream genes using the "Add nearest gene to List of DNA Motifs" feature and Microsoft Excel. The results were manually inspected for those boxes within 500 bp upstream of a gene on the same strand. Figures for publication were prepared using ImageQuant TL (Amersham Biosciences), Adobe Photoshop CS3 (Adobe Systems Inc) and Corel Graphics Suite X5 (Corel Corporation).

Results

Purification of Spo0A and its DNA binding domain

In order to characterize *C. difficile* Spo0A, the full length protein and its DNA binding domain (DBD) were expressed as a C-terminally 6xHis-tagged protein in the heterologous host *Escherichia coli* (Figure 1A) and purified to near homogeneity using metal affinity chromatography (Figure 1A; lanes P). Full length protein was used to raise antibodies to detect Spo0A in total lysates of *C. difficile* strains, and the purified DNA binding domain was used in subsequent *in vitro* binding assays (see below).

Spo0A is expressed throughout growth

We determined the expression of *C. difficile* Spo0A throughout growth. We found that the protein is present in lysates from exponential to stationary growth phase cells. We performed immunoblotting using polyclonal antibodies against *C. difficile* Spo0A on total lysates of wild type and spo0A mutant cells grown in a tryptone-yeast based medium (TTY). We found a clear signal of the size expected for full length Spo0A (~31 kDa) as early as 3 hours post inoculation (exponential growth phase), through transition phase (8h) as well as 24 and 48 hours post inoculation (stationary growth phase) (Figure 1B; 630Δ*erm*). The signals were specific for *C. difficile* Spo0A as they were absent from lysates from the *C. difficile spo0A* mutant (Figure 1B, CT::spo0A). We obtained similar results in other media, such as the commonly used supplemented brain heart infusion broth (BHIS; data not shown).

To determine relative levels of Spo0A throughout growth, we performed an immunoblot experiment using fluorescent antibodies, which gives more quantitative information compared to the use of horseradish peroxidase conjugated antibodies in our hands. We found that the levels of Spo0A increases approximately 20-fold from 6 hours post inoculation and remains at similar levels from 8 to 48 hours post inoculation (Figure 1C).

Though it should be noted that the Western blots do not provide information on the phosphorylation state of the protein, we conclude that the protein in active

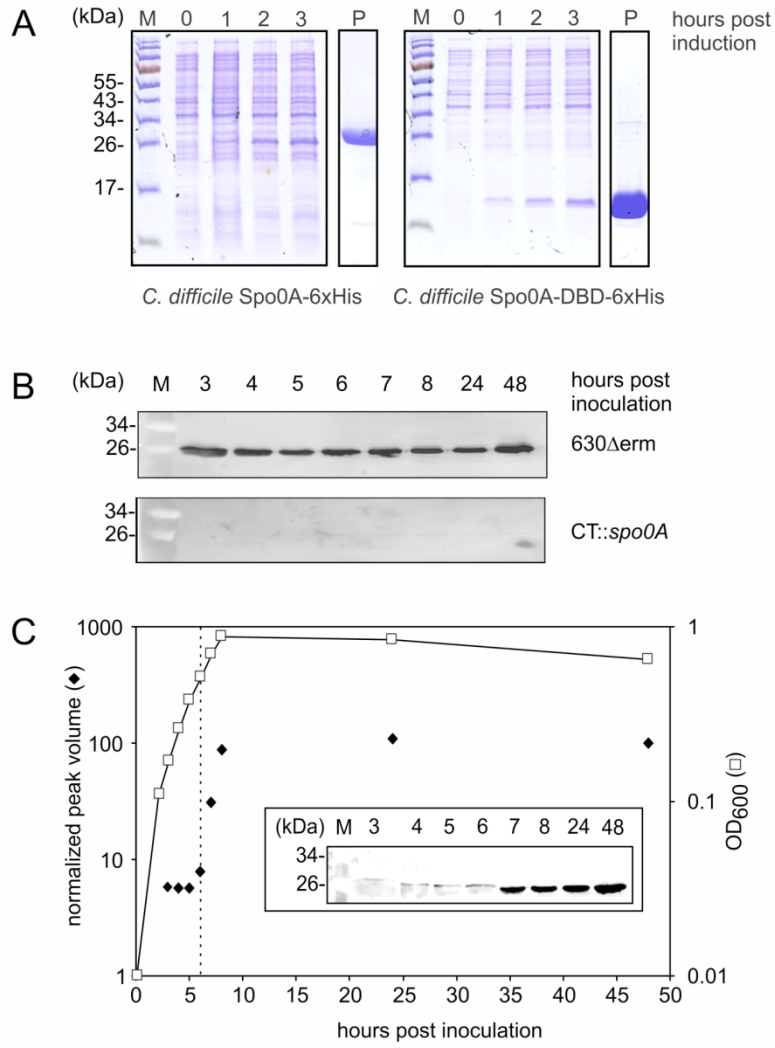


Figure 1: Purification and detection of *C. difficile* Spo0A. **(A)** Heterologous overproduction of Spo0A-6xHis and Spo0A-DBD-6xHis in *E. coli* Rosetta(DE3) pLysS. M = molecular weight marker, numbers indicate hours after induction with 1mM IPTG. P = metal affinity purified protein. Lysates were separated on a 12% SDS-PAGE. **(B)** *and* **(C).** Immunoblot detection of Spo0A in total cell lysates of a *C. difficile spo0A* mutant (CT::*spo0A*) and a wild type strain (630Δ*erm*). Times indicated range from early exponential (3h post inoculation) to late stationary growth phase (48h post inoculation). Sample volumes were corrected for OD_{600} to ensure loading of similar amounts of total cell lysate in each lane. For details see Materials and Methods. M = molecular weight marker. **(B)** ECL+ detection. **(C)** Fluorescent detection. Y-axes show peak volumes normalized to values at 48 hours post inoculation (closed diamonds; left axis) and optical density readings at 600nM (open squares; right axis). Inset shows the blot on which the curve is based. Vertical dashed line indicates the moment Spo0A levels increase sharply (6 hours post inoculation).

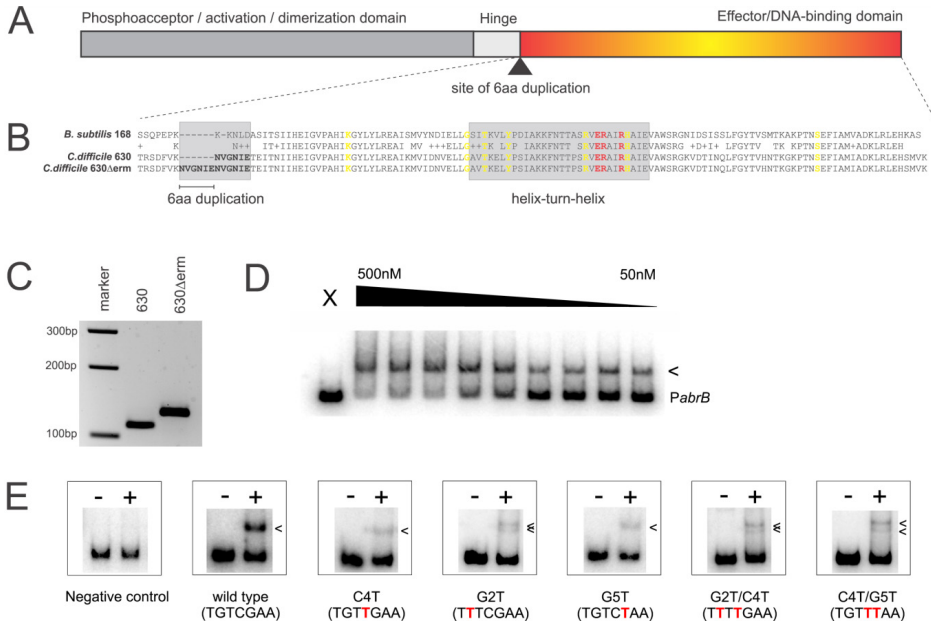


Figure 2: (A) Spo0A box is important for high affinity binding by *C. difficile* Spo0A. **(A)** Domain organization of Spo0A. The site of the duplication in strain 630Δ*erm* identified in this study is indicated by an arrow. **(B)** Sequence alignment of the C-terminal regions of the Spo0A proteins of *B. subtilis* 168 and *C. difficile* strains 630 and 630Δ*erm*. Residues identified in structural studies using Bacillus Spo0A as involved in backbone interactions are indicated in yellow, residues forming base-specific contacts are indicated in red (13). The region of the 6aa duplication and the helix-turn-helix motifs are boxed in gray and the duplication in strain 630Δ*erm* is underlined. **(C)** PCR showing the presence of the duplication near the DNA binding domain in *C difficile* 630Δ*erm* compared to 630. **(D)** Electrophoretic mobility shift assay using purified *C. difficile* Spo0A-DBD-6xHis and a radiolabeled P*abrB* DNA fragment. X = no protein control, the triangle indicates 1.3-fold serial dilutions of protein to the indicated concentrations. The arrow indicates a DNA:protein complex. **(E)** Electrophoretic mobility shift assays without (-) or with (+) 150 nM Spo0A-DBD-6xHis added to radiolabeled P*abrB* fragments carrying mutations in the consensus Spo0A box (in red). Arrows indicate DNA:protein complexes. The negative control is P*citG* from *B. subtilis*.

or inactive form is present throughout growth and is more abundant in stationary growth phase.

Spo0A of *C. difficile* strain 630Δ*erm* contains a 6-aminoacid duplication

BLAST homology searches readily identify a homolog of the well-characterized *B. subtilis* Spo0A protein in *C. difficile* 630 (CD1214) and previous work demonstrated that a spo0A mutant (an insertional inactivation of cd1214) – as expected – no longer forms spores (41). In silico analyses suggest a similar secondary structure for both proteins (Figure 2A), with a conserved dimerization and DNA binding domain, separated by a poorly conserved hinge region (7,12).We compared the sequence of CD1214 obtained from our lab strain 630⊿*erm* (38) to that of the published *C. difficile*630 genome (42). Strain 630⊿*erm* is a spontaneous erythromycin sensitive strain, which is commonly used in mutagenesis studies and was obtained by serial passaging of strain 630 (33,38). The 630⊿*erm spo0A* sequence (Genbank accession no *JX050222*) was derived from the expression plasmids constructed for this study, and confirmed in a whole genome sequence of strain 630⊿*erm* generated in our lab (data not shown). We found that 630⊿*erm* spo0A contains an 18 base pair direct repeat,resulting in a 6 amino acid (NVGNIE) duplication compared to the published reference sequence. The duplication maps to a region of the protein with relatively low sequence conservation (hinge), flanking the highly conserved DNA binding domain (Figure 2A and B). We verified the absence of this duplication in strain 630 by PCR (Figure. 2C) as well as sequencing from the chromosomal DNA of *C. difficile* 630 (data not shown), to rule out an error in the original genome sequence and to demonstrate that the difference in size of the PCR product was specific to the 18bp insertion. In addition, we checked several other strains of PCR ribotypes 12 (to which 630 and 630⊿*erm* belong) by PCR, but the duplication was found to be unique to 630⊿*erm* among the isolates tested (data not shown).

C. difficile Spo0A-DBD shows similar specificity as _B. subtilis_ Spo0A-DBD

Next, we examined the conservation of the Spo0A-DBD between *B. subtilis* and *C. difficile*. In *B. subtilis* amino acid residues contacting the backbone of the DNA and interacting with specific residues of the Spo0A binding sequence have been defined (13). We found that all these residues were conserved in the *C. difficile* protein sequence (Figure 2B), indicating that the protein likely recognizes a similar motif.

C. difficile Spo0A-DBD shows similar specificity as _B. subtilis_ Spo0A-DBD

Next, we examined the conservation of the DNA binding domain of Spo0A (Spo0A-DBD) between *B. subtilis* and *C. difficile*. In *B. subtilis* amino acid residues

contacting the backbone of the DNA and interacting with specific residues of the Spo0A binding sequence have been defined (13). We found that all these residues were conserved in the *C. difficile* protein sequence (Figure 2B), indicating that the protein likely recognizes a similar motif.

DNA binding by full length Spo0A in *B. subtilis* requires phosphorylation dependent dimerization (8,9). However, it was shown that the isolated DBD is capable of binding to legitimate targets of the full length protein (10). Analogously, we purified the *C. difficile* Spo0A-DBD for use in *in vitro* binding assays. As no direct targets for the *C. difficile* protein have been reported so far, we used the upstream region of the *abrB* gene (P*abrB*) of *B. subtilis*. P*abrB* is commonly used as a high-affinity control in binding assays with the *B. subtilis* Spo0A or Spo0A-DBD protein (43,44). It is noteworthy that we failed to identify a homolog of *abrB* in *C. difficile* using BLAST, indicating that potential indirect regulation by Spo0A cannot occur through *abrB* in *C. difficile* as it does in *B. subtilis*. We found that *C. difficile* Spo0A-DBD bound with high affinity to P*abrB* (Figure 2D and E). We performed electrophoretic mobility shift assays (EMSAs) using radiolabeled P*abrB* and increasing amounts of purified *C. difficile* Spo0A-DBD that was purified using a C-terminal 6xHis-tag. The addition of protein leads to a dose-dependent retardation of the DNA fragment with an apparent K_D of <50nM. In the same range of protein concentrations, no binding was observed for a negative control (a DNA fragment of *B. subtilis citG* (*45*)) (Figure 2E), suggesting that binding was specific for the *abrB* promoter region. *B. subtilis* Spo0A recognizes a distinct sequence (0A box), that is characterized by a 7bp core motif (TGTCGAA) (10,11). Structural studies have revealed that the protein makes specific contacts with the G at position 2 (G2), and the C at position 4 (C4) and 5 (G5) of this motif (13). We introduced G2A, C4A, G5A, G2A/C4A and C4A/G5A mutations in the perfect consensus core 0A-box present in P*abrB*. We found that the affinity of *C. difficile* Spo0A for these mutated P*abrB* fragments was highly reduced (Figure 2E). We performed EMSAs using radiolabeled P*abrB* containing the mutated core sequence. For the single point mutations in the DNA, the affinity decreased ~10-fold. There did not seem to be an additive effect of a second point mutation for the two combinations tested. None of the mutations abolished binding of *C. difficile* Spo0A completely, most likely as the result of binding of Spo0A to other (non-consensus) 0A boxes in the *abrB* promoter (44).

Taken together, we conclude that the guanine and cytosine residues in the core TGTCGAA motif of P*abrB* are important for specific binding of this fragment by *C. difficile* Spo0A-DBD.

The presence of a consensus Spo0A box has predictive value for binding by *C. difficile* Spo0A-DBD

Above, we have established that the Spo0A-DBD of *C. difficile* is highly homologous to that of the *B. subtilis* Spo0A protein, and that the proteins recognize a similar consensus sequence (Figure. 2). Based on this information, we identified the several genes as putative direct targets of *C. difficile* Spo0A. We queried the *C. difficile* 630 genome sequence for perfect matches to the core 0A box using Genome2D (40). Such an analysis revealed the presence of 102 matching motifs, of which 45 were located within 500bp of the initiating ATG of an open reading frame on the same strand (data not shown). Our attention was drawn to *spo0A* and *sigH*, as these two genes were previously found to be regulated by Spo0A in *B. subtilis* and/or play important roles in sporulation (3,46–48). We found that *C. difficile* Spo0A bound to DNA sequences upstream of *spo0A* and *sigH*.

We performed EMSAs with DNA encompassing 220-281bp upstream of the initiating ATG codon of the *spo0A*, *sigH* and *spoVG* open reading frames. We found that the addition of Spo0A-DBD to the reactions caused retardation of the *spo0A* and *sigH* DNA fragments (Figure 3A), but not of a *spoVG* fragment which did not contain a consensus 0A box (Figure 3B). It should be noted that the affinity of Spo0A-DBD for the region upstream of *spo0A* was the highest we have

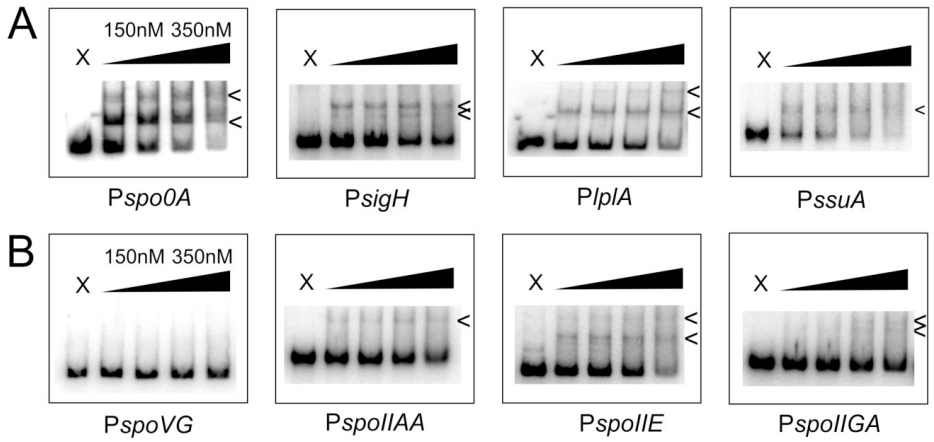


Figure 3: *C. difficile* Spo0A binds to predicted and expected target sequences. Electrophoretic mobility shift assay using purified *C. difficile* Spo0A-DBD-6xHis and a radiolabeled DNA fragments. X = no protein control, the triangle indicates 1.3-fold serial dilutions of protein to the indicated concentrations. The arrows indicate DNA:protein complexes. **(A)** *In silico* predicted target sequences upstream of the genes encoding Spo0A (*spo0A*), σH (*sigH*), a lipoate ligase (*lplA*) and an aliphatic sulphonates ABC transporter (*ssuA*). **(B)** Target sequences predicted on the basis of findings in other organisms: *spoIIAA*, *spoIIE* and *spoIIGA*. The DNA upstream of *spoVG* serves as a negative control.

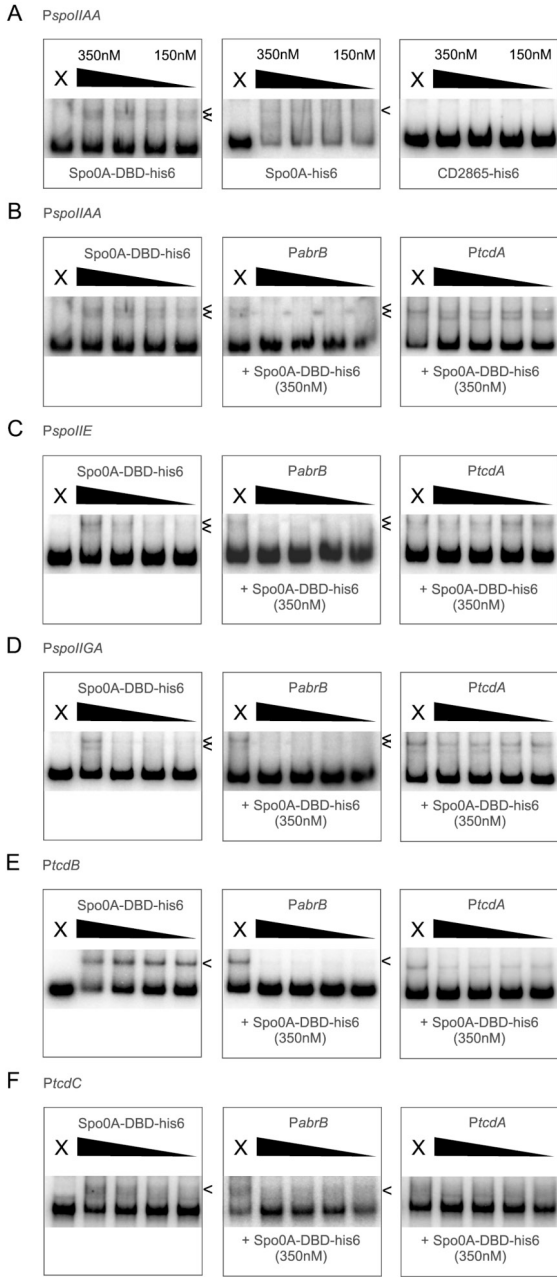


Figure 4: Specificity controls for binding by Spo0A-DBD-his6. Arrows indicate the position of shifted species (DNA:protein complexes). Titrations with PCR fragments of P*abrB* (containing a high affinity binding site) and P*tcdA* (lacking such a site) correspond to approximately 0.1nM/µL - 0.03 nM/µL

(A) Comparison of binding of Spo0A-DBD-his6, Spo0A-his6 and CD2195-his6 binding to the upstream region of *spoIIAA*.

(B) Binding of Spo0A-DBD-his6 to the upstream region of *spoIIAA* is reversed by the addition of P*abrB*, but not by the addition of P*tcdA*).

(C) Binding of Spo0A-DBD-his6 to the upstream region of *spoIIE* is reversed by the addition of P*abrB*, but not by the addition of P*tcdA*.

(D) Binding of Spo0A-DBD-his6 to the upstream region of *spoIIGA* is reversed by the addition of P*abrB*, but not by the addition of P*tcdA*.

(E) Binding of Spo0A-DBD-his6 to the upstream region of *tcdB* is reversed by the addition of P*abrB*, but not by the addition of P*tcdA*.

(F) Binding of Spo0A-DBD-his6 to the upstream region of *tcdC* is not or moderately affected by the addition of P*abrB* and/or P*tcdA*.

observed so far for any *C. difficile* DNA. Moreover, the presence of multiple shifted species could indicate the presence of more than one strong binding site. These results establish that *spo0A* and *sigH* are likely legitimate targets of Spo0A in *C. difficile*, and confirm that *spoVG* is not, in line with results obtained in *B. subtilis* (10).

We were interested to see if Spo0A in *C. difficile* could potentially regulate genes that have no documented function in sporulation. Our *in silico* analysis identified several genes with no obvious link to sporulation that had a consensus 0A box within 100bp upstream of their start codon. This positioning is similar to that observed for *spo0A* (-75) and *sigH* (-78). We confirmed *in vitro* binding of the *C. difficile* Spo0A-DBD to the promoter regions of *lplA* and *ssuA*.

We carried out EMSA experiments using probes that included the perfect consensus site and purified Spo0A-DBD protein. We observed binding of the protein to fragments upstream of the *lplA* gene (CD1654; box at -67) and the *ssuA* gene (CD1484; box at -82) (Figure 3A). The *lplA* gene encodes a predicted lipoate-protein ligase, and *ssuA* is annotated as an aliphatic sulfonates ABC transporter; to our knowledge, neither of these have been directly implicated in sporulation or have found to be targets for Spo0A in other organisms.

Together our results establish the potential for binding of Spo0A to DNA upstream of *spo0A* and *sigH*, two genes that are important for sporulation, and indicate that Spo0A may have functions that go beyond the regulation of sporulation in *C. difficile*.

C. difficile Spo0A-DBD binds to early sporulation promoters

It has been established that a *spo0A* mutant of *C. difficile* does not produce any spores, consistent with a crucial role in the sporulation pathway (33). However, the *in silico* identification of upstream regions with a consensus Spo0A binding site did not point to any of the early sporulation genes (downstream of *spo0A* itself) as direct targets of Spo0A. This is likely the result of variations in the 0A-box in these promoters that were disregarded in the box search. In support of this, many well-characterized legitimate direct targets of *B. subtilis* Spo0A (such as *spoIIAA* and *spoIIE*) do a 100% match to the core motif, but rather one or more near-consensus boxes (5,49). We found that Spo0A-DBD of *C. difficile* is capable of binding the DNA upstream of the open reading frames of *spoIIAA, spoIIE* and *spoIIGA* with low affinity.

We performed EMSA experiments using increasing amounts of purified Spo0A-DBD from *C. difficile* 630Δ*erm* and the DNA fragments indicated above (Figure 3B). For *spoIIAA* (encoding an anti-anti sigma-factor) and *spoIIE* (encoding a serine phosphatase), we observed a low intensity shifted species at concentrations as low as 150 nM. For *spoIIGA* (encoding a sporulation specific protease) we

observed the shifted species only at higher concentrations of protein (>200nM). The negative control (*spoVG*) did not demonstrate binding of Spo0A-DBD at these concentrations. Moreover, the. shift we observed was reversible using unlabeled DNA containing a high affinity binding site, but not using unlabeled DNA that lacked such a site not contain (Figure 4 B-D). Therefore, we consider the binding to *spoIIAA*, *spoIIE* and *spoIIGA* genes to be specific, despite the fact that increasing the amount of protein did not seem to cause a significant increase in the amount of DNA in the complex. Together, these results suggest that Spo0A in *C. difficile* might regulate the transcription of at least a subset of early sporulation genes by direct binding to their promoter regions.

C. difficile Spo0A-DBD binds to DNA upstream of *tcdB*

It has previously been reported that the deletion of Spo0A in *C. difficile* results in a significantly lower toxin production and a ~1000-fold reduction in the toxicity of culture supernatant derived from *spo0A* negative cells towards Vero cells (35). Considering the absence of a homolog of the *abrB* repressor, direct binding of Spo0A and concomitant activation of toxin gene transcription is a likely mechanism through which this could occur. We found evidence for direct binding of Spo0A-DBD to the region upstream of *tcdB*, encoding one of the major clostridial toxin genes, and possibly *tcdC*, but this did not seem to result in lower toxin levels in our hands.

We performed EMSAs using DNA upstream of *tcdR* (encoding a sigma factor responsible for the activation of toxin gene transcription), *tcdB* (encoding toxin B), *tcdA* (encoding toxin A). In order to test regions upstream of all open reading frames in the PaLoc, we also tested binding of Spo0A to DNA upstream of *tcdE* (encoding a holin-like protein (50,51)) and *tcdC* (encoding a putative negative regulator of toxin production (52–54)), even though this regulator does not have a significant effect on toxin levels under the conditions we used (55,56). Of the regions tested, we only observed a clear shifted species, indicative of Spo0A binding, for *tcdB* (Figure 5A); the shifted species in our EMSA assay was reversed by the addition of unlabeled DNA containing a high affinity binding site, but not by DNA lacking such a site (Figure 4E).

For *tcdC,* some smearing was observed at all concentrations of proteins tested (Figure 5A), and there did not seem to be a clear effect of the addition of unlabeled DNA fragments (Figure 4F). The probes for *tcdA*, *tcdE* and *tcdR* were indistinguishable from those obtained with our negative control, *spoVG*.

We wanted to determine if toxin levels in culture supernatants were directly or indirectly affected by Spo0A, as was previously suggested. We found no lower toxicity towards Vero cells of culture supernatants derived from spo0A mutant cells compared to wild type.

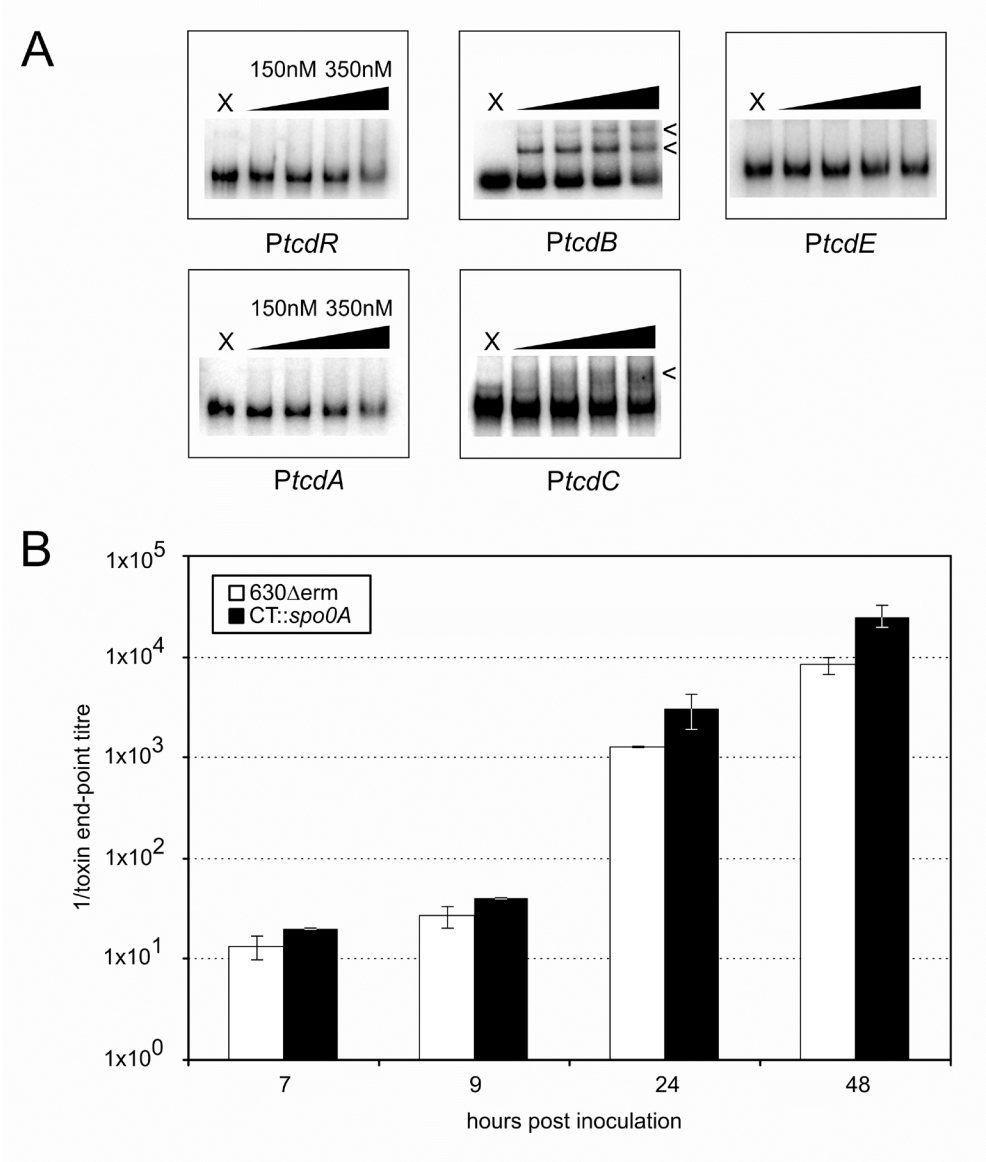


Figure 5: *C. difficile* Spo0A does not contribute to toxin production directly. **(A)** Electrophoretic mobility shift assays using purified *C. difficile* Spo0A-DBD-his6 and radiolabeled DNA fragments corresponding to the upstream regions of the indicated toxin-related genes. X = no protein control, the triangle indicates 1.3-fold serial dilutions of protein to the indicated concentrations. Arrows indicate (possible) protein:DNA complexes. **(B)** Toxin end-point titre of culture supernatant from wild type (630Δ*erm*; white bars) or *spo0A* mutant (CT::*spo0A*; black bars) *C. difficile* cells. For details see Materials and Methods. Error bars represent standard error of the mean (n=3). None of the observed differences were significant in an independent sample t-test.

We grew three independent biological replicates of a wild type (630Δ*erm*) or Clostron-generated *spo0A* mutant (CT::*spo0A* - a kind gift of the Minton lab) in glucose-free TTY medium. We harvested culture supernatant at late-exponential phase (approximately 7 hours post inoculation), the transition phase between exponential and stationary growth phase (approximately 9 hours post inoculation), as well as two time points in stationary phase (24 and 48 hours post inoculation) and determined the toxin endpoint titres (see Materials and Methods). In contrast to previous findings, we observed a small (≤4-fold) increase in the toxicity of supernatants derived from spo0A mutant cells compared to wild type, but in all cases this difference was not statistically significant (p>0.05, independent sample t-test). In other medium (BHIS), we observed no differences at all (data not shown). We conclude that Spo0A does not positively affect toxin production in *C. difficile* 630Δ*erm* and the in vivo relevance of the binding to regions upstream of *tcdB* and/or *tcdC* is therefore limited under our experimental conditions.

5

Discussion

The Spo0A-box of *C. difficile*

In *B. subtilis*, the binding site of Spo0A on target DNA has been well-characterized, through a combination of in vitro binding assays, determination of in vivo binding profiles and mutagenesis of regulated promoter sequences. This work has led to the identification of a conserved core motif, TGTCGAA, or Spo0A box (5,10,11,45). Depending on the analysis, this motif is flanked by one or more adenine or thymine residues (10,11). Interestingly, many target genes do not harbor a perfect match to this consensus sequence, but rather contain one or more degenerate motifs. The differences in these motifs may reflect different promoter architectures (e.g. AT content), modes of action (e.g. activation or repression) or levels of regulation. Spo0A genes in *B. subtilis* can be divided in different classes that respond to different levels of phosphorylated Spo0A (43,57).

For *C. difficile*, we conclude that the Spo0A protein likely recognizes a motif that is similar to the *B. subtilis* Spo0A box on the basis of four lines of evidence; 1. All DNA binding/contacting residues are conserved (Figure 2B), 2. *C. difficile* Spo0A can bind with high affinity to a target of *B. subtilis* Spo0A (Figure 2D), 3. Mutagenesis of key residues in the *B. subtilis* Spo0A box reduces affinity of *C. difficile* Spo0A for DNA (Figure 2E) and 4. A *B. subtilis* Spo0A box has predictive value for DNA binding by *C. difficile* Spo0A (Figure 3A). It is conceivable that our model system, using the purified DNA binding domain, does not accurately reflect binding to all target sites, if target site selectivity is determined in part by other parts by of

the full length protein. It is likely that differences do exist between the preferred binding sites for both proteins that will be evident when a comprehensive analysis is performed of in vivo DNA binding of *C. difficile* Spo0A; based on the limited data set of this study, a MEME analysis (58) already suggests possible differences in the extended Spo0A motif (W.K. Smits, unpublished observations). These differences may relate to the much higher AT content of C. difficile compared to B. subtilis (71 vs. 56.5 %, respectively), or phosphorylation dependent dimerization, for instance.

Regulation of sporulation; feedback regulation of *spo0A*

The initiation of sporulation in *B. subtilis* is subject to complex regulation (for review see ref (1,59)). The activation of Spo0A is controlled by a multi-component phosphorelay that can integrate environmental cues (60) and ensures a gradual increase in the level of phosphorylated Spo0A in the cell (57). In addition, the transcription of the spo0A gene is controlled by multiple feedback loops. For instance, Spo0A regulates its own transcription by binding to the spo0A promoter (46), as well as by indirectly stimulating the transcription of *sigH*, encoding a sigma factor that recognizes the *spo0A* promoter (48). In *C. difficile*, there are some interesting differences and similarities in the regulatory pathways. Most notably, there seems to be no phosphorelay (2) and the phosphorylation state of Spo0A is supposedly controlled by orphan histidine kinases (35). The transcription of *spo0A* in *C. difficile* is under control of the transition state sigma factor Sigma H (37), as it is in *B. subtilis* (61). Our data indicate that both spo0A and sigH could be targets for direct regulation by Spo0A in *C. difficile* (Figure 3A), raising the possibility of auto-regulation of *spo0A*. The putative direct regulation of *sigH* by Spo0A may reflect that the *C. difficile* genome does not harbor a homolog of the pleiotropic regulator AbrB, which is responsible for the Spo0A-dependent regulation of *sigH* in *B. subtilis* (48). Consistent with a model in which *spo0A* is positively auto-regulated, we noted a sharp increase in the levels of Spo0A as cells approach the stationary growth phase (Figure 1C).

Downstream of Spo0A, we found binding of Spo0A to DNA upstream of several early sporulation genes, such as *spoIIAA*, *spoIIE*, and *spoIIGA* (Figure 3B). All these observations are consistent with direct regulation of these genes by Spo0A in other organisms (5,45,49,62), and the conservation of the sporulation pathway (2).

Regulation of processes outside sporulation

Though Spo0A is the key regulator for sporulation in Firmicutes, it regulates numerous other processes in various bacteria. In the non-pathogenic *B. subtilis*, for instance, the protein also affects competence development, biofilm formation, the

production of and resistance to antimicrobial compounds, chromosome dynamics and aspects of phage biology (10,14–16). Importantly, several of these processes are indirectly regulated, through the Spo0A-dependent repression of *abrB*. Additionally, transcription of *abrB* responds already to low levels of Spo0A~P (43). As a result these effects are detectable in late-exponential and early stationary phase, as some Spo0A is present throughout growth in *B. subtilis* cells.

Though *abrB* is absent from *C. difficile*, this does not exclude the possibility of indirect transcriptional regulation through Spo0A-dependent effects on other regulators. Alternatively, Spo0A may exert a direct effect. In *Clostridium acetobutylicum* and *C. beijerinckii*, Spo0A is a direct regulator of solvent formation, as well as sporulation (22,23). It seems therefore conceivable that Spo0A in *C. difficile* also affects aspects of metabolism. In this respect, it is important to note that also in *C. difficile* Spo0A is detectable from early exponential growth phase on (Figure 1B).

We observed direct binding of *C. difficile* Spo0A to the promoter region of *sigH* (Figure 3A). This gene encodes the key sigma factor for the transition phase, and regulates processes outside sporulation as well (37). Moreover, we found significant levels of Spo0A from early stationary phase on (Figure 1B and unpublished observations), indicating the regulatory actions of Spo0A need not be limited to stationary phase in *C. difficile*. In line with this idea, we found a potential regulatory link between Spo0A and two genes that to our knowledge are not related to the sporulation process, the lipoate ligase *lplA* and the aliphatic sulfonates transporter *ssuA* (Figure 3A). The presence of a putative Spo0A binding site upstream of these genes, as well as the spacing compared to the start codon, is conserved in the problematic Stoke-Mandeville strain (R20291), a member of RT 027. This could indicate that these aspects of regulation by Spo0A are conserved in multiple strains of *C. difficile*.

It should be noted that our work so far has been limited to an in vitro analysis of Spo0A binding, and therefore does not indicate whether activation or repression of the putative target genes occurs in vivo. To answer this question, detailed transcriptome and/or proteome studies have to be performed. In order to distinguish direct from indirect effects, in vivo binding profiles of Spo0A should be performed. The antibodies generated for this study should prove to be useful for this type of experiments.

Regulation of toxin production

Amongst the pathogenic Firmicutes, Spo0A has been reported to affect toxin production in multiple species. In *B. anthracis* a *spo0A* mutation results in elevated levels of AbrB, and concomitantly lower levels of the toxin genes *pagA*, *cya* and *lef* that are under AbrB control (17). Similarly, the production of the emetic toxin

cereulide in *B. cereus* is greatly repressed in a spo0A mutant, in an AbrB-dependent manner (63). In contrast, Spo0A directly represses the expression of the *cry* toxin genes in *B. thuringiensis* and a *spo0A* mutant is therefore a hyper-producer of the insecticidal crystal protein (18,21). In *Clostridium perfringens* TpeL, a member of the large clostridial toxins just like TcdA and TcdB, is directly dependent on Spo0A (64) and also the production of enterotoxin in this organism seems to be (indirectly) dependent on sporulation (65,66).

In *C. difficile* an insertional *spo0A* mutant generated using Clostron technology was reported to have ~10-fold reduced levels of Toxin A (TcdA), both intracellularly and extracellularly as well as ~1000-fold reduced toxicity towards Vero cells, which are primarily sensitive towards Toxin B (TcdB) (35). Our in vitro binding data indicate a potential binding site for Spo0A upstream of *tcdB* and possibly *tcdC* (Figure 5A). However, the in vivo relevance of this binding seems limited as in our hands an independently derived but otherwise identical mutant (a kind gift of the Minton lab; (33)) did not demonstrate a reduced toxicity towards Vero cells. In contrast, we found that in TTY medium toxin levels were slightly elevated in *spo0A* mutant cells compared to wild type (≤2-fold in exponential phase cells up to 4-fold in late-stationary phase cells). The small, and not significant, differences in toxin levels in our experiments might be attributed to differences in the susceptibility of cells for lysis rather than the production of toxin, but could also indicate a negative regulatory effect of Spo0A on toxin production. In support of the latter hypothesis, it was recently reported that a spo0A mutant of *C. difficile* strain R20291 (a RT 027/BI/NAP1 epidemic strain) demonstrates ~10-fold higher toxin levels than its isogenic wild type 30h post inoculation, and is significantly more virulent in a mouse model of disease (34).

The differences between Underwood *et al* (35) on the one hand and our study as well as the study of Deakin and coworkers (34) on the other hand may be explained by differences in experimental conditions, such as the medium used. However, we observed no difference in cytotoxicity between supernatant derived from wild type or *spo0A* mutant cells when they were grown in BHIS, a medium nearly identical to that used previously (data not shown). Alternatively, the differences could indicate integration of the group II intron at more than one location in the chromosome in the strain used in Underwood *et al* (35). In the absence of a complementation experiment and/or Southern blot data, this remains to be established.

In summary, our data are consistent with a model in which the regulation of the major clostridial toxins in *C. difficile* is not positively affected by Spo0A, in contrast to previous findings and other pathogenic Clostridia. Whether Spo0A is truly a negative regulator of toxin production remains to be confirmed using in vitro and in vivo transcription assays.

Concluding remarks

In the present study we have for the first time demonstrated direct binding of the DNA binding domain of *C. difficile* Spo0A to putative target DNA. This work has revealed that aspects of Spo0A binding are conserved between *Bacillus* and *C. difficile* (0A box, possible auto-regulation and binding to early sporulation promoters), whereas others are not (the absence of *abrB* as a direct target in *C. difficile*, binding to DNA upstream of *lplA*, *ssuA*). The effects of Spo0A on toxin production may be similar to those observed for *B. thuringiensis* (18,21). Future work will be aimed at determining the effect of Spo0A on the transcription of the putative target genes, and carry out a comprehensive analysis of Spo0A binding in vivo. The identification of genes affected by Spo0A in *C. difficile* may shed light on the role of the protein in virulence and pathogenesis of this organism.

Reference List

1. Errington J (2003) Regulation of endospore formation in *Bacillus subtilis*. Nat Rev Microbiol 1: 117-126. 10.1038/nrmicro750.
2. Paredes CJ, Alsaker KV, Papoutsakis ET (2005) A comparative genomic view of clostridial sporulation and physiology. Nat Rev Microbiol 3: 969-978. nrmicro1288 (pii);10.1038/nrmicro128.
3. Ferrari FA, Trach K, LeCoq D, Spence J, Ferrari E *et al.* (1985) Characterization of the *spo0A* locus and its deduced product. Proc Natl Acad Sci U S A 82: 2647-2651.
4. Bird TH, Grimsley JK, Hoch JA, Spiegelman GB (1993) Phosphorylation of Spo0A activates its stimulation of *in vitro* transcription from the *Bacillus subtilis spoIIG* operon. Mol Microbiol 9: 741-749.
5. Baldus JM, Green BD, Youngman P, Moran CP, Jr. (1994) Phosphorylation of *Bacillus subtilis* transcription factor Spo0A stimulates transcription from the *spoIIG* promoter by enhancing binding to weak 0A boxes. J Bacteriol 176: 296-306.
6. Burbulys D, Trach KA, Hoch JA (1991) Initiation of sporulation in *B. subtilis* is controlled by a multicomponent phosphorelay. Cell 64: 545-552. 0092-8674(91)90238-T (pii).
7. Grimsley JK, Tjalkens RB, Strauch MA, Bird TH, Spiegelman GB *et al.* (1994) Subunit composition and domain structure of the Spo0A sporulation transcription factor of *Bacillus subtilis*. J Biol Chem 269: 16977-16982.
8. Lewis RJ, Scott DJ, Brannigan JA, Ladds JC, Cervin MA *et al.* (2002) Dimer formation and transcription activation in the sporulation response regulator Spo0A. J Mol Biol 316: 235-245. 10.1006/jmbi.2001.5331 (doi);S0022283601953318 (pii).
9. Muchova K, Lewis RJ, Perecko D, Brannigan JA, Ladds JC *et al.* (2004) Dimer-induced signal propagation in Spo0A. Mol Microbiol 53: 829-842. 10.1111/j.1365-2958.2004.04171.x (doi);MMI4171 (pii).
10. Molle V, Fujita M, Jensen ST, Eichenberger P, Gonzalez-Pastor JE *et al.* (2003) The Spo0A regulon of Bacillus *subtilis*. Mol Microbiol 50: 1683-1701. 3818 (pii).
11. Liu J, Tan K, Stormo GD (2003) Computational identification of the Spo0A-phosphate regulon that is essential for the cellular differentiation and development in Gram-positive spore-forming bacteria. Nucleic Acids Res 31: 6891-6903.
12. Lewis RJ, Krzywda S, Brannigan JA, Turkenburg JP, Muchova K *et al.* (2000) The trans-activation domain of the sporulation response regulator Spo0A revealed by X-ray crystallography. Mol Microbiol 38: 198-212. mmi2134 (pii).
13. Zhao H, Msadek T, Zapf J, Madhusudan, Hoch JA *et al.* (2002) DNA complexed structure of the key transcription factor initiating development in sporulating bacteria. Structure 10: 1041-1050. S0969212602008031 (pii).
14. Castilla-Llorente V, Munoz-Espin D, Villar L, Salas M, Meijer WJ (2006) Spo0A, the key transcriptional regulator for entrance into sporulation, is an inhibitor of DNA replication. EMBO J 25: 3890-3899. 7601266 (pii);10.1038/sj.emboj.7601266.
15. Hahn J, Roggiani M, Dubnau D (1995) The major role of Spo0A in genetic competence is to downregulate *abrB*, an essential competence gene. J Bacteriol 177: 3601-3605.
16. Hamon MA, Lazazzera BA (2001) The sporulation transcription factor Spo0A is required for biofilm development in *Bacillus subtilis*. Mol Microbiol 42: 1199-1209. 2709 (pii).

17. Saile E, Koehler TM (2002) Control of anthrax toxin gene expression by the transition state regulator *abrB*. J Bacteriol 184: 370-380.
18. Agaisse H, Lereclus D (1994) Expression in *Bacillus subtilis* of the Bacillus *thuringiensis cryIIIA* toxin gene is not dependent on a sporulation-specific sigma factor and is increased in a *spo0A* mutant. J Bacteriol 176: 4734-4741.
19. Baum JA, Malvar T (1995) Regulation of insecticidal crystal protein production in Bacillus *thuringiensis*. Mol Microbiol 18: 1-12.
20. Chen HJ, Tsai TK, Pan SC, Lin JS, Tseng CL *et al.* (2010) The master transcription factor Spo0A is required for poly(3-hydroxybutyrate) (PHB) accumulation and expression of genes involved in PHB biosynthesis in *Bacillus thuringiensis*. FEMS Microbiol Lett 304: 74-81. FML1888 (pii);10.1111/j.1574-6968.2010.01888.x.
21. Poncet S, Dervyn E, Klier A, Rapoport G (1997) Spo0A represses transcription of the *cry* toxin genes in *Bacillus thuringiensis*. Microbiology 143 (Pt 8): 2743-2751.
22. Ravagnani A, Jennert KC, Steiner E, Grunberg R, Jefferies JR *et al.* (2000) Spo0A directly controls the switch from acid to solvent production in solvent-forming clostridia. Mol Microbiol 37: 1172-1185. mmi2071 (pii).
23. Harris LM, Welker NE, Papoutsakis ET (2002) Northern, morphological, and fermentation analysis of *spo0A* inactivation and overexpression in *Clostridium acetobutylicum* ATCC 824. J Bacteriol 184: 3586-3597.
24. Rupnik M, Wilcox MH, Gerding DN (2009) *Clostridium difficile* infection: new developments in epidemiology and pathogenesis. Nat Rev Microbiol 7: 526-536. nrmicro2164 (pii);10.1038/nrmicro2164.
25. Vedantam G, Clark A, Chu M, McQuade R, Mallozzi M *et al.* (2012) *Clostridium difficile* infection: Toxins and non-toxin virulence factors, and their contributions to disease establishment and host response. Gut Microbes 3. 19399 (pii).
26. Freeman J, Bauer MP, Baines SD, Corver J, Fawley WN *et al.* (2010) The changing epidemiology of *Clostridium difficile* infections. Clin Microbiol Rev 23: 529-549. 23/3/529.
27. Goorhuis A, Bakker D, Corver J, Debast SB, Harmanus C *et al.* (2008) Emergence of *Clostridium difficile* infection due to a new hypervirulent strain, polymerase chain reaction Ribotype 078. Clin Infect Dis 47: 1162-1170. 10.1086/592257.
28. Warny M, Pepin J, Fang A, Killgore G, Thompson A *et al.* (2005) Toxin production by an emerging strain of *Clostridium difficile* associated with outbreaks of severe disease in North America and Europe. Lancet 366: 1079-1084. S0140-6736(05)67420-X.
29. Lyras D, O'Connor JR, Howarth PM, Sambol SP, Carter GP *et al.* (2009) Toxin B is essential for virulence of *Clostridium difficile*. Nature 458: 1176-1179. nature07822 (pii);10.1038/nature07822.
30. Kuehne SA, Cartman ST, Heap JT, Kelly ML, Cockayne A *et al.* (2010) The role of Toxin A and Toxin B in *Clostridium difficile* infection. Nature 467: 711-713. nature09397 (pii);10.1038/nature09397 (doi).
31. Perelle S, Gibert M, Bourlioux P, Corthier G, Popoff MR (1997) Production of a complete binary toxin (actin-specific ADP-ribosyltransferase) by *Clostridium difficile* CD196. Infect Immun 65: 1402-1407.
32. Stubbs S, Rupnik M, Gibert M, Brazier J, Duerden B *et al.* (2000) Production of actin-specific ADP-ribosyltransferase (binary toxin) by strains of *Clostridium difficile*. FEMS Microbiol Lett 186: 307-312. S0378-1097(00)00162-2 (pii).

33. Heap JT, Pennington OJ, Cartman ST, Carter GP, Minton NP (2007) The ClosTron: a universal gene knock-out system for the genus *Clostridium* . J Microbiol Methods 70: 452-464. S0167-7012(07)00208-4 (pii);10.1016/j.mimet.2007.05.021 (doi).
34. Deakin LJ, Clare S, Fagan RP, Dawson LF, Pickard DJ *et al.* (2012) *Clostridium difficile spo0A* gene is a persistence and transmission factor. Infect Immun . IAI.00147-12 (pii);10.1128/IAI.00147-12 (doi).
35. Underwood S, Guan S, Vijayasubhash V, Baines SD, Graham L *et al.* (2009) Characterization of the sporulation initiation pathway of *Clostridium difficile* and its role in toxin production. J Bacteriol 191: 7296-7305. JB.00882-09 (pii);10.1128/JB.00882-09 (doi).
36. Steiner E, Dago AE, Young DI, Heap JT, Minton NP *et al.* (2011) Multiple orphan histidine kinases interact directly with Spo0A to control the initiation of endospore formation in *Clostridium acetobutylicum*. Mol Microbiol 80: 641-654. 10.1111/j.1365-2958.2011.07608.x (doi).
37. Saujet L, Monot M, Dupuy B, Soutourina O, Martin-Verstraete I (2011) The key sigma factor of transition phase, SigH, controls sporulation, metabolism, and virulence factor expression in *Clostridium difficile*. J Bacteriol 193: 3186-3196. JB.00272-11 (pii);10.1128/JB.00272-11 (doi).
38. Hussain HA, Roberts AP, Mullany P (2005) Generation of an erythromycin-sensitive derivative of *Clostridium difficile* strain 630 (630Deltaerm) and demonstration that the conjugative transposon Tn*916*DeltaE enters the genome of this strain at multiple sites. J Med Microbiol 54: 137-141.
39. Snijder EJ, Bredenbeek PJ, Dobbe JC, Thiel V, Ziebuhr J *et al.* (2003) Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. J Mol Biol 331: 991-1004. S0022283603008659 (pii).
40. Baerends RJ, Smits WK, de Jong A, Hamoen LW, Kok J *et al.* (2004) Genome2D: a visualization tool for the rapid analysis of bacterial transcriptome data. Genome Biol 5: R37. 10.1186/gb-2004-5-5-r37 (doi);gb-2004-5-5-r37 (pii).
41. Heap JT, Cartman ST, Kuehne SA, Cooksley C, Minton NP (2010) ClosTron-targeted mutagenesis. Methods Mol Biol 646: 165-182. 10.1007/978-1-60327-365-7_11 (doi).
42. Sebaihia M, Wren BW, Mullany P, Fairweather NF, Minton N *et al.* (2006) The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome. Nat Genet 38: 779-786. ng1830 (pii);10.1038/ng1830 (doi).
43. Fujita M, Gonzalez-Pastor JE, Losick R (2005) High- and low-threshold genes in the Spo0A regulon of Bacillus *subtilis*. J Bacteriol 187: 1357-1368. 187/4/1357 (pii);10.1128/JB.187.4.1357-1368.2005 (doi).
44. Strauch M, Webb V, Spiegelman G, Hoch JA (1990) The SpoOA protein of *Bacillus subtilis* is a repressor of the *abrB* gene. Proc Natl Acad Sci U S A 87: 1801-1805.
45. Baldus JM, Buckner CM, Moran CP, Jr. (1995) Evidence that the transcriptional activator Spo0A interacts with two sigma factors in Bacillus *subtilis*. Mol Microbiol 17: 281-290.
46. Strauch MA, Trach KA, Day J, Hoch JA (1992) Spo0A activates and represses its own synthesis by binding at its dual promoters. Biochimie 74: 619-626.
47. Weir J, Predich M, Dubnau E, Nair G, Smith I (1991) Regulation of *spo0H*, a gene coding for the Bacillus *subtilis* sigma H factor. J Bacteriol 173: 521-529.

48. Fujita M, Sadaie Y (1998) Feedback loops involving Spo0A and AbrB in *in vitro* transcription of the genes involved in the initiation of sporulation in *Bacillus subtilis*. J Biochem 124: 98-104.
49. York K, Kenney TJ, Satola S, Moran CP, Jr., Poth H *et al.* (1992) Spo0A controls the sigma A-dependent activation of *Bacillus subtilis* sporulation-specific transcription unit *spoIIE*. J Bacteriol 174: 2648-2658.
50. Olling A, Seehase S, Minton NP, Tatge H, Schroter S *et al.* (2012) Release of TcdA and TcdB from *Clostridium difficile* cdi 630 is not affected by functional inactivation of the *tcdE* gene. Microb Pathog 52: 92-100. S0882-4010(11)00191-4 (pii);10.1016/j.micpath.2011.10.009 (doi).
51. Tan KS, Wee BY, Song KP (2001) Evidence for holin function of *tcdE* gene in the pathogenicity of *Clostridium difficile*. J Med Microbiol 50: 613-619.
52. Matamouros S, England P, Dupuy B (2007) *Clostridium difficile* toxin expression is inhibited by the novel regulator TcdC. Mol Microbiol 64: 1274-1288. MMI5739 (pii).
53. Dupuy B, Govind R, Antunes A, Matamouros S (2008) *Clostridium difficile* toxin synthesis is negatively regulated by TcdC. J Med Microbiol 57: 685-689. 57/6/685 (pii);10.1099/jmm.0.47775-0 (doi).
54. Carter GP, Douce GR, Govind R, Howarth PM, Mackin KE *et al.* (2011) The anti-sigma factor TcdC modulates hypervirulence in an epidemic BI/NAP1/027 clinical isolate of *Clostridium difficile*. PLoS Pathog 7: e1002317. 10.1371/journal.ppat.1002317 (doi);PPATHOGENS-D-11-00972 (pii).
55. Cartman ST, Kelly ML, Heeg D, Heap JT, Minton NP (2012) Precise Manipulation of the *Clostridium difficile* Chromosome Reveals a Lack of Association Between *tcdC* Genotype and Toxin Production. Appl Environ Microbiol . AEM.00249-12 (pii);10.1128/AEM.00249-12 (doi).
56. Bakker D, Smits WK, Kuijper EJ, Corver J (2012) TcdC Does Not Significantly Repress Toxin Expression in *Clostridium difficile* 630Δ*erm*. PLoS ONE 7(8): e43247. doi:10.1371/journal.pone.0043247
57. Fujita M, Losick R (2005) Evidence that entry into sporulation in *Bacillus subtilis* is governed by a gradual increase in the level and activity of the master regulator Spo0A. Genes Dev 19: 2236-2244. 19/18/2236 (pii);10.1101/gad.1335705 (doi).
58. Bailey TL, Boden M, Buske FA, Frith M, Grant CE *et al.* (2009) MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res 37: W202-W208. gkp335 (pii);10.1093/nar/gkp335 (doi).
59. Piggot PJ, Hilbert DW (2004) Sporulation of *Bacillus subtilis*. Curr Opin Microbiol 7: 579-586. S1369-5274(04)00125-0 (pii);10.1016/j.mib.2004.10.001 (doi).
60. Ireton K, Rudner DZ, Siranosian KJ, Grossman AD (1993) Integration of multiple developmental signals in *Bacillus subtilis* through the Spo0A transcription factor. Genes Dev 7: 283-294.
61. Predich M, Nair G, Smith I (1992) *Bacillus subtilis* early sporulation genes *kinA*, *spo0F*, and *spo0A* are transcribed by the RNA polymerase containing sigma H. J Bacteriol 174: 2771-2778.
62. Satola S, Kirchman PA, Moran CP, Jr. (1991) Spo0A binds to a promoter used by sigma A RNA polymerase during sporulation in *Bacillus subtilis*. Proc Natl Acad Sci U S A 88: 4533-4537.

63. Lucking G, Dommel MK, Scherer S, Fouet A, Ehling-Schulz M (2009) Cereulide synthesis in emetic *Bacillus cereus* is controlled by the transition state regulator AbrB, but not by the virulence regulator PlcR. Microbiology 155: 922-931. 155/3/922 (pii);10.1099/mic.0.024125-0 (doi).
64. Paredes-Sabja D, Sarker N, Sarker MR (2011) *Clostridium perfringens tpeL* is expressed during sporulation. Microb Pathog 51: 384-388. S0882-4010(11)00123-9 (pii);10.1016/j.micpath.2011.05.006 (doi).
65. Duncan CL, Strong DH, Sebald M (1972) Sporulation and enterotoxin production by mutants of *Clostridium perfringens*. J Bacteriol 110: 378-391.
66. Zhao Y, Melville SB (1998) Identification and characterization of sporulation-dependent promoters upstream of the enterotoxin gene (*cpe*) of *Clostridium perfringens*. J Bacteriol 180: 136-142.

6

TcdC does not significantly repress toxin expression in *Clostridium difficile* 630Δ*erm*

D. Bakker, W.K. Smits, E.J. Kuijper and J. Corver

Department of Medical Microbiology, Center of Infectious Diseases, Leiden University Medical Center, the Netherlands.

PLoS One 2012

Abstract

In the past decade, *Clostridium difficile* has emerged as an important gut pathogen. Symptoms of *C. difficile* infection range from mild diarrhea to pseudomembranous colitis, sometimes resulting in colectomy or death. The main virulence factors of *C. difficile* are toxin A and toxin B. Besides the genes encoding these toxins (*tcdA* and *tcdB*), the pathogenicity locus (PaLoc) also contains genes encoding a sigma factor (*tcdR*) and a putative anti-sigma factor (*tcdC*). The important role of TcdR as a sigma factor for toxin expression is undisputed, whereas the role of TcdC as an anti-sigma factor, inhibiting toxin expression, is currently the subject of debate.

To clarify the role of TcdC in toxin expression, we generated an isogenic ClosTron-based mutant of *tcdC* in *Clostridium difficile* strain 630Δ*erm* (CT::*tcdC*) and determined the transcription levels of the PaLoc genes and the expression levels of the toxins in the wild type strain and the *tcdC* mutant strain.

We found only minor differences in transcription levels of the PaLoc genes between the wild type and CT::*tcdC* strains and total toxin levels did not significantly differ either. These results suggest that in *C. difficile* 630Δ*erm* TcdC is not a major regulator of toxin expression under the conditions tested.

Introduction

Clostridium difficile is an anaerobic, Gram-positive, spore forming rod shaped bacterium that can cause disease with a wide variety of symptoms, ranging from mild diarrhea to severe forms of pseudomembranous colitis (1–3). Since 2004, numerous countries have reported outbreaks in health-care facilities caused by hypervirulent *C. difficile* PCR Ribotype (RT) 027 (1–6). *Clostridium difficile* infection (CDI) caused by RT 027 is associated with a more severe course of the disease and a higher mortality rate than other ribotypes (1,3,6). Recently, increasing numbers of the hypervirulent RT 078 are reported (7). *C. difficile* RT 078 is more frequently associated with community acquired CDI and affects a younger population than RT 027 (6–9). Furthermore, CDI caused by RT 078 is associated with an increased morbidity compared to other ribotypes (8).

The main virulence factors of the enteropathogenic *C. difficile* are the two large clostridial Toxins, Toxin A (TcdA) and Toxin B (TcdB). These toxins are glycosyltransferases that inactivate Rho, Rac and Cdc42, thereby disrupting the cytoskeleton and tight junctions of the cells, resulting in apoptosis (10). This induces an inflammatory response and degradation of the intestinal epithelial cell layer. Besides the genes encoding these toxins (*tcdA* and *tcdB*), the Pathogenicity Locus (PaLoc) also contains genes encoding a sigma factor (*tcdR*) and a putative anti-sigma factor (*tcdC*) (11–13). In between the toxin genes the *tcdE* gene is situated, which encodes a putative holin protein (14). Interestingly, both hypervirulent RTs 027 and 078 have been shown to contain mutations in the *tcdC* gene, encoding the proposed negative regulator of toxin gene transcription, and this has been proposed as a possible explanation for their increased virulence (8,15).

The exponential growth phase of *C. difficile* has been reported to be associated with a high transcription level of the *tcdC* gene and low transcription levels of *tcdR* and the toxin genes, whereas the stationary growth phase is associated with a low transcription level of the *tcdC* gene and high transcription levels of *tcdR* and the toxin genes in strain VPI10463 (16). The synthesis and secretion of the toxins is increased upon entry into the stationary growth phase (16–19). The decreasing transcription of *tcdC* correlates with diminishing TcdC protein levels in stationary growth phase (16,20).

TcdR is an alternative sigma factor that positively regulates toxin production (11,12). The direct interaction of TcdR and the RNA polymerase core enzyme mediates recognition of the toxin promoters and the *tcdR* promoter (11,12,21). TcdC has been reported to act like an anti-sigma factor for toxin production by destabilizing the TcdR-RNA polymerase core enzyme complex in a way that is not yet fully understood (12).

The reported inverse correlation between the transcription of *tcdC* and the toxin genes and the expression patterns of the corresponding proteins, together with the biochemical data, has led to the prevailing model that TcdC is an important repressor of toxin expression (12,16,17,20). This model seems to be supported by the finding that the absence of a functional TcdC caused by a frame shift mutation (Δ117 bp) in the *tcdC* gene is linked to a supposed increased toxin production in certain (hyper) virulent strains (15,22).

Recently, some doubts were raised about the importance of TcdC for regulation of toxin expression on the basis of two findings. First, two studies have found increasing levels of *tcdC* transcription in time that coincide with increasing transcription of the toxin genes and increasing amounts of toxin production (18,19). Second, there is a great variability in toxin expression levels among (hyper) virulent strains, even though these generally carry mutations in *tcdC* (15,18,19). Therefore, a minor (or modulatory) role for TcdC in the regulation of toxin expression was proposed (18,19).

Here, we sought to clarify the role of TcdC in regulation of the toxin production by generating an isogenic *tcdC* mutant (CT::*tcdC*) using the ClosTron technology**.** We find only minor differences in transcription levels of the PaLoc genes between the wild type and CT::*tcdC* strains and the expressed total toxin levels did not significantly differ, suggesting that the role of TcdC in toxin regulation is not of significance under the conditions tested in *C. difficile* strain 630Δ*erm*.

Material & Methods

Bacterial strains and growth conditions

The *Clostridium difficile* and *Escherichia coli* (*E. coli*) strains and plasmids used in this study are described in Table 1. *E. coli* strains were grown in Luria Bertani (LB, USB cooperation) medium supplemented with appropriate antibiotics when required. *C. difficile* strains were grown anaerobically in a microaerobic cabinet (Don Whitley DG 250) at 37 ºC in pre-reduced 3% Bacto Tryptose (Difco), 2% Yeast extract (Difco) and 0.1% thioglycolate (pH 7.4) medium (TY) or Brain Heart Infusion broth (Oxoid) supplemented with 0.5% yeast extract and 0.01% L-cysteine (Sigma) (BHIS) (40,41). When required, the broths were supplemented with appropriated antibiotics. For RNA extraction and toxin quantification, *C. difficile* 630Δ*erm* (wild type) and two independent isogenic *tcdC* mutant strains (CT::*tcdC*) were serially diluted and pre-cultured (overnight) in pre-reduced TY broth. Mid-logarithmic growth phase pre-cultures (OD_{600} 0.4-0.8) were used to inoculate pre-reduced TY broth to a starting OD_{600} of 0.05 (± 0.01). Optical density readings and samples for

total toxin quantification were taken hourly in the exponential growth phase (until 8 hours post inoculation) and at 12, 24 and 48 hours post inoculation. Samples for RNA extraction were taken at 6, 8, 12, and 24 hours post inoculation. Samples for Western blot detection of TcdC were taken at 8 hours post inoculation. We routinely monitored the purity of the *C. difficile* cultures by culturing on appropriate agar plates and performed control PCRs to ensure that the insertional disruption of the *tcdC* gene had remained intact during our experiments. All experiments were performed six times.

Table 1: Strains and plasmids used in this study.

Strains	Description	Origin
Escheria coli		
DH5α	Erythromycin^S, Lincomycin^S	Laboratory stock
CA434	Erythromycin^S, Lincomycin^S, Kanamycin^R, plasmid R702	(41)
Clostridium difficile		
630Δ*erm* (wt)	Erythromycin^s, Lincomycin^S	(34)
Leeds_035	RT 035, *tcdC* negative, PaLoc negative	(44)
*CT::tcdC*1	630Δ*erm*Δ*tcdC*623as , Erythromycin^R, Lincomycin^R	This study
*CT::tcdC*2	630Δ*erm*Δ*tcdC*623as , Erythromycin^R, Lincomycin^R	This study
Plasmids		
pMTL007C-E2	Thiamphenicol^R, Erythromycin^S	(25)
pDB001	pMTL007C- E2:*tcdC*623as	This study

Generation of *tcdC* mutant strains

We generated two independent isogenic *tcdC* mutants by insertional inactivation of the *tcdC* gene in the wild type strain 630Δ*erm* using ClosTron technology (24,25). Briefly, the Perutka algorithm on the ClosTron website (http://www.clostron.com) was used to design primers (Table 2) for retargeting the Group II intron (Sigma; Targetron). The retargeted intron was cloned using the restriction enzymes BsrGI and HindIII into plasmids pMTL007C-E2 and the constructs were verified by sequencing (25). The verified plasmid (pDB001) was transformed to *E. coli* CA434 and transferred to the wild type strain 630Δ*erm* via conjugation (34,41). The selection of *C. difficile* transconjugants was done by subculturing on pre-reduced BHIS agar supplemented with thiamphenicol (Sigma; 10μg/mL) and *C. difficile* selective supplement (Oxoid). This was followed by several rounds of subculturing on pre-reduced BHIS agar supplemented with lincomycin (Sigma; 20 μg/mL) and *C. difficile* selective supplement to promote integration of the GroupII intron into the gene of interest. Chromosomal DNA isolated from the transconjugants using

a QIAamp blood kit (Qiagen) was used in conventional PCRs and sequence runs to confirm the disruption of *tcdC* and the nucleotide position of the insertion in the *tcdC* gene. Primers used for cloning and sequencing are listed in Table 2.

Complementation can be a valuable control for knockout studies. However, as our *tcdC* mutant strains have no clearly detectable phenotype regarding toxin production, complemented mutant strains are expected to be comparable to wild type and *tcdC* mutant strains, as also reported recently in an independent study (38). Therefore, a complementation study would not add to the message this manuscript.

Southern blots

Southern blot analysis was performed to verify a specific single integration into the genome. Genomic DNA was extracted using a Phenol-chloroform extraction (42). Four μg of genomic DNA was digested with EcoRV enzyme and separated on a 0.8% agarose/0.5xTris-acetate-EDTA gel by electrophoresis. DNA was transferred onto a Hybond N+ filter (Amersham) in 10x saline sodium citrate (SSC) solution. The filter was washed in 2X SSC and baked at 80°C for 2 hours. Prehybridization of the filter was done for 2 hrs at 60°C in 5x SCC, 5x Denhart and 100 mg/mL of yeast tRNA. Probes specific for the group II intron (EBS2-*tcdC*623as/Sal-R1), *ermB* gene (oWKS1131/oWKS1132) and *tcdC* gene (*tcdC*5-*tcdC*6) were generated. Primers are listed in Table 2. The generated probes (100 ng) were radiolabeled (^{32}P dATP) using Klenow enzyme (Roche) and overnight hybridized in 10 mL fresh pre-hybridization buffer at 60°C. The filter was washed for 30 min in 2x SCC, 0.5% SDS, 30 min in 1X SSC, 0.5% SDS and 30 min in 0.5X SSC, 0.5% SDS and analyzed using phosphorimage screen and a Typhoon 9410 scanner (GE healthcare).

Western blots

Antibodies against TcdC were generated by immunizing rabbits with a synthetic peptide (CQLARTPDDYKYKKV) representing a specific TcdC epitope (Genscript). Note that this epitope is located before the Clostron insertion site, and would therefore also be expected to detect truncated TcdC protein, would this be produced. Western blots were performed as follows. *C. difficile* (2 mL) cultures were harvested by centrifugation (2 min, 11.000 xg, 4ºC) and washed with Phosphate Buffered Saline (PBS). The bacterial pellets were resuspended in PBS containing protease inhibitor cocktail (Complete, Roche) and lysed by sonification. The bacterial lysates were centrifuged at low speed (3 min, 1000xg, 4ºC) to remove unbroken bacterial cells (20). To separate the cytosolic proteins from the membrane associated proteins the bacterial supernatant was centrifuged at 200.000 xg, 4ºC for 1 hr (20). The pelleted membrane associated proteins were resuspended in 10 mM Tris-HCl (pH 7.4), 5 mM EDTA with 2% Triton X-100

Table 2. Primers and probes used in this study.

Primers	Sequence (5'- 3')	
IBS-*tcdC*623as	AAAAAAGCTTATAATTATCCTTAGTTATCGTTC-CAGTGCGCCCAGATAGGGTG	This study
EBS2-*tcdC*623as	TGAACGCAAGTTTCTAATTTCGATTATAAC-TCGATAGAGGAAAGTGTCT	This study
EBS1d-*tcdC*623as	CAGATTGTACAAATGTGGTGATAACAGA-TAAGTCGTTCCAGCTAACTTACCTTTCTTTGT	This study
EBS universal	Intron mutagenesis/Control PCR/ CGAAAT-TAGAAACTTGCGTTCAGTAAAC TGAACG-CAAGTTTCTAATTTCGATTATAACTCGATAGAG-GAAAGTGTCT	(25)
*tcdC*1	Control PCR/ATGTTTTCTAAAAAAAATGAT	This study
*tcdC*2	Control PCR/TTAATTAATTTTCTCTACAGCT	This study
tcdR Forward	Multiplex 1/ATAATGATGCCCACAAGATGATTTAG	This study
tcdR Reverse	Multiplex 1/AAAGAAGTGATCTATATCATCAGT-TAC	This study
tcdR probe	Multiplex 1/TEX-TATGACCTGAACCACCTTCCAT-TCTCC-BHQ-2	This study
tcdB forward	Multiplex 1/ATAATGATGCCCACAAGATGATTTAG	This study
tcdB Reverse	Multiplex 1/AAAGAAGTGATCTATATCATCAGT-TAC	This study
tcdB probe	Multiplex 1/TEX-TATGACCTGAACCACCTTCCAT-TCTCC-BHQ-2	This study
tcdE Forward	Multiplex 2/ATTTGATACATTATTAGGATGTT-TAAG	This study
tcdE Reverse	Multiplex 2/AAATATACATGCTATCATTGCTAC	This study
tcdE probe	Multiplex 2/FAM-TGATTCCTCCATCTATTC-CAAAACTAGAA-BHQ-1	This study
tcdA forward	Multiplex 1/AATTCCAATACAAGCCCTGTAG	This study
tcdA Reverse	Multiplex 1/TATCAGCCCATTGTTTTATGTATTC	This study
tcdA probe	Multiplex 1/FAM-ATCACTGACTTCTCCACCTATC-CATACAA-BHQ-1	This study
*tcdC*3	Multiplex 1/CATAATTTCCAGACACAGCTAATC	This study
*tcdC*4	Multiplex 1/GGATATGATACTGGTATTACT-TATGAC	This study
tcdC probe	Multiplex 1/YAK-TGCACCTCATCACCATCTTCAA-TAACTTG-BHQ1	This study
rspJ Forward	GATCACAAGTTTCAGGACCTG	This study
rspJ Reverse	GTCTTAGGTGTTGGATTAGC	This study
tcdC5	CATATCCTTTCTTCTCCTCTTC	This study
tcdC6	AATTGTCTGATGCTGAACC	This study
oWKS-1131	AAAGCGATGCCGAGAATCTG	This study
oWKS-1132	TCTCGGAGTATACGGCTCTG	This study
Sal-R1	ATTACTGTGACTGGTTTGCACCACCCTCTTCG	(45)

for 30 min at room temperature. Equal amounts of the resuspended membrane associated proteins were separated on 15% SDS PAGE gel and transferred onto polyvinyl difluoride (PVDF) membranes. Similarly generated membranes with the transferred membrane associated proteins of a RT 035 (PaLoc negative) strain were used for pre incubation of the TcdC antibodies. The membranes were probed with the pre incubated TcdC antibody and an antibody against the β subunit of the *E. coli* F_0F_1 ATPase that cross reacts with the homologous protein in *C. difficile* (20,43). The probed membranes were analyzed using secondary anti-mouse horse radish peroxidase conjugated antibodies (Dako), a chemiluminescence detection kit (Amersham) and a Typhoon 9410 scanner (GE healthcare).

RNA extraction

Five mL of the *C. difficile* cultures were 1:1 diluted with ice cold methanol and stored overnight at -80°C. Bacterial pellets, obtained by centrifugation (20 min, 3000 x g, 4°C), were resuspended into 200 µL lysisbuffer (100 mM EDTA, 200 mM Tris-HCl pH 7.0, 50 mg/mL lysozyme) and incubated for 1 hr at 37°C. Tri-pure reagent (Roche) was used for the extraction of RNA according to the manufacturer's instruction with minor modifications. Briefly, 1 mL Tri-pure was added to the lysed bacterial pellets and incubated for 5 min at room temperature. Per 1 mL Tri-pure, 200 µL chloroform was added and carefully shaken by hand for 3 min, followed by an incubation of 2-5 min at room temperature. The aqueous phase was collected after centrifugation (12,000 x g for 15 min at 4°C) and transferred to a fresh tube. RNA was precipitated by mixing the aqueous phase with 500 µL isopropanol, followed by an incubation of 10 min at room temperature. The precipitated RNA was collected by centrifugation (12,000 x g, 10 min, 4°C) and resuspended in 100 µL DNase/RNase free water. The RNA was re-precipitated overnight at -80°C with ammonium acetate (Fluka; 10 mM) and 3 volumes of absolute ethanol. The re-precipitated RNA was washed once with 80% ethanol and dissolved in 50 µL DNase/RNase free water. The RNA was treated twice with a TurboDNase (Ambion) according to the manufacturer's instruction followed by another Tri-pure RNA isolation. The quality and purity of the extracted RNA was assessed using a RNA nano chip on an Agilent Bioanalyzer.

Transcriptional analysis of the PaLoc genes

A RevertAid™ H Minus Reverse Transcriptase kit (Fermentas) was used to synthesize cDNA according to the manufacturer's instruction. Random hexamers were used to convert 750 ng RNA into cDNA. The synthesized cDNA was treated with RNase (Qiagen) for 1 hour at 37°C and stored at -20°C. The software program Molecular Beacon (Premier Biosoft) was used to design primer pairs

and probes (Table 2) for the 2 multiplex quantitative PCRs (qPCR), based on the available genome of *C. difficile* strain 630 (35). All primer pairs were first tested by conventional PCR and multiplex PCR to confirm specificity and amplicon sizes. The primer pair and the probe for the amplification of the *tcdC* gene are in front of the insertion site in the *tcdC* gene (Figure 1A), allowing detection of *tcdC* transcription levels in wild type and CT::*tcdC* strains. The real-time multiplex qPCR amplification of the PaLoc genes and the reference gene encoding for a ribosomal protein (*rpsJ*) was performed on a CFX96 real-time PCR detection system (Biorad) (31). The amplification efficiencies of the PaLoc and reference genes were determined using serially diluted genomic DNA (standard curve). The manually calculated efficiencies and the reference gene *rpsJ* were used to normalize the expression levels of the PaLoc genes. The amplification was performed in a 25 µL final volume. The first real-time multiplex qPCR (target genes: *tcdA*, *tcdA* and *tcdC*) contained 25 µL Hotstar mastermix (Qiagen), forward and reverse primers (80 nm each primer), 2.5 mM $MgCl_2$, 100 nM of each probe and 2 µL synthesized cDNA. The second multiplex real-time multiplex Q-PCR (target genes: *tcdR*, *tcdE*) contained 25 µL Hotstar mastermix (Qiagen), forward and reverse primers (80 nm each primer), 3.5 mM $MgCl_2$, 100 nM of each probe and 2 µL synthesized cDNA. The real-time qPCR to quantify the reference gene *rpsJ* contained 25 µL Hotstar mastermix (Qiagen), forward and reverse primers (80 nm each primer), 2.5 mM $MgCl_2$, 0.06% SYBRgreen (Sigma) and 2 µL synthesized cDNA. The real-time qPCR protocol included an enzyme activation step for 15 min at 95°C, followed by 50 cycles of amplification; 95°C for 30 sec, 52°C for 30 sec and 72°C for 30 sec.

Relative quantification of toxin expression

Total toxin amounts were quantified using 2 assays; a toxin end point titer assay and a commercial available ELISA (Ridascreen, Biopharma). The supernatants of culture samples (1 mL) were collected after centrifugation (30 min, 3000 xg, 4ºC), filter sterilized (0.45 µM cellulose acetate membrane) and stored at 4°C.

For the toxin end point titer assay, Vero cells were seeded into a 96 wells plate at a density of 1x 10^4 cells per well and incubated overnight at 37°C and 5% CO_2. The filter sterilized supernatants of 5, 8, 12, 24 and 48 hours post inoculation (hpi) were diluted 2, 10^1, 10^2, 10^3, 10^4 and 10^5 fold in cell culture medium (Dulbecco modified Eagle medium (Lonza) supplemented with penicillin 100 u/mL, streptomycin 100 U/mL, fetal calf serum(10%). Fifty µL of the dilutions were added onto the Vero cell monolayers and incubated for 1 hr at 37ºC and 5% CO_2. For the neutralization assay a 2-fold dilution of each tested time point (5, 8, 12, 24 and 48 hpi) was pre-incubated with a 1/100 diluted anti-toxin (Techlab) for 1 hr at 37°C and 5% CO_2. After the pre-incubation, 50 µL was added onto the Vero cell monolayers. The

incubated bacterial supernatants were aspirated off after one hour and replaced with 100 µL cell culture medium. After 24 hrs of incubation the end-point titer was determined of each diluted time point (26). The end-point titer was defined as the first dilution at which the Vero Cell morphology was indistinguishable from the neutralized 2-fold diluted supernatants (26). The enzyme immunoassay (Ridascreen, Biopharma) was performed according manufacture's protocol.

Statistical analysis

Statistical analysis was performed using the software package SPSS 18 (IBM). An independent sample t-test was employed to compare the strains at different time points.

Results

The importance of TcdC for regulation of toxin expression was recently challenged by two studies (18,19). It was proposed, based on the increasing transcription levels of the PaLoc genes in time and the variability in toxin expression levels among virulent strains, that TcdC has a minor or modulatory role on toxin expression rather than a major role as previously assumed. In this study we sought to clarify the role of TcdC for toxin expression by generating an isogenic *tcdC* mutant. As toxin gene expression is subject to complex regulation influenced by glucose and cysteine, we performed our experiments in a trypton-yeast (TY) based broth (17,23). TY broth does not contain glucose and no cysteine was added. We verified that in TY broth earlier and higher expression of toxins was achieved in comparison to the commonly used Brain Heart Infusion (BHI) broth (data not shown).

Generation and characterization of a TcdC mutant

TcdC consists of three domains: a hydrophobic domain, a proposed dimerization domain and a proposed C-terminal repressor domain (Figure 1A) (12). We successfully disrupted the *tcdC* gene in the region coding for the repressor domain using ClosTron technology. Disruption of genes using the ClosTron technology results in stable mutants and no or non-functional proteins (24–26). The genotype of the disruption was confirmed with conventional PCRs using the *tcdC*2 primer and the EBS universal primer and with primer pairs (*tcdC*1 and *tcdC*2) flanking the ClosTron insertion site (Figure B). Sequence analysis confirmed that the disruption was in the proposed repressor domain of the *tcdC* gene at the expected site (data not shown). In addition, Southern blot analysis using intron-, *ermB* and *tcdC*- specific probes clearly confirmed a specific single insertion of the Group II intron in the genome (Figure 1C).

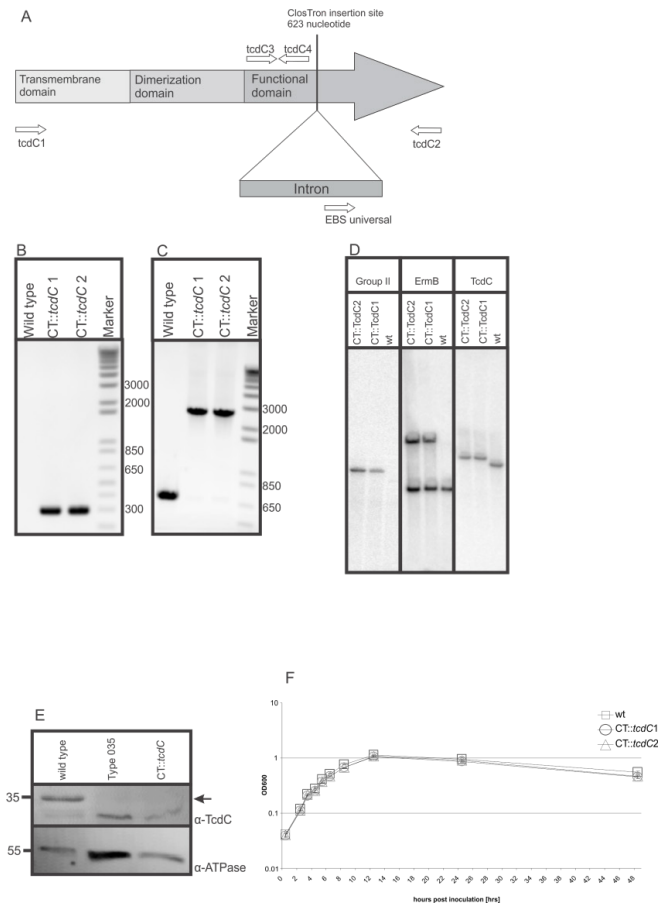


Figure 1: Characterization of the *C. difficile tcdC* mutant. **(A)** Schematic representation of 3 different domains of TcdC and the intron insertion site for the inactivation of TcdC. The arrows in the putative repressor domain represent the locations and orientation of the primers used in the RT-q-PCR and conventional control PCRs. **(B and C)** PCR confirmation of the wild type strain and the CT::*tcdC* mutant. The primer EBS universal and *tcdC*2 generated a PCR product of 302 bp for the CT::*tcdC* strains. Primers *tcdC*1 and *tcdC*2 generated a 699 bp PCR product for the wild type and for the CT::*tcdC* strain a PCR product of circa 2800 bp. **(D)** Southern blot analysis of EcoRV digested genomic DNA of wild type and CT::*tcdC* strains with a Group II intron, *ermB* gene and *tcdC* specific probes. Note that probing with the *ermB* probe results in 2 bands for the CT::*tcdC* strains, since wild type already carries a copy of the *ermB* gene in the genome (35). **(E)** Western blot analysis of TcdC production in wild type and CT::*tcdC* strain 8 hours post inoculation. The arrow indicates the location of TcdC protein based on MW and absence of the protein in the PaLoc negative Type 035 strain. Note that cross-reaction of TcdC antibody with a protein of similar MW was also observed in Carter *et al.* (30). **(F)** Growth curves of *C. difficile* 630Δ*erm* and *C. difficile* CT::*tcdC* mutant strains. The absorbance (OD_{600}) was measured over 48 hrs of growth in TY medium. The error bars indicate the standard error of the mean of six experiments.

Western blot analysis, using antibodies against TcdC, confirmed that the isogenic *tcdC* mutant no longer expressed TcdC (Figure 1D). A control blot using antibodies against F_0F_1, indicated that the lack of signal in the TcdC Western blot was not a results of lower amounts of proteins loaded in the lanes of RT 035 (a PaLoc negative strain) and the *tcdC* mutant.

The growth kinetics of the wild type and CT::*tcdC* strains showed no significant differences in various media tested (Figure 1E and data not shown). In TY broth, which does not contain glucose or added cysteine, the wild type strain and the CT::*tcdC* strains showed an exponential growth phase in the first 8 hours post inoculation and after 12 hours post inoculation both strains entered into the stationary growth phase (Figure 1E). Conventional control PCRs confirmed that the disruption of the *tcdC* gene had remained intact during our growth curves experiments (data not shown).

Comparable relative transcription levels of PaLoc genes in wild type and CT::*tcdC*

In order to determine the influence of TcdC on the transcription levels of the PaLoc genes we compared the relative transcription levels of the PaLoc genes of wild type and CT::*tcdC* strains by reverse transcriptase quantitative real-time PCR (RT-qPCR). We found comparable transcription levels of all PaLoc genes in wild type and CT::*tcdC* strains.

Overall, the logarithmic growth phase was associated with lower transcription levels of the PaLoc genes and by entering into the stationary phase increasing transcription levels of PaLoc genes were found, as previously described for *tcdR, tcdE*, *tcdB* and *tcdA* (16,18,19) and *tcdC* (18,19) (Figure 2). The transcription levels of *tcdR* in wild type and CT::*tcdC* strains increased approximately 100-fold between 6 and 24 hours post inoculation (Figure 2A). Though the expression of *tcdR* was, on average, 3-fold higher at the various time points in the CT::*tcdC* strains compared to the wild type, this difference was not statistically significant (Figure 2A, all *p values* ≥ 0.088). Similarly, we observed a 10- to 100-fold increase in the transcription levels of *tcdB* (Figure 2B), *tcdE* (Figure 2C), *tcdA* (Figure 2D) and *tcdC* (Figure 2E) when comparing values from the logarithmic growth phase with those observed in the stationary growth phase. The expression levels of *tcdB*, *tcdE*, *tcdA* and *tcdC* were, on average, 1.5-fold, 2.5-fold, 1.4-fold and 1.7-fold higher, respectively, in the CT::*tcdC* strains compared to the wild type. With one exception, these differences were not found to be significant. The type level at 8 hours post inoculation (Figure 2D). However, no significant differences are found between the wild type and CT::*tcdC* strains at any of the other time points. transcription level of *tcdB* in the CT::*tcdC*1 strain is significantly (*P*=0.046) higher compared to wild Therefore, we

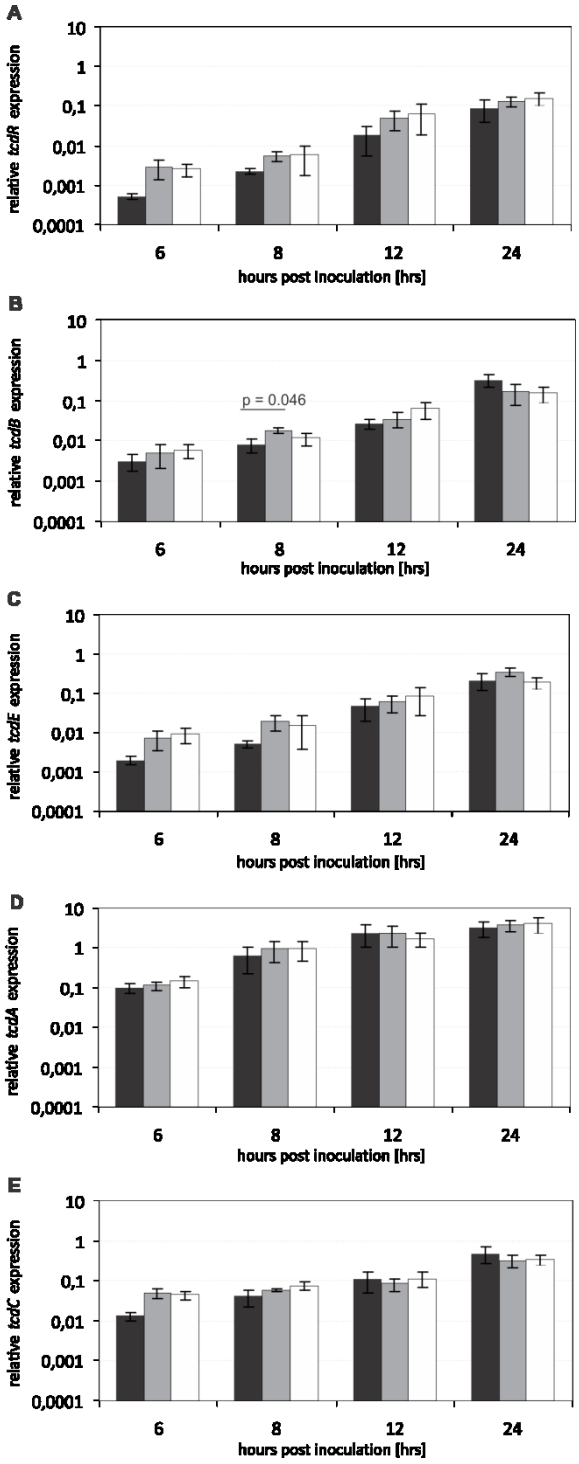


Figure 2: The relative PaLoc gene expression profiles of wild type and CT::*tcdC* in time. The error bars indicate the standard error of the mean (n=6). The asterisk (*) indicate a significant difference between wild type and CT::*tcdC* strain. Values are normalized to *rpsJ* expression. Wild type corresponds to black bars, CT::*tcdC*1 mutant strains to gray bars, CT::*tcdC*2 to the white bars.
(A) The relative expression of *tcdR*.
(B) The relative expression of *tcdB*.
(C) The relative expression of *tcdE* (
D) The relative expression of *tcdA*.
(E) The relative expression of *tcdC*.

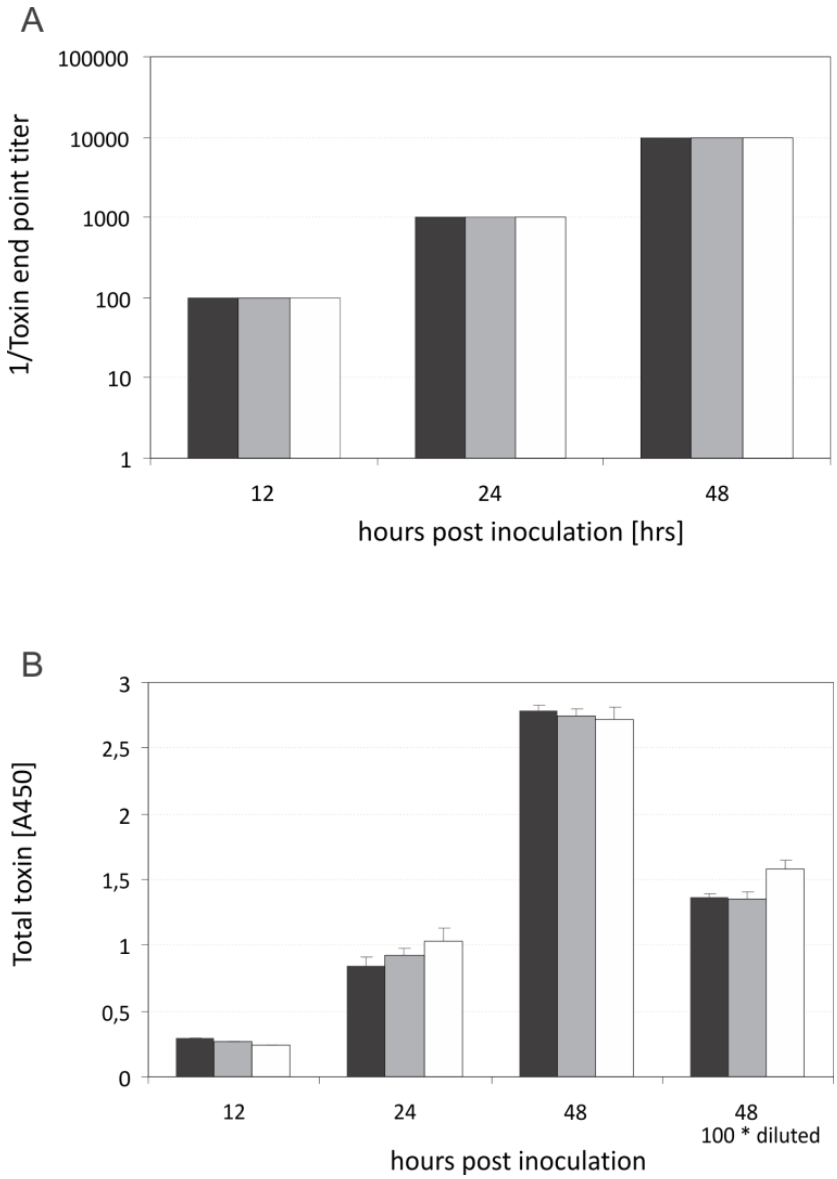


Figure 3. The toxin production profiles of wild type and CT::*tcdC* mutant strains in time. Wild type corresponds to black bars, CT::*tcdC*1 mutant strains to gray bars, CT::*tcdC*2 to the white bars. Total toxin amounts were quantified by using two independent assays. **(A)** The supernatants were incubated in a ten fold dilutions series on Vero cell monolayers. After 24 hrs the cytotoxic effects were quantified by determing the toxin end point titer. Values are given as means (n=6). **(B)** An enzyme immunoassay was used for direct quantification of the secreted toxins according manufacters protocol. The supernatants of 12 and 24 hours post inoculation were 10 times diluted. The supernatants of 48 hours post inoculation were diluted 10 and 100 times. Values are given as means ± standard error of the mean (n=6).

conclude that the disruption of the *tcdC* gene does not result in a consistently and significantly increased transcription level of the PaLoc genes.

Comparable toxin expression in wild type and CT::*tcdC*

Considering the small increase in PaLoc gene expression in the CT::*tcdC* mutants observed in the RT-qPCR experiments, we were interested to see if this difference translated into higher toxin levels. We determined toxin levels using two independent assays, but found no consistent difference between wild type and mutant cells. First, filter sterilized bacterial supernatants were incubated on a Vero cell (a kind gift of Dr. E.J. Snijder (27)) monolayer and cytotoxic effects were quantified after 24 hours by determining the end-point titer (Figure 3A) (26). In the exponential growth phase (5 and 8 hours post inoculation) no cytotoxic effects were detectable (data not shown). In the stationary growth phase (12, 24 and 48 hours post inoculation) we observed increasing cytotoxic effects, indicative of the presence of toxin. Importantly, the observed cytotoxic effects were specific for *C. difficile* Toxin A and B, as a pre-incubation of the filter sterilized bacterial supernatants with anti-toxin, a polyclonal antibody against Toxin A and Toxin B (Techlab), resulted in complete neutralization of cytotoxic effects on the Vero cells at all time points (data not shown). The *tcdC* mutant strains showed no significant differences in toxin levels compared to the wild type strain (Figure3A). Next, we used an enzyme immunoassay (Ridascreen, Biopharma) for the direct detection and relative quantification of the secreted toxins. In the exponential growth phase (5 and 8 hours post inoculation) no toxins were detectable (data not shown), consistent with the lack of toxicity

Towards Vero cells described above. In the stationary growth phase (12, 24 and 48 hours post inoculation) increasing toxin levels were detectable. When we compared the toxin levels at various time points, there were equal amounts of toxins in the wild type and *tcdC* mutant strains. We conclude that the disruption of the *tcdC* gene does not result in consistently and significantly increased toxin levels.

Discussion

C. difficile infections caused by the (hyper-)virulent RT 027 (NAP1/REA B1) and RT 078 (NAP7/REA BK) are associated with an increased morbidity and severity of disease compared to other types (1,3). This increase is suggested to be linked to toxin hyper production (3,22,28). A potential mechanism by which this could occur is through inactivation of a negative regulator of the toxin gene transcription. TcdC has been identified as a negative regulator of toxin production (12).

In the currently prevailing model, a major role for TcdC in the repression of toxin genes has been proposed on the basis of three lines of evidence. First, in *C. difficile* VPI10463 (a high toxin producing strain that also expresses high levels of TcdC (18,29)), an inverse correlation between the transcription of *tcdC* and the genes encoding the toxins is found (16,18,29). This correlation for TcdC is also observed in protein levels (20). Second, elegant *in vitro* experiments have established that heterologously produced and purified TcdC protein can interfere with TcdR-mediated transcription of toxin genes in a way that is not yet fully understood (12). Finally, a frame shift mutation (Δ117 bp) in *tcdC*, that results in a non-functional protein, is associated with increased toxin production in certain (hyper)virulent strains (15,22).

Recently, it was reported that the introduction of a functional *tcdC* gene from a high toxin-producing strain that lacks any of the hyper virulence associated *tcdC* mutations (VPI10463, RT 087) into an epidemic strain carrying a non-functional *tcdC* (M7404, RT 027/NAP1/REA B1) can reduce toxin expression levels and moderately attenuate virulence (30). This data seems consistent with the model discussed above. However, it is unclear how the levels of TcdC in the complemented strain relate to the physiological levels of the protein prior to the inactivation of TcdC in this strain background. The introduced *tcdC* gene, including its transcription signals, was derived from a different genetic background (VPI10463, RT 087) and was introduced on a multicopy plasmid. In addition, the reintroduction of TcdC in a strain lacking a functional TcdC, may affect processes that are not normally affected. Finally, the experiments were not corrected for the additional copies of the *tcdC* promoter that could result in the titration of regulators binding to those sequences.

In an alternative approach that addresses many of the issues above, the role of *tcdC* in toxin expression could be addressed by removing it from a background in which it is normally functional. To this end, we generated an isogenic ClosTron-based *tcdC* mutant strain that could be directly compared to its wild type counterpart, in which the TcdC protein was expected to be functional. Our data obtained with these mutant strains show that TcdC does not exert a major or even significant effect on the transcription of the PaLoc genes or the expression levels of the toxins under the conditions tested.

Our experiments were performed in a glucose free TY broth medium, since glucose is a known repressor of toxin production (17). Indeed, we observed earlier and higher levels of toxin production in TY broth than in the commonly used Brain-Heat-Infusion broth (BHIS) based media, that does contain low amounts of glucose (0.2%) and to which frequently cysteine is added. However, also in BHIS we did not observe a significant effect of a *tcdC* deletion on toxin expression (data not shown).

We controlled critical parameters in our experiments by performing conventional PCRs which confirmed that the disruption of *tcdC* remained intact throughout the growth curve. Western blot analysis with antibodies raised against a TcdC epitope confirmed that the disruption of the *tcdC* gene resulted in the absence of TcdC protein (Figure 1D). The disruption of the *tcdC* gene did not affect the growth kinetics compared to the wild type strain (Figure 1E).

In the RT-qPCR experiments, sample to sample variation was corrected by normalizing to the reference gene *rpsJ* (31). The *rpsJ* gene was selected for normalization, since *rpsJ* was overall the highest ranked reference gene regarding gene expression stability (31).Reverse transcription was carried out using random hexamers, to prevent gene specific biases (32). PCR efficiency in the qPCR was determined using a standard curve for each gene, enabling post run correction (33). To obtain objective data concerning the quantification of the secreted toxins, we used an end point titer assay and an enzyme immunoassay rather than a manual (subjective) cell scoring system (26).

The trends observed in the transcription of the PaLoc genes and the expression of the toxins generally conform to previously reported data (18,19). It should be noted that the up-regulation in time of *tcdC* transcription was not observed in earlier studies on *C. difficile* VPI10463 (16) but is consistent with more recent reports (18,19). We observed an increase in transcription of the PaLoc genes in time, and a concomitant increase in toxicity of culture supernatant in stationary phase that can be attributed to the toxins as it is fully neutralized by anti-toxin against Toxin A and B.

The disruption of the *tcdC* gene resulted in an on average 1.7 fold higher transcription level of *tcdC* in time compared to the wild type strain, although this difference was not found to be statistically significant. It should be noted that we detect these differences because the real time PCR probe detects a region of the gene upstream of the ClosTron insertion site (Figure 1A). This finding might indicate some kind of feedback mechanism on TcdC expression. Similar to *tcdC* gene expression, the disruption of *tcdC* resulted in a slightly higher transcription level of the other PaLoc genes, although this was generally not significant. Moreover, the increased transcription level of the toxin genes did not result in a detectable increase in toxin levels as measured with two independent assays.

Based on the paradigm that TcdC is a major suppressor of toxin production we expected precocious and significantly elevated transcription levels of *tcdA*, *tcdB, tcdE* and *tcdR* in the CT::*tcdC* strains compared to wild type. However, our data indicate that TcdC exerts a moderate, if any, effect on the transcriptional levels of the PaLoc genes and the expression of toxins in *C. difficile* 630Δ*erm* under the conditions tested.

Clostridium difficile strain 630Δ*erm* is a derivative of the clinical isolate 630 (34,35), a RT 012 strain. RT 012 strains constitute 4% of the clinically isolated toxinogenic isolates in Europe (7). *Clostridium difficile* 630 (RT 012)-derived strains are commonly used to investigate virulence of mutants (26,36,37).

An independent study, published during the preparation of this manuscript, reached a similar conclusion with respect to the role of TcdC in toxin regulation in *C. difficile* 630Δ*erm* using an allelic exchange technique (38). In that paper reintroduction of a single functional copy of *tcdC* at its native locus did not affect toxin production in strain R20291 either (38). R20291 is a strain from problematic RT 027 (NAP1/REA B1) that was isolated following an outbreak in Stoke Mandeville, UK.

Our work and that of Cartman and coworkers (38) seem at odds with the previous reports that clearly demonstrate that TcdC can act as a repressor for toxin gene expression (12,30). However, we cannot exclude the possibility that TcdC exerts a more profound effect under specific conditions, or in other strains of *C. difficile* than 630Δ*erm* and R20291. It should be clear though that *in vivo* relevance of TcdC for toxin regulation in these two strains is limited.

In conclusion, we suggest that TcdC might have a modulatory role in regulating toxin expression, and that TcdC functionality is therefore not a major determinant of the virulence of *C. difficile*. This is supported by the lack of correlation between virulence, toxin production and *tcdC* gene variants that was noted by several other studies (18,19,30,39).

Reference List

1. Goorhuis, A., K. T. van der, N. Vaessen, F. W. Dekker, B. R. van den, C. Harmanus, H. S. van den, D. W. Notermans, and E. J. Kuijper. 2007. Spread and epidemiology of *Clostridium difficile* polymerase chain reaction Ribotype 027/toxinotype III in The Netherlands. Clin.Infect.Dis. 45:695-703.
2. Kuijper, E. J., B. Coignard, and P. Tull. 2006. Emergence of *Clostridium difficile*-associated disease in North America and Europe. Clin.Microbiol.Infect. 12 Suppl 6:2-18.
3. Loo, V. G., L. Poirier, M. A. Miller, M. Oughton, M. D. Libman, S. Michaud, A. M. Bourgault, T. Nguyen, C. Frenette, M. Kelly, A. Vibien, P. Brassard, S. Fenn, K. Dewar, T. J. Hudson, R. Horn, P. Rene, Y. Monczak, and A. Dascal. 2005. A predominantly clonal multi-institutional outbreak of *Clostridium difficile*-associated diarrhea with high morbidity and mortality. N.Engl.J.Med. 353:2442-2449.
4. Kuijper, E. J., F. Barbut, J. S. Brazier, N. Kleinkauf, T. Eckmanns, M. L. Lambert, D. Drudy, F. Fitzpatrick, C. Wiuff, D. J. Brown, J. E. Coia, H. Pituch, P. Reichert, J. Even, J. Mossong, A. F. Widmer, K. E. Olsen, F. Allerberger, D. W. Notermans, M. Delmee, B. Coignard, M. Wilcox, B. Patel, R. Frei, E. Nagy, E. Bouza, M. Marin, T. Akerlund, A. Virolainen-Julkunen, O. Lyytikainen, S. Kotila, A. Ingebretsen, B. Smyth, P. Rooney, I. R. Poxton, and D. L. Monnet. 2008. Update of *Clostridium difficile* infection due to PCR Ribotype 027 in Europe, 2008. Euro.Surveill 13.
5. Paltansing, S., R. J. van den Berg, R. A. Guseinova, C. E. Visser, d. van, V, and E. J. Kuijper. 2007. Characteristics and incidence of *Clostridium difficile*-associated disease in The Netherlands, 2005. Clin.Microbiol.Infect. 13:1058-1064.
6. Pepin, J., L. Valiquette, M. E. Alary, P. Villemure, A. Pelletier, K. Forget, K. Pepin, and D. Chouinard. 2004. *Clostridium difficile*-associated diarrhea in a region of Quebec from 1991 to 2003: a changing pattern of disease severity. CMAJ. 171:466-472.
7. Bauer, M. P., D. W. Notermans, B. H. van Benthem, J. S. Brazier, M. H. Wilcox, M. Rupnik, D. L. Monnet, J. T. van Dissel, and E. J. Kuijper. 2011. *Clostridium difficile* infection in Europe: a hospital-based survey. Lancet. 377:63-73.
8. Goorhuis, A., D. Bakker, J. Corver, S. B. Debast, C. Harmanus, D. W. Notermans, A. A. Bergwerff, F. W. Dekker, and E. J. Kuijper. 2008. Emergence of *Clostridium difficile* infection due to a new hypervirulent strain, polymerase chain reaction Ribotype 078. Clin.Infect.Dis. 47:1162-1170.
9. Wilcox, M. H. *Clostridium difficile* Ribotyping Network for England and Northern Ireland. 209.
10. Just, I. and R. Gerhard. 2004. Large clostridial cytotoxins. Rev.Physiol Biochem. Pharmacol. 152:23-47.
11. Mani, N. and B. Dupuy. 2001. Regulation of toxin synthesis in *Clostridium difficile* by an alternative RNA polymerase sigma factor. Proc.Natl.Acad.Sci.U.S.A 98:5844-5849.
12. Matamouros, S., P. England, and B. Dupuy. 2007. *Clostridium difficile* toxin expression is inhibited by the novel regulator TcdC. Mol.Microbiol. 64:1274-1288.
13. Voth, D. E. and J. D. Ballard. 2005. *Clostridium difficile* toxins: mechanism of action and role in disease. Clin.Microbiol.Rev. 18:247-263.
14. Tan, K. S., B. Y. Wee, and K. P. Song. 2001. Evidence for holin function of *tcdE* gene in the pathogenicity of *Clostridium difficile*. J.Med.Microbiol. 50:613-619.

15. Curry, S. R., J. W. Marsh, C. A. Muto, M. M. O'Leary, A. W. Pasculle, and L. H. Harrison. 2007. *tcdC* genotypes associated with severe TcdC truncation in an epidemic clone and other strains of *Clostridium difficile*. J.Clin.Microbiol. 45:215-221.
16. Hundsberger, T., V. Braun, M. Weidmann, P. Leukel, M. Sauerborn, and C. von Eichel-Streiber. 1997. Transcription analysis of the genes *tcdA-E* of the pathogenicity locus of *Clostridium difficile*. Eur.J.Biochem. 244:735-742.
17. Dupuy, B. and A. L. Sonenshein. 1998. Regulated transcription of *Clostridium difficile* toxin genes. Mol.Microbiol. 27:107-120.
18. Merrigan, M., A. Venugopal, M. Mallozzi, B. Roxas, V. K. Viswanathan, S. Johnson, D. N. Gerding, and G. Vedantam. 2010. Human hypervirulent *Clostridium difficile* strains exhibit increased sporulation as well as robust toxin production. J.Bacteriol. 192:4904-4911.
19. Vohra, P. and I. R. Poxton. 2011. Comparison of toxin and spore production in clinically relevant strains of *Clostridium difficile*. Microbiology 157:1343-1353.
20. Govind, R., G. Vediyappan, R. D. Rolfe, and J. A. Fralick. 2006. Evidence that *Clostridium difficile* TcdC is a membrane-associated protein. J.Bacteriol. 188:3716-3720.
21. Mani, N., D. Lyras, L. Barroso, P. Howarth, T. Wilkins, J. I. Rood, A. L. Sonenshein, and B. Dupuy. 2002. Environmental response and autoregulation of *Clostridium difficile* TxeR, a sigma factor for toxin gene expression. J.Bacteriol. 184:5971-5978.
22. Warny, M., J. Pepin, A. Fang, G. Killgore, A. Thompson, J. Brazier, E. Frost, and L. C. McDonald. 2005. Toxin production by an emerging strain of *Clostridium difficile* associated with outbreaks of severe disease in North America and Europe. Lancet 366:1079-1084.
23. Karlsson, S., A. Lindberg, E. Norin, L. G. Burman, and T. Akerlund. 2000. Toxins, butyric acid, and other short-chain fatty acids are coordinately expressed and down-regulated by cysteine in *Clostridium difficile*. Infect.Immun. 68:5881-5888.
24. Heap, J. T., O. J. Pennington, S. T. Cartman, G. P. Carter, and N. P. Minton. 2007. The ClosTron: a universal gene knock-out system for the genus *Clostridium* . J.Microbiol. Methods 70:452-464.
25. Heap, J. T., S. A. Kuehne, M. Ehsaan, S. T. Cartman, C. M. Cooksley, J. C. Scott, and N. P. Minton. 2010. The ClosTron: Mutagenesis in *Clostridium* refined and streamlined. J.Microbiol.Methods 80:49-55.
26. Kuehne, S. A., S. T. Cartman, J. T. Heap, M. L. Kelly, A. Cockayne, and N. P. Minton. 2010. The role of toxin A and toxin B in *Clostridium difficile* infection. Nature 467:711-713.
27. Sebaihia, M., B. W. Wren, P. Mullany, N. F. Fairweather, N. Minton, R. Stabler, N. R. Thomson, A. P. Roberts, A. M. Cerdeno-Tarraga, H. Wang, M. T. Holden, A. Wright, C. Churcher, M. A. Quail, S. Baker, N. Bason, K. Brooks, T. Chillingworth, A. Cronin, P. Davis, L. Dowd, A. Fraser, T. Feltwell, Z. Hance, S. Holroyd, K. Jagels, S. Moule, K. Mungall, C. Price, E. Rabbinowitsch, S. Sharp, M. Simmonds, K. Stevens, L. Unwin, S. Whithead, B. Dupuy, G. Dougan, B. Barrell, and J. Parkhill. 2006. The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome. Nat. Genet. 38:779-786.
28. Carter, G. P., G. R. Douce, R. Govind, P. M. Howarth, K. E. Mackin, J. Spencer, A. M. Buckley, A. Antunes, D. Kotsanas, G. A. Jenkin, B. Dupuy, J. I. Rood, and D. Lyras. 2011. The Anti-Sigma Factor TcdC Modulates Hypervirulence in an Epidemic BI/NAP1/027 Clinical Isolate of *Clostridium difficile*. PLoS.Pathog. 7:e1002317.

29. Snijder, E. J., P. J. Bredenbeek, J. C. Dobbe, V. Thiel, J. Ziebuhr, L. L. Poon, Y. Guan, M. Rozanov, W. J. Spaan, and A. E. Gorbalenya. 2003. Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. J.Mol.Biol. 331:991-1004.
30. O'Connor, J. R., S. Johnson, and D. N. Gerding. 2009. *Clostridium difficile* infection caused by the epidemic BI/NAP1/027 strain. Gastroenterology 136:1913-1924.
31. 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.
32. Metcalf, D., S. Sharif, and J. S. Weese. 2010. Evaluation of candidate reference genes in *Clostridium difficile* for gene expression normalization. Anaerobe. 16:439-443.
33. Pfaffl, M. W. 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 29:e45.
34. Graeber, K., A. Linkies, A. T. Wood, and G. Leubner-Metzger. 2011. A guideline to family-wide comparative state-of-the-art quantitative RT-PCR analysis exemplified with a brassicaceae cross-species seed germination case study. Plant Cell. 23:2045-2063.
35. Hussain, H. A., A. P. Roberts, and P. Mullany. 2005. Generation of an erythromycin-sensitive derivative of *Clostridium difficile* strain 630 (630Deltaerm) and demonstration that the conjugative transposon Tn*916*DeltaE enters the genome of this strain at multiple sites. J.Med.Microbiol. 54:137-141.
36. Ho, T. D. and C. D. Ellermeier. 2011. PrsW is required for colonization, resistance to antimicrobial peptides, and expression of extracytoplasmic function sigma factors in *Clostridium difficile*. Infect.Immun. 79:3229-3238.
37. Lyras, D., J. R. O'Connor, P. M. Howarth, S. P. Sambol, G. P. Carter, T. Phumoonna, R. Poon, V. Adams, G. Vedantam, S. Johnson, D. N. Gerding, and J. I. Rood. 2009. Toxin B is essential for virulence of *Clostridium difficile*. Nature 458:1176-1179.
38. Cartman, S. T., M. L. Kelly, D. Heeg, J. T. Heap, and N. P. Minton. 2012. Precise Manipulation of the *Clostridium difficile* Chromosome Reveals a Lack of Association between the *tcdC* Genotype and Toxin Production. Appl.Environ.Microbiol. 78:4683-4690.
39. Murray, R., D. Boyd, P. N. Levett, M. R. Mulvey, and M. J. Alfa. 2009. Truncation in the *tcdC* region of the *Clostridium difficile* PathLoc of clinical isolates does not predict increased biological activity of Toxin B or Toxin A. BMC.Infect.Dis. 9:103.
40. Garnier, T. and S. T. Cole. 1986. Characterization of a bacteriocinogenic plasmid from *Clostridium* perfringens and molecular genetic analysis of the bacteriocin-encoding gene. J.Bacteriol. 168:1189-1196.
41. Purdy, D., T. A. O'Keeffe, M. Elmore, M. Herbert, A. McLeod, M. Bokori-Brown, A. Ostrowski, and N. P. Minton. 2002. Conjugative transfer of clostridial shuttle vectors from Escherichia coli to *Clostridium difficile* through circumvention of the restriction barrier. Mol.Microbiol. 46:439-452.
42. Knetsch, C. W., E. M. Terveer, C. Lauber, A. E. Gorbalenya, C. Harmanus, E. J. Kuijper, J. Corver, and H. C. van Leeuwen. 2012. Comparative analysis of an expanded *Clostridium difficile* reference strain collection reveals genetic diversity and evolution through six lineages. Infect.Genet.Evol.
43. Wren, B. W. and S. Tabaqchali. 1987. Restriction endonuclease DNA analysis of *Clostridium difficile*. J.Clin.Microbiol. 25:2402-2404.

44. Dawson, L. F., E. H. Donahue, S. T. Cartman, R. H. Barton, J. Bundy, R. McNerney, N. P. Minton, and B. W. Wren. 2011. The analysis of para-cresol production and tolerance in *Clostridium difficile* 027 and 012 strains. BMC.Microbiol. 11:86.:86.
45. Deckers-Hebestreit, G. and K. Altendorf. 1986. Accessibility of F0 subunits from *Escherichia coli* ATP synthase. A study with subunit specific antisera. Eur.J.Biochem. 161:225-231.

7

Clostridium difficile TcdC protein binds four stranded G- quadruplex structures

H C. van Leeuwen[1], Dennis Bakker[1], Philip Steindel[2], Ed J. Kuijper[1], Jeroen Corver[1]

[1] *Department of Medical Microbiology, Center of Infectious Diseases, Leiden University Medical Center, the Netherlands.*
[2] *Department of Biochemistry, Brandeis University, MS009, 415 South Street, Waltham, MA 02454*

Nucleic Acids Res 2013

Abstract

Clostridium difficile infections are increasing worldwide due to emergence of virulent strains. Infections can result in diarrhea and potentially fatal pseudomembranous colitis. The main virulence factors of *C.difficile* are clostridial toxins TcdA and TcdB. Transcription of the toxins is positively regulated by the sigma factor TcdR. Negative regulation is believed to occur through TcdC, a proposed anti-sigma factor. Here, we describe the biochemical properties of TcdC in order to understand the mechanism of TcdC action.

Bioinformatic analysis of the TcdC protein sequence predicted the presence of a hydrophobic stretch (amino-acids 30-50), a potential dimerization domain (aa 90-130) and a C-terminal OB-fold (oligonucleotide-binding).

Gel filtration chromatography of two truncated recombinant TcdC proteins (TcdCΔ1-89 and TcdCΔ1-130), showed that the domain between aa 90 and 130 is involved in dimerization.

Binding of recombinant TcdC to single-stranded DNA was studied using a single-stranded SELEX approach. This involved specific binding of ss-DNA sequences from a pool of random oligonucleotides, as monitored by electrophoretic-mobility shift assay. Analysis of the oligonucleotides bound showed that the OB-fold domain of TcdC can bind specifically to DNA folded into G-quadruplex structures containing repetitive guanine nucleotides forming a four-stranded structure.

In summary, we provide evidence for DNA binding of TcdC, which suggests an alternative function for this proposed anti-sigma factor.

Introduction

Clostridium difficile is a spore forming, anaerobic bacterium that can cause antibiotic associated diarrheal disease in humans. In the past decade, the incidence, complications and mortality of *Clostridium difficile*-associated infection (CDI) have increased dramatically due to the emergence of new hypervirulent PCR Ribotypes (RT) (1-4). Virulence of *C.difficile* has been linked to the production of two toxin molecules, Toxin A and Toxin B, which are encoded within the Pathogenicity Locus (PaLoc). These toxins cause intestinal damage and ultimately clinical disease (5). Both toxins have the same enzymatic activity. Upon entering intestinal epithelial cells they catalyze the transfer of glucose onto the Rho family of GTPases, leading to reorganization of the actin cytoskeleton, complete rounding of cells and destruction of the intestinal barrier function. This causes diarrhea and in some cases may lead to a severe inflammatory response and pseudomembranous colitis.

The mechanisms that regulate the levels of toxin synthesis are slowly being unraveled. Toxin genes, *tcdA* and *tcdB*, are located on the PaLoc together with two regulatory genes *tcdR* and *tcdC*, and *tcdE*, which encodes a holin like protein that may facilitate the release of the toxins into the extracellular environment (6). TcdR has been demonstrated to activate gene expression of both toxins as a specific RNA polymerase sigma factor belonging to the subgroup of extracytoplasmic function (ECF) σ^{70}-family of RNA polymerase sigma factors (7). Members of this group include several σ-factors involved in positive regulation of potent toxins such as botulinum neurotoxin (BotR of *Clostridium botulinum*) and tetanus neurotoxin (TetR of *Clostridium tetani*) (8). Toxin expression is also influenced by the nutritional status of the bacteria; a rapidly metabolizable carbon source such as glucose inhibits toxin expression (9). In addition, general regulatory molecules, such as CodY and CcpA are known to influence toxin synthesis (10;11). In *C.difficile*, TcdR not only stimulates toxin gene transcription but also activates its own expression suggesting a large overshoot in protein expression once activated (7). A negatively acting mechanism therefore is required to put a limit on this system during unrestricted growth of *C.difficile*.

Activation of bacterial gene expression by specific σ-factors is often subject to control by specific antagonists, called anti-sigma factors (12). Generally they sequester their cognate σ factor, preventing it from interacting with the RNA polymerase. Encoded within the PaLoc is TcdC, which has been postulated to act as an anti-sigma factor and negatively regulate toxin production. TcdC transcription pattern was reported to be inverse to TcdR and the toxins as it is highly transcribed and expressed during the exponential growth phase, while its expression is strongly reduced as the growth rate slows in stationary phase (13). This inverse correlation

suggested that TcdC interferes with toxin gene expression. However, more recent studies have shown that this inverse correlation cannot be confirmed using quantitative RT-PCR (14-16). This suggests that TcdC may not be as important in toxin regulation as previously thought.

A direct inhibitory effect on transcription of *tcdC* has been shown *in vitro*. The TcdR-RNA-polymerase-DNA complex is destabilized by TcdC, preventing initiation of transcription. However, once a stable open complex is formed with the promoter no inhibition by TcdC occurs (17). The target of TcdC in (prevention of) complex formation is unclear; no interaction with the TcdR-RNAP-complex was found nor does TcdC bind to dsDNA in the promoter, suggesting a potentially unique inhibitory mechanism of TcdC.

Recent *in vivo* studies on the importance of TcdC on toxin expression show contradictory results. *TcdC* complementation of strain M7404, a toxinogenic strain that lacks a functional *tcdC* gene, results in a reduced amount of produced toxin and an attenuated phenotype in hamsters (18). In contrast, complementation of strain R20291, another strain that lacks a functional *tcdC* gene, with a functional *tcdC* gene did not alter the toxin titers (19). In addition, knockout of *tcdC* in strain 630Δ*erm* did not result in an increased level of toxins produced, nor did it result in increased toxin mRNA production (14).

Since the suggested anti sigma function of TcdC is not undisputed and because the mechanism by which TcdC is supposed to inhibit TcdR mediated transcription is unknown, we aimed to further characterize the biochemical properties of TcdC. Through *in silico* analyses we found that TcdC contains a predicted single-stranded nucleic acid binding fold (OB-fold). In this paper we show for the first time through a combination of *in silico* analysis and biochemical experiments that TcdC can bind to nucleic acids.

Material and Methods

Construction of plasmids

To construct his10-tagged TcdC expression plasmids, the sequence was amplified by PCR from *C.difficile* strain 630 genomic DNA, using specific primers, see Table 1. The PCR products were digested with NdeI and XhoI or NdeI and BamHI and ligated into pET16b (Novagen) similarly digested with NdeI and XhoI/BamHI. This resulted in the construction of TcdC expression vectors containing a 10-His-tag at its N-terminus.

Table 1: Primers used to generate bacterial expression constructs.

Delta1-50 TcdC
Forward primer, TATGCATATGGGATATGATACTGGTATTAC Reverse primer, TTTTCTCGAGTTAATTAATTTTCTCTACAGCTATCCC
Delta1-89 TcdC
Forward primer, GTTCCATATGAAAGACGACGAAAAGAAAGCTATTG Reverse primer as for Delta1-50 TcdC
Delta1-130 TcdC
Forward primer, TATGCATATGGGATATGATACTGGTATTAC Reverse primer as for as for Delta1-50 TcdC
Delta Delta1-89; Delta208-232 TcdC (90-207)
Forward primer as for Delta1-89 TcdC Reverse primer, TACTGGATCCTTTAAGCACTTATACCTCTTATAG

DNA binding studies

Probes used for band shift assays were obtained from Eurogentec (Maastricht, The Netherlands), end labeled with T4-polynucleotide kinase and ^{32}P-γ-ATP and purified using Micro Bio-Spin Colums P-30 Tris RNase Free (Biorad) according to manufacturer's instructions. Binding reactions were carried out for 60 minutes on ice in 20 ul binding buffer (20 mM Hepes-KOH pH 7.5, 50 mM NaCl, 40 mM KCl, 7% glycerol, 1 mM EDTA, 0.1 mM DTT and 0.25 pmol probe (12.5 nM). Free DNA and protein-DNA complexes were separated on a 7% polyacrylamide gel (37.5:1) run in 0.5xTBE. Dried gels were exposed to a Biorad phosphoimaging screen-K and scanned on a Typhoon 9410 from GE Healthcare. The equilibrium dissociation constant (Kd) was calculated at half saturation Kd=Pt-Db (Db=DNA bound, 6 nM). The Pt (total protein concentration) was calculated using a deduced molecular mass of 18.8 kD for the His10-TcdCΔ1-89 protein.

Single stranded SELEX

The random site used in the first selection round was 5' A G T G C A G T G G A T C C T G T C G - N N N N N N N N N N N N N N N - AGGCGAATTCAGTCCAAGTG 3', ^{32}P-labelled at the 5'end. Binding reactions were as described above with 10 or 100 ng purified TcdC Δ1-89. In the first selection round 10 ul of 50% Cobalt^{2+}-beads (Clontech) in binding buffer was added to the binding reaction and incubated for 30 min at 4°C under continuous rotation. Subsequently the beads were spun down (30 sec. at 100 g) and washed three times in 100 ul binding buffer. Bound random oligo was eluted in 20 ul binding buffer with 250 mM Imidazol and subsequent heating heating 10 minutes at 95 °°C. Beads were spun down and supernatant collected. 5 ul of supernatant was amplified

using 30 PCR cycles with primers, AGTCAGTGCAGTGGATCCTGTCG (forward) and CACTTGGACTGAATTCGCCTC (reverse). The resulting 53 bp PCR product was digested with N.BtsI nicking endonuclease (New England Biolabs) and labeled using ^{32}p-γ-ATP. Digested and labeled probe was separated on a 12% polyacrylamide gel (19:1). Following exposure of the wet gel to X-ray film (Fujifilm, super RX), the highest band corresponding to the uncleaved top stand was cut out and eluted according to the QIAEX II Gel Extraction Kit for polyacrylamide Gels (Qiagen). 5 ul of the extracted probe was used in a next round of selection using bandshift assay. The Protein-DNA complexes were separated on a 7% polyacrylamide gel as described above. After exposure to film the bound probe (shift) was cut out and eluted for an additional round of selection.

Quadruplex staining

Specific staining of quadruplex forming DNA was performed according to Yang *et al.* (20). Briefly, polyacrylamide gels were incubated in 20 uM ETC ($C_{39}H_{47}N_3O_6S_4$, Organica Feinchemie GmbH Wolfen) in phosphate buffer saline (PBS) for 30 minutes. Rinsed five times with water and then scanned in a Typhoon 9410 (GE Healthcare), excitation 532 nm and emission 610 nm.

Bioinformatics analysis

For all bioinformatic analyses protein sequence Q189K7 (*C.difficile* strain 630) was used. Predictions of coiled-coil helices were carried out using the Multicoil Scoring Form ((21); http://groups.csail.mit.edu/cb/multicoil/cgi-bin/multicoil.cgi). All predictions were performed using standard settings. All sequence alignments were performed by use of Clustal Omega - Multiple Sequence Alignment, available from EMBL-EBI European bioinformatics institute (http://www.ebi.ac.uk/Tools/msa/clustalo/; Sievers *et al.*, Mol Syst Biol. 2011).

Structural models of the TcdC conserved C-terminal domain were generated by the automated I-TASSER (threading, assembly and refinement) simulation method; http://zhanglab.ccmb.med.umich.edu/I-TASSER/ (22;23). Predictions were done using the standard parameters. As part of sequence homology detection, the protein alignment was analyzed using HHpred at the Max-Planck Institute for Developmental Biology (http://toolkit.tuebingen.mpg.de/hhpred; (24)). Predictions were done using the following parameters: Selected, database pfamA_v26.0; Max. MSA Generation iterations, 0. Other parameters set at default.

Protein purification

E.coli (BL21) lysates (50 mM sodium phosphate buffer, pH 8.0, 5 mM beta-mercaptoethanol, 0.1% NP40, 300 mM NaCl) containing histidine tagged proteins

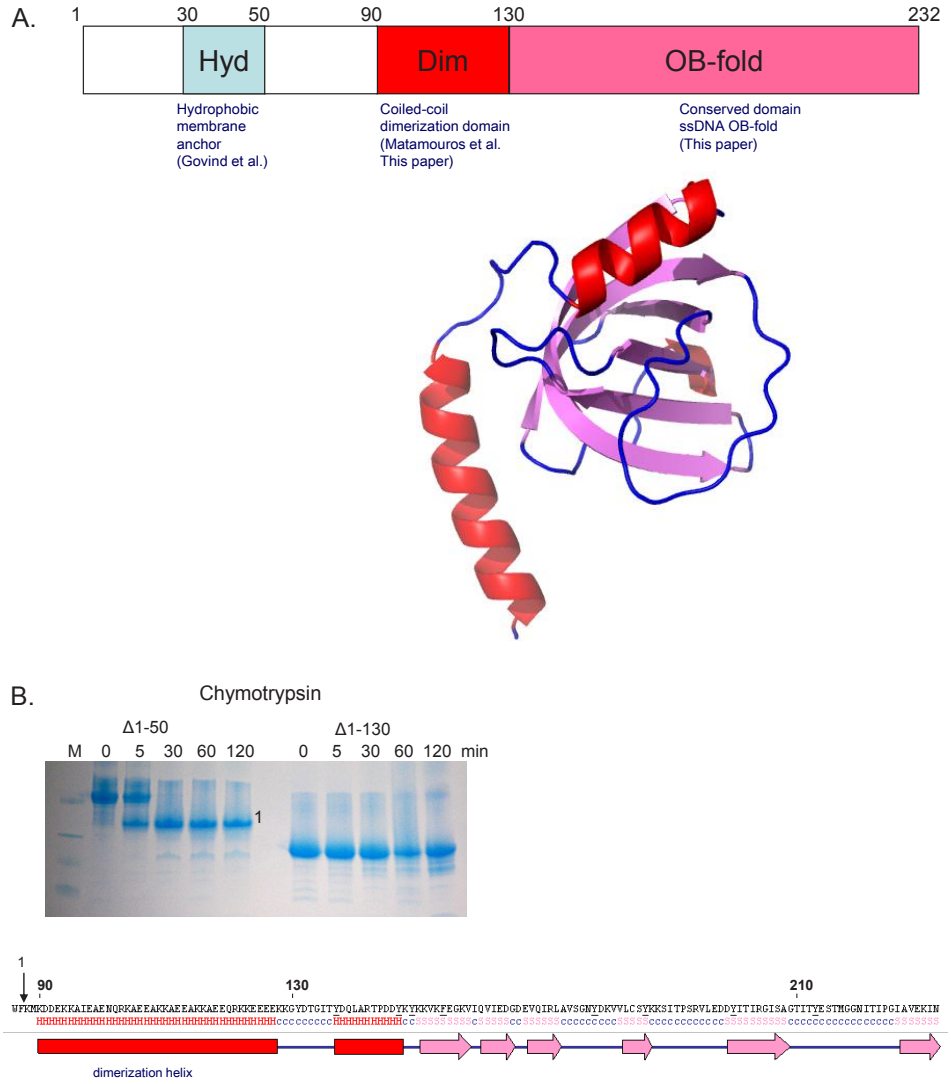


Figure 1: Domain organization of *Clostridium difficile* TcdC. **(A)**. Schematic presentation of TcdC with three predicted domains; Hyd= hydrophobic membrane anchor, Dim:dimerization domain and OB fold = conserved C-terminal domain containing a predicted oligonucleotide binding fold. Bottom, spatial structure of the TcdC- C-terminal domain (aa 90-232) predicted by the automated simulation method I-TASSER (Roy *et al.*, 2010) and represented with a ribbon diagram. N and C indicate amino-terminus and carboxy-terminus. **(B)** Limited proteolysis of TcdC⊿1-50 and TcdC⊿1-130. Histidine-tagged proteins were digested with chymotrypsin for 5, 30, 60, or 120 min, and the fragments were resolved by SDS-PAGE. Amino-terminal sequences were identified using an Edman degradation. Site of chymotrypsin cleavage (arrow 1) indicated in the primary amino acids sequence and TcdC predicted secondary structure elements. H=helix (red box), C=coil (blue line), S=sheet (pink arrow).

were loaded on a 1 mL Ni-NTA column (Qiagen). The column was washed with 20 mL wash buffer (50 mM sodium phosphate buffer pH 7.0, 300 mM NaCl, 5 mM mercaptoethanol, 5% glycerol, 20 mM Imidazol). The His-tagged proteins eluted at ~200 mM imidazole when employing a 25 mL linear gradient ranging from 20 to 250 mM imidazole. Peak fractions containing the His-tagged proteins were pooled and 200ul loaded onto a superdex 75 gelfiltration column equilibrated and run in 50 mM sodium phosphate buffer pH 7.0, 150 mM NaCl, 5 mM mercaptoethanol, 5% glycerol. Protein concentrations were measured and peak fractions were used for DNA binding studies.

Results

In silico analysis of TcdC

In many cases sequence similarity allows the inference of protein function. At the primary amino acid sequence level, the C-terminal domain of TcdC (residues 130-232, conserved domain, Figure 1A) has sequence identity (conservation) to potential/putative protein homologues from both anaerobic and facultative aerobic members of the Firmicutes phylum (Figure 2). Though several TcdC homologues have been identified, none of them have been characterized biochemically in detail and therefore do not provide a clue to the TcdC mechanism of action.

As the primary sequence gave no indication to its function we used computational protein structure prediction for detecting remote homologous templates. Structural models of the TcdC conserved C-terminal domain were generated by the automated I-TASSER (threading, assembly and refinement) simulation method (22;23). The best model (Figure 1A) was predicted to be composed of a five-stranded closed beta-barrel connected by large loops (C-score = -0.9; C-score is a confidence score for the predicted model with a C-score score above cutoff of − 1.5 is used to select models of correct topology). Matching the best predicted model with proteins from the PDB databasesrevealed a nucleic acid binding OB-fold (oligonucleotide binding, IPR016027) containing domain in all the 10 top matches/best scoring templates (TM = 0.8 – 0.7; TM-score >0.5 indicates a model of correct match topology). The core of the OB-fold forms a surface to bind to single-stranded DNA (ssDNA) or RNA (25;26). Variations in folds, loops and amino-acids in the binding interface determine ligand and sequence specificity. Members of this OB-fold group include proteins critical for DNA replication protein (RPA), DNA recombination (RuvA), translation (tRNA synthetase anticodon binding protein) and telomere-end-binding proteins (hPot-1)(25;26).

I-TASSER folding of the region preceding the conserved domain (aa 90-130) predicted a large helix (Figure 1A) containing many positive and negatively charged amino-acids. Such a helix clearly can form a charged coiled-coil motif with another molecule thereby forming an intertwined dimer as was predicted by Matamouros *et. al.* (17). The coiled-coil prediction was confirmed using the Multicoil Scoring Form (21), which calculated a maximum coiled-coil probability of 0.861 of this region (data not shown).

In addition to the protein structure prediction, we used the TcdC conserved domain protein alignment (Figure 2) for a highly sensitive profile-based search (24;27). Using the TcdC-conserved domain multiple sequence alignment rather than a single sequence as a query increases sensitivity and allows for homology detection of protein families. Pairwise comparison of the TcdC profile with the PFAM database resulted in a hit with PF12869, tRNA_anti-like family containing the nucleic acid-binding OB fold (E-value 1.3e-6, probability 98.3). In summary, these *in silico* analyses clearly suggest that TcdC forms a dimeric ssDNA binding OB-protein fold.

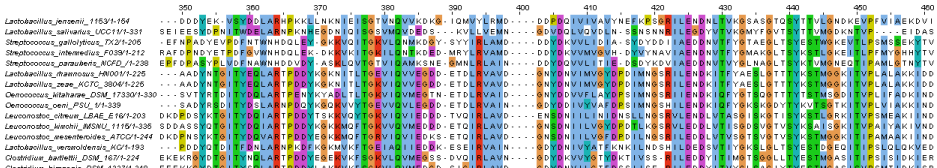


Figure 2: TcdC conserved domain protein alignment with protein homologues from members of the Firmicutes phylum.Numbering top according to alignment; Numbers bottom according to *Clostridium difficile* TcdC amino acids.

Limited proteolysis suggests a folded structure of the TcdC conserved domain

In order to confirm the predicted domains and borders of the nucleic acid binding OB-fold of TcdC, we cloned the *tcdC* gene including 10-histidine codons at the N-terminus into a bacterial expression vector. In order to produce soluble protein expression in *E.coli*, the first 50 amino-acids, which contain the reported hydrophobic membrane anchor (13) (Figure 1A) were removed. His-10 tagged TcdC was overexpressed in *E.coli* (BL21) and purified using a nickel affinity column.

To investigate the local conformation of this TcdC protein, we used limited proteolytic digestion. Protease resistance is an indication of structured protein sequences, as folded structure is usually protected from proteolytic degradation. TcdCΔ1-50 was digested with chymotrypsin, which cleaves after aromatic amino acids. Despite the presence of 11 potential cleavage sites (W, Y, F), chymotrypsin digestion of TcdCΔ1-50 led to only one distinct fragment (Figure 1B). To identify this fragment, the proteolytic product was subjected to N-terminal sequencing using Edman degradation. The identified N-terminal sequence (KMKD) corresponds to residue 88 of TcdC directly adjacent to the large coiled-coil helix. When we tested TcdCΔ1-130 (Figure 1B), corresponding to the OB-fold domain, we observe hardly any cleavage.

Taken together, these studies provide strong support for a folded structure of the TcdC conserved domain, including the dimerization domain, resistant to proteolytic cleavage.

TcdC contains a dimerization domain

Consistent with the proteolytic protection assay and prediction of the coiled-coil dimerisation helix, we constructed an expression vector containing the TcdC conserved domain including this putative dimerisation domain (TcdCaa90-232, here named TcdCΔ1-89). In addition a construct without the dimerization domain (TcdCΔ1-130) was generated. Both proteins were subsequently purified using nickel–affinity chromatography and gel-filtration (Figure 3).

Besides extra purity the latter column allows for separation by size and thus molecular weight estimation (Figure 4). Indeed the apparent molecular weight of TcdCΔ1-89 of 35 kDa, with a predicted molecular weight of 18 kDa, fits a dimeric protein. In contrast, TcdCΔ1-130, with a predicted molecular weight of 14 kDa and apparent molecular weight of 14 kDa, fits a monomeric protein. This confirms that the region between AA 90 – 130 contains a dimerization domain. Dimerization was confirmed using a cross linking with glutaraldehyde which can form stable intersubunit covalent bonds (Figure 4C). This experiment shows that TcdC Δ1-130 forms no visible dimers after crosslinking, while TcdC Δ1-89 forms

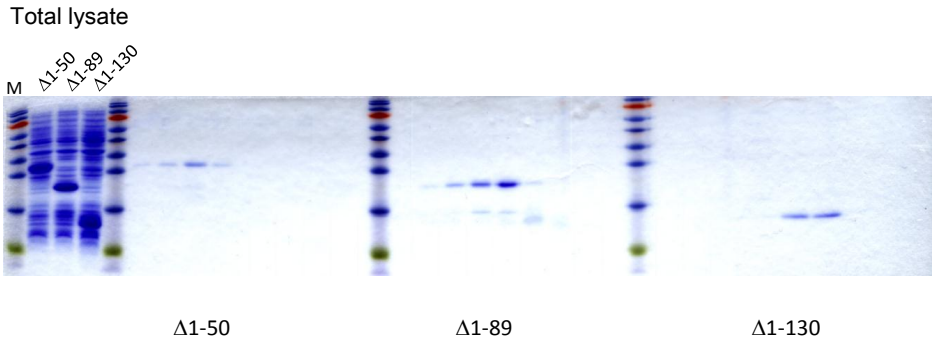


Figure 3. Cobalt2+ agarose and gel filtration purified His10-TcdC Δ1-50 , Δ1-89 and Δ1-130.

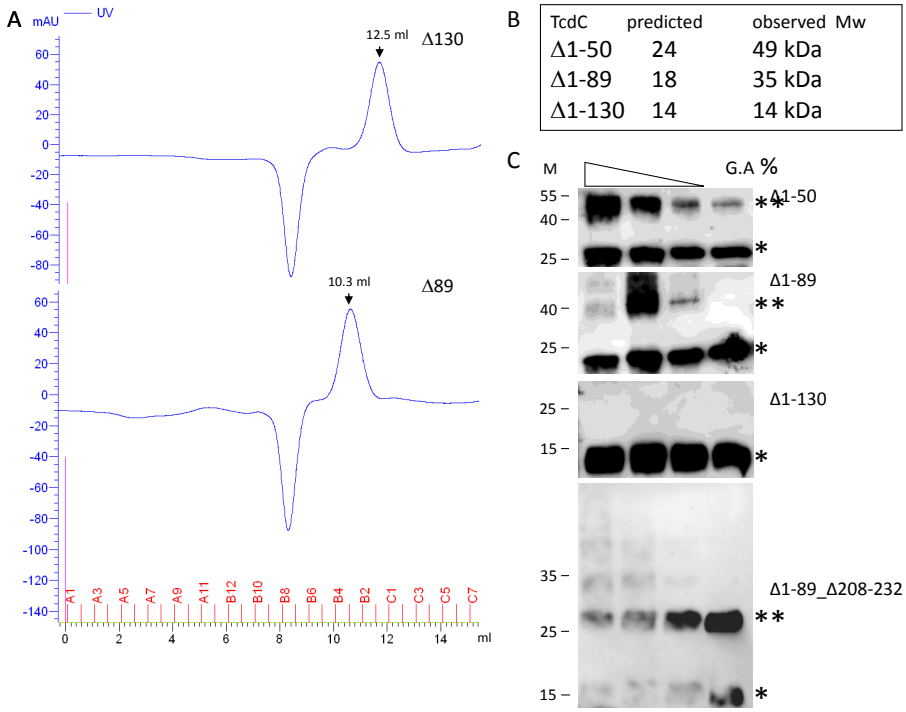


Figure 4: Gel filtration separation and cross linking of TcdC deletion mutants show that the TcdC dimerization domain is located between AA90-130. **(A)** Chromatographic separation of purified TcdC proteins on a superdex 75 column. Arrows indicate the elution points (mL) for the proteins. **(B)** Predicted molecular weight and observed molecular weight on the GF column of the respective deletion mutants. **(C)** TcdC dimerization as evidenced by glutaraldehyde cross-linking. 25 ng of TcdC was incubated for 30 minutes in the presence or absence of glutaraldehyde. Glutaraldehyde (G.A.) concentrations were 0.006% , 0.003%, 0.0006% and 0% (from left to right). After incubation, proteins were analyzed by Western blot analysis, using His-tag specific monoclonal antibody (Abcam). * indicates the bands corresponding to monomeric protein. ** corresponding to the dimer**.**

dimers already at low concentration of glutaraldehyde, thereby confirming the gel-filtration experiments.

TcdC does not bind to *tcdA* promoter elements

Based on existing evidence for the TcdC point of action i.e. destabilizing open complex formation before transcription initiation (17), we tested binding to (ss-) DNA corresponding to the region of the *tcdA* promoter that undergoes melting during transcription (opening of the double helix, resulting in exposed ssDNA) (28;29). Using protein TcdC Δ1-89 in a mobility shift assay, we tested binding to the *tcdA* double stranded promoter (-32 to +22 relative to transcription start), single stranded promoter top strand (non-template), single stranded promoter bottom strand (template), open promoter complex (region -10 to -6 or -13 to +4 open) as well as the *tcdA* mRNA gene transcript (+1 to +22) and the DNA-RNA hybrid (see Table 2). Surprisingly, we found no DNA binding for any of these fragments (data not shown). Also partially double stranded/single stranded overhang (5'and 3') and forked templates of the promoter (see Table 2) showed no binding. Finally, we tested a synthetic Holliday junction (30;31), which can be found at replication origins and recombination junctions, but found no binding.

TcdC binding sites selected through Selex

Because of our unsuccessful attempt to find the TcdC DNA binding site directly, we adapted a single-stranded SELEX (Systematic Evolution of Ligands by EXponential enrichment, (32), a procedure that allows extraction of oligomers with an optimal binding affinity from an initially random pool of oligonucleotides (Figure 5). After site-selection and PCR amplification, single stranded DNA is recreated using (asymmetric -) nicking of the bottom strand of the amplified selected sites followed by denaturation. These sites are subsequently used in an additional selection round thereby increasing the specificity of the selection procedure (Figure 6). Initial selection of his10-TcdCΔ1-89 bound fragments from the pool of ss-oligonucleotides, containing a stretch of 15 random nucleotides was performed through $Cobalt^{2+}$-agarose beads pull down (round 1, Figure 2). Two additional selection rounds were carried out using separation of bound DNA fragments on a polyacrylamide gel (round 2 and 3, Figure 6).

During these selection rounds we observed a higher molecular weight product (HMW) arise, which is bound and shifted by TcdCΔ1-89 (Figure 6). Each round showed a clear enrichment of the amount of higher molecular weight product being bound and shifted in the presence of TcdCΔ1-89. After three rounds of selection the enriched sites were cloned and sequenced. Table 3, shows the individual sites selected by the TcdCΔ1-89 ssSELEX. Most of the sequences selected (17/18)

Table 2: Primers used to test TcdC ss/ds DNA binding.

***tcdA* promoter double strand promoter** 795576..795629
-10 +1 5'CAAATTACTATCAGACAATCT**CCTTAT**CTAAT**A**GAAGAGTCAATTAACTAATTG 3' 3'GTTTAATGATAGTCTGTTAGAGGAATAGATTATCTTCTCAGTTAATTGATTAAC 5'
***tcdA* promoter template strand**
5'CAATTAGTTAATTGACTCTTCTATTAGATAAGGAGATTGTCTGATAGTAATTTG 3'
***tcdA* promoter non-template strand**
5'CAAATTACTATCAGACAATCTCCTTATCTAATAGAAGAGTCAATTAACTAATTG 3'
***tcdA* promoter large open promoter -13 to +4**
CTCCTTATCTAAT**A**GAA 5'CAAATTACTATCAGACAAT GAGTCAATTAACTAATTG 3' 3'GTTTAATGATAGTCTGTTA CTCAGTTAATTGATTAAC 5' CTCCTTATCTAATAGAA
***tcdA* promoter small open promoter -10 to -6**
TTAT 5'CAAATTACTATCAGACAATCT**CC** CTAAT**A**GAAGAGTCAATTAACTAATTG 3' 3'GTTTAATGATAGTCTGTTAGAGG GATTATCTTCTCAGTTAATTGATTAAC 5' TTAT
***tcdA* promoter 3'overhang**
5'CAAATTACTATCAGACAATCTCCTTATCTAAT**A**GAAGAGTCAATTAACTAATTG 3' 3'GTTTAATGATAGTCTGTTAGAGGAATAGATTA 5'
***tcdA* promoter 5'overhang**
5'CAAATTACTATCAGACAATCTCCTTATCTAAT**A**GAAGAGTCAATTAACTAATTG 3' 3' TCTTCTCAGTTAATTGATTAAC 5'
***tcdA* promoter vorked template**
TCTTCTCAGTTAATTGATTAAC 3' 5' CAAATTACTATCAGACAATCTCCTTATCTAAT 3'GTTTAATGATAGTCTGTTAGAGGAATAGATTA TCTTCTCAGTTAATTGATTAAC 5'
***tcdA* RNA transcript +1 to +22**
5'AGAAGAGUCAAUUAACUAAUUG 3'
***tcdA* DNA-RNA hybrid**
5'**A**GAAGAGUCAAUUAACUAAUUG 3' 3'GTTTAATGATAGTCTGTTAGAGGAATAGATTATCTTCTCAGTTAATTGATTAAC 5'
Four way junction
5'GACGCTGCCGAATTCTGGCTTGCTAGGACATCTTTGCCCACGTTGACCC 3' 5'TGGGTCAACGTGGGCAAAGATGTCCTAGCAATGTAATCGTCTATGACGTT 3' 5'CAACGTCATAGACGATTACATTGCTAGGACATGCTGTCTAGAGACTATCGA 3' 5'ATCGATAGTCTCTAGACAGCATGTCCTAGCAAGCCAGAATTCGGCAGCGT 3'

contain a stretch of 3 Gs (highlighted bold) and 2/3rd of the selected clones contains an A nucleotide preceding this G-stretch. 6 out of 18 clones contain two stretches of 3 Gs. Next, we selected two sites obtainedfrom the ssSELEX. One clone with a single aGGG consensus site (clone #2, Table 3) and one with a double consensus site (clone #5, Table 3) and tested these individually on a polyacrylamide gel. As shown in Figure 7A, both clones form the HMW product also observed in the ssSELEX, although in different efficiencies (on average ~90% of #5 formed HMW product compared to ~10% of clone #2). When we tested TcdC binding the HMW band was readily shifted, thereby confirming their efficient binding selection. In contrast to the HMW product, none of the ssDNA product was shifted (Figure 3A). Assuming a simple binding model, gel retardation experiments allow for a quick estimate of the protein equilibrium binding constants (Kd) at half saturation using the formula Kd=Pt-Db (the total protein and DNA concentrations at 50% binding). Quantification of the binding revealed a protein dissociation constant of ~30 nM (Figure 7B). When we assume that TcdC binds as a dimer, the Kd is ~15 nM. Typically, affinity of OB-fold proteins range from 1nM to 100 nM (33). Upon titration of TcdC, at least three complexes can be observed, which suggest multiple binding sites on the quadruplex structure (Figure 7B), with the first shifted complex representing the highest affinity binding site. Alternatively the dimeric protein might bind more than one quadruplex structure (each monomer bindings separate quadruplexes) thus forming larger complexes.

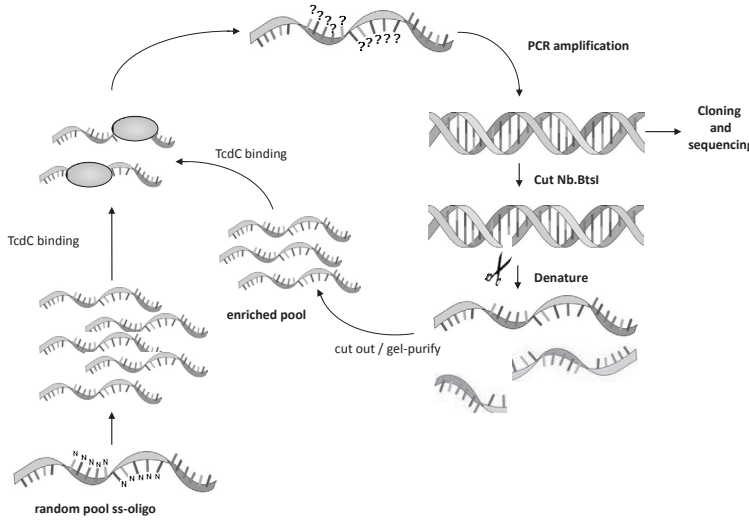


Figure 5: In vitro selection of TcdC binding sites (ssSELEX) starting with a random ssDNA oligonucleotide library, including generation of single stranded DNA after PCR amplification using a sequence specific nicking endonuclease (Nb.BtsI) that cleaves only one strand on a double stranded DNA product.

Characterization of the TcdC bound sequences

To characterize the structure of the HMW product and demonstrate that the HMW complex is the result of intra- or intermolecular structures, we heated radiolabeled clone #5 in the presence of formamide thereby denaturing DNA duplexes and secondary structures. Upon heating, the HMW product shifts to a lower molecular weight (Figure 8A), suggesting that the DNA element is a multiplex forming secondary structures. The fact that this unusual HMW structure is likely to be a multiplex and contains stretches of GGGs suggested that it could form a so called G- quadruplex, a four-stranded helical structure with four guanine bases from each strand forming hydrogen-bonds G-tetrads). Three (or more) guanine tetrads can stack on top of each other to form a G-quadruplex. We therefore tested the HMW product with ETC ($C_{39}H_{47}N_3O_6S_4$), an extended aromatic cyanine dye(that specifically recognizes stacked G-quadruplexes (20).

Figure 8B shows that ETC specifically stains the HMW product but not duplex and single-strand DNA (compare to Figure 8C), confirming that the selected element forms a G-quadruplex. To further support the quadruplex structure (QS) proposed for the #5 sequence we analyzed the HMW formation with a number of point mutations. We did not observe changes in the quadruplex HMWs when

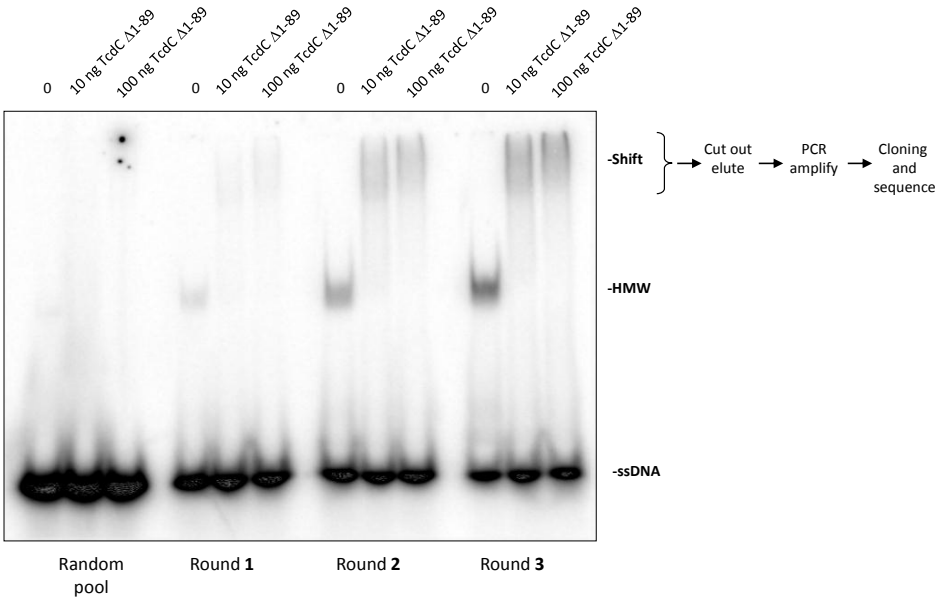


Figure 6. Mobility shift assay of TcdC Δ1-89 selected binding sites. ssDNA selected at each round (see 'Materials and Methods') were used as probes in gel mobility analysis. The selection rounds are indicated. Shifted probes in round 2 and 3 were cut out and eluted as indicated and cloned after the last round.

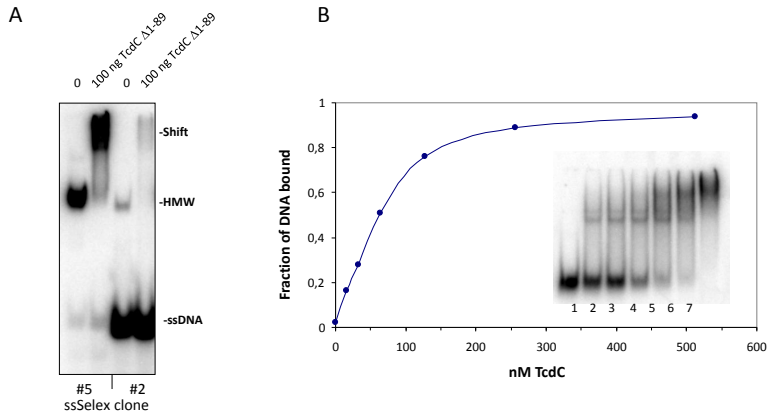


Figure 7.Binding characteristic of TcdC ⊿1-89. **A**. Binding to two selected clones (#2 and #5 corresponding to Table 3) was tested using 100 ng of TcdC⊿1-89 in a bandshift assay. ssDNA = single stranded; HMW= higher molecular weight product formed; shift indicated protein-DNA complex). **B**. Determination of the dissociation constant of TcdC⊿1-89 for #5 HMW binding. Lanes 1 to 7 of the inset gel shows increasing concentrations of TcdC⊿1-89 3 ng, 6 ng, 12 ng, 25ng, 50 ng and 100 ng. Band shift assays were performed as described under material and methods.

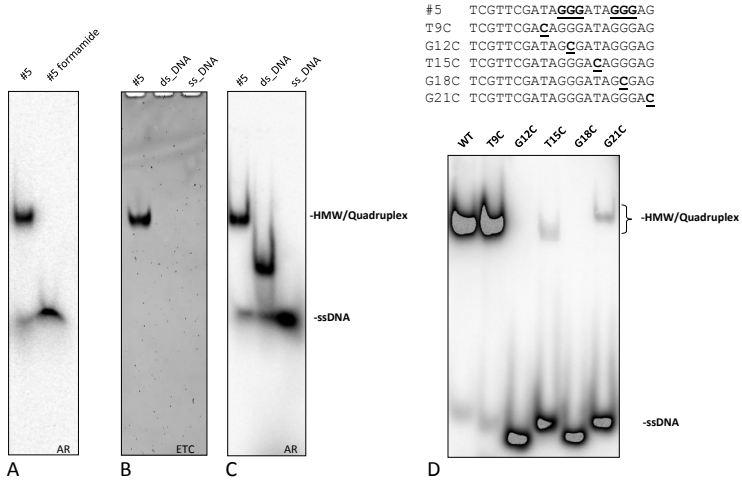


Figure 8: Characterization of a TcdC bound sequence element. **(A)** Higher molecular weight product (HMW) is formed by inter molecular interactions which are lost at 5 min 95°C in formamide. DNA was 5'end labeled with ^{32}Pγ-phosphate for visualization using storage phosphor screen autoradiography (AR). **(B)** Recognition of HMW product by ETC: a quadruplex specific stain (see 'Materials and Methods'). **(C)** Loading control of gel-samples loaded in panel B. Each DNA was 5'end labeled with ^{32}P-γ-phosphate see 'Materials and Methods'). **(D)** Point mutations affecting G-quadruplex formation. G-stretches and mutated positions are in bold underlined. Probes were ^{32}P-labelled and separated on a 10% polyacrylamide gel. Quadruplex is indicated.

the T9 preceding the G-stretches was replaced with a C (Figure 8D). However mutations at positions 12, 15, 18 and 21 of the #5 sequence resulted in significant alteration of migration on the gel, presumably due to a loss of QS. Especially when guanines involved in a G-tetrad formation, G12 and G15, were substituted with a C, the mutations completely abrogated the capacity of the sequence to fold into a G-quadruplex.

Table 3: Selected binding sites for the TcdCΔ1-89 protein in a single stranded SELEX.

#1	TCGGTGTGTT**GGG**TG**AGGG**AC
#2	TCGGCCTGGATACAT**AGGG**AC
#3	TCGGAATGACTGGCGT**GGG**AC
#4	TCGC**GGG**TGGCTGGA**AGGG**AC
#5	TCGTTCGAT**AGGG**AT**AGGG**AC
#6	TCGTTGTCTGGTC**AGGG**GGAC
#7	TCGAGCTATAGGT**GGG**TAGAC
#8	TCGGT**AGGGG**G**AGGG**A**GGG**AC
#9	TCGACAAAGCAT**GGG**TCCGAC
#10	TCGGTCTTTTG**GGG**TAAGGAC
#11	TCGTTTAGG**AGGG**TCTAGAC
#12	TCGAATATG**GGG**AAGTAGGAC
#13	TCGATTTG**GGG**ACTGCTGGAC
#14	TCGCGTC**AGG**AGGTGTTAGAC
#15	TCGCGG**AGGG**AAC**GGG**TGGAC
#16	TCGTAA**AGGG**TGATTCTGGAC
#17	TCGG**AGGG**CCAGGTCGTGAC
#18	TCG**AGGG**TTACCGTAG**GGG**AC
consensus	**aGGG**

Oligonucleotides sequences obtained are aligned. Stretches of 3G or longer and corresponding to the consensus are highlighted in bold. Underlined are the constant sequences flanking the randomized 15 nucleotides.

TcdC binds as a dimer

Above we have shown that the dimerization coiled-coil helix forms a proteolytically protected structure together with the OB-fold. We were interested to determine if dimerization is required for efficient recognition and binding of the G-quadruplex. Therefore electrophoretic mobility shift assay was carried out with TcdCΔ1-130, which behaves as a monomer. Figure 9 shows that no binding occurred with purified TcdCΔ1-130 indicating requirement of the TcdC dimerization domain for efficient binding. To exclude the possibility that the loss of binding by TcdCΔ1-130, is caused by direct of binding of the coiled-coil domain to the quadruplex we tested a TcdC protein, which does contain the dimerization helix (aa90-130) but misses the C-terminal part of the OB-fold containing a loop forming part of the putative

ssDNA-binding channel (aa208-232). Binding of this protein, TcdC Δ1-89_Δ208-232 (Figure 9), to the quadruplex structure was tested and showed no binding confirming that the coiled-coil domain is not directly involved in quadruplex binding.

Discussion

TcdC has been described to act as a factor responsible for inhibition of transcription of the toxin genes. Here we describe that the conserved carboxy terminal domain of TcdC is predicted to form a coiled five-stranded beta-sheet capped by an alpha helix (Figure 1A). This common fold has been described in different proteins which bind oligonucleotides or oligosaccharides and thus named OB-fold (oligonucleotide / oligosaccharide binding) (34). Using ssSELEX, a method to determine the binding site of TcdC, we found that the optimal binding site forms a G-quadruplex.

G-quadruplexes are nucleic acid sequences rich in guanine and capable of constituting a four-stranded structure. These four stranded structures are stabilized through hydrogen bonds between four guanine bases forming a square planar structure called a guanine tetrad (35). Typically three guanine tetrads can stack on top of each other to form a G-quadruplex. Quadruplexes can be formed in DNA as well as RNA and can be very diverse as GGG interactions can form intramolecularly as well as intermolecular. Intermolecular quadruplexes can be arranged from two strands (each containing two GGG stretches) or four strands (one GGG stretch each). The spacing (loops) between the GGG-stretches can vary between 1-7 nucleotides (36). *In silico* studies have shown that putative Quadruplex structures (pQS) are abundant in prokaryotic as well as eukaryotic gene promoters and G-rich telomeres found at the end of chromosomes (37-40). The presence of QS in human gene promoters has been shown to result in transcriptional repression (38). When QS are present in the 5' untranslated region of the mRNA it can interfere with ribosome binding and translation initiation (41;42).When we analyzed the whole *C.difficile* (strain 630) genome in Quadfinder, an online server for prediction of quadruplex-forming motifs in nucleotide sequences (36), we found 5 pQS. Unfortunately none of these were located in the PaLoc, where TcdC is speculated to act. Three pQS are present within open reading frames (CD1092A, CD1115 and CD1849) and two in the 3'UTR of genes (CD0938 and CD2929) in the strand that is complementary to mRNA (producing CCC stretches in the mRNA). It should be noted that prediction programs can only identify intramolecular QS (4 GGG stretches on the same strand) not the bimolecular or tetramolecular forms.

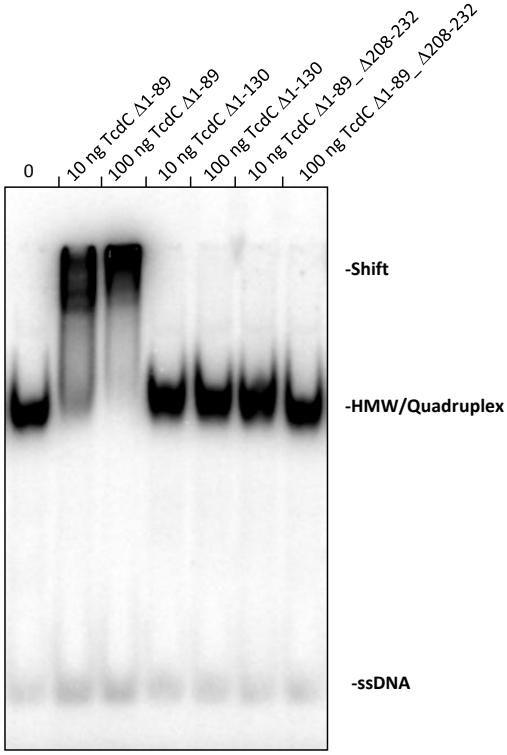


Figure 9: TcdC dimerization domain is required for DNA binding. TcdC dimer (Δ1-89) or monomer (Δ1-130) was incubated with oligo #5. Only the TcdC dimer was able to bind to the quadruplex DNA, as evidenced by the shifted DNA.

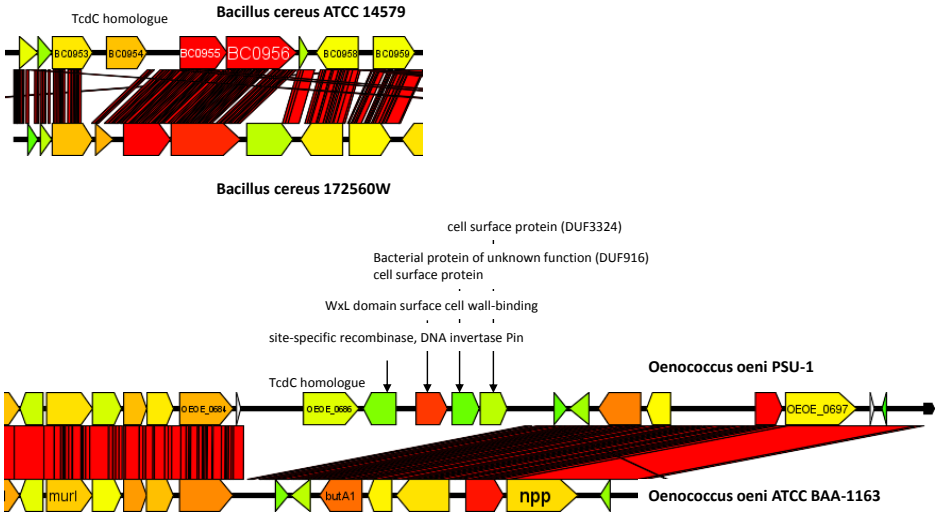


Figure 10: Pairwise comparison between TcdC homologues containing Firmicutes. Two examples are shown.Top; comparison between Bacillus cereus ATCC 14579 and Bacillus cereus 172560W showing a single insertion of the TcdC homologue BC0954. Bottom; comparison between Oenococcus oeni PSU-1 and Oenococcus oeni ATCC BAA-1163 showing an insertion of five genes including the TcdC homologue OEOE_068.

Telomers contain single stranded repeats of TAGGG found at the end of chromosomes protecting them from exonuclease degradation. Intramolecular QS of these TAGGG repeats play a role in telomere maintenance (43). However, in order to efficiently replicate the lagging strand of these telomeres these QS must be disrupted, thereby permitting processive telomere elongation. At least two proteins have been reported to bind and unfold these quadruplex structures, human POT1 and RPA, both characterized by the presence of an OB-fold (44-47). Despite the similarity in protein fold and DNA recognition site between TcdC (aGGG) and these OB-fold proteins (TAGGG), we could not find identical DNA contacting amino-acids when the TcdC structure was superimposed on the hPOT-1-DNA co-crystal (43). Although the circular genomes of Firmicutes do not contain single-stranded ends, a role for the OB-fold containing TcdC homologues in destabilizing alternative DNA secondary structures could be envisaged.

An *alternative* mechanism of TcdC action might be exemplified by the eukaryotic RNA polymerase II complex, which includes a subunit (rpb7) with an OB-fold. It was speculated that this OB-fold domain, which is located at the RNA exit path, binds RNA as it exits the enzyme thereby stabilizing the early transcribing complex (48). An opposite effect i.e. destabilizing the initiation complex by an OB-fold protein, such as TcdC, could be pictured. An unexpected G-quadruplex structure is described in the crystal structure of a bacterial -10 promoter element, 5'TGTACAATGGG 3' (-14 to -4), complexed with sigma factor Taqσ^A (49). In this structure the downstream $G_{-6}G_{-5}G_{-4}$ do not interact with the protein but twist away from the protein-DNA complex and form G-quadruplexes with other (symmetry-related) GGG motifs. The relevance in this complex was not clear and such a GGG-motif is absent next to the *tcdA* -10 promoter element.

TcdC is part of the pathogenicity locus, a well defined genetic element that is present at identical locations in the chromosome of pathogenic *C.difficile* strains. In non-toxinogenic strains, however, it is completely absent. These observations have led to the suggestion that the PaLoc may be associated with a (bacteriophage) transposable genetic element (50). Examining the genomic location of TcdC homologues of other Firmicutes showed that several of these family members are located on insertional-elements. For example the TcdC homologue (E-value 1e-23) of *Oenococussoeni* strain PSU-1 (Figure 10) is part of an insertion containing 4 additional genes encoding three putative cell-wall proteins and one site specific recombinase. In contrast, a TcdC homologue of *Bacillus cereus* strain 172560 (E-value 1e-41) is inserted without any additional genes.

It is interesting to mention that the TcdC variants present in *LactoBacillus* and *Leuconostoc* are also found in their homologous phages (i.e. *LactoBacillus* phages A2 and Lrm1 and *Leuconostoc* phage phiMH1).In these phages the TcdC homologues

are present in the lysis/lysogeny genetic switch operon, located between the CI repressor and Int, integrase, suggesting that TcdC is part of the regulatory decision circuit.

Our overall data suggests that *C.difficile* TcdC forms an OB-fold that binds quadruplex structures. However, the *in vivo* relevance remains unreported. Extensive investigations showed no binding to *tcdA* promoter elements. Clearly, quadruplex structures play a role in gene regulation and expression. Unfortunately, no multiple G-stretches are found within the PaLoc where TcdC is thought to exert its function. It may well be that the single stranded regions of the quadruplex mimics another structure bound by TcdC and the quadruplex is an approximation of the optimal structural binding determinant. It remains to be established in which way the capability of dimeric TcdC to bind G-quadruplexes demonstrated in this study relates to its role as a transcriptional repressor or another cellular function.

Reference List

1. Goorhuis,A., Bakker,D., Corver,J., Debast,S.B., Harmanus,C., Notermans,D.W., Bergwerff,A.A., Dekker,F.W. and Kuijper,E.J. (2008) Emergence of *Clostridium difficile* infection due to a new hypervirulent strain, polymerase chain reaction Ribotype 078. *Clin. Infect. Dis.*, 47, 1162-1170.
2. Kuijper,E.J., Barbut,F., Brazier,J.S., Kleinkauf,N., Eckmanns,T., Lambert,M.L., Drudy,D., Fitzpatrick,F., Wiuff,C., Brown,D.J. *et al.* (2008) Update of *Clostridium difficile* infection due to PCR Ribotype 027 in Europe, 2008. *Euro. Surveill*, 13.
3. McDonald,L.C., Killgore,G.E., Thompson,A., Owens,R.C., Jr., Kazakova,S.V., Sambol,S.P., Johnson,S. and Gerding,D.N. (2005) An epidemic, toxin gene-variant strain of *Clostridium difficile*. *N. Engl. J. Med.*, 353, 2433-2441.
4. Pepin,J., Valiquette,L. and Cossette,B. (2005) Mortality attributable to nosocomial *Clostridium difficile*-associated disease during an epidemic caused by a hypervirulent strain in Quebec. *CMAJ.*, 173, 1037-1042.
5. Rupnik,M., Wilcox,M.H. and Gerding,D.N. (2009) *Clostridium difficile* infection: new developments in epidemiology and pathogenesis. *Nat. Rev. Microbiol.*, 7, 526-536.
6. Govind,R. and Dupuy,B. (2012) Secretion of *Clostridium difficile* toxins A and B requires the holin-like protein TcdE. *PLoS. Pathog.*, 8, e1002727.
7. Mani,N. and Dupuy,B. (2001) Regulation of toxin synthesis in *Clostridium difficile* by an alternative RNA polymerase sigma factor. *Proc. Natl. Acad. Sci. U. S. A*, 98, 5844-5849.
8. Raffestin,S., Dupuy,B., Marvaud,J.C. and Popoff,M.R. (2005) BotR/A and TetR are alternative RNA polymerase sigma factors controlling the expression of the neurotoxin and associated protein genes in *Clostridium* botulinum type A and *Clostridium* tetani. *Mol. Microbiol.*, 55, 235-249.
9. Mani,N., Lyras,D., Barroso,L., Howarth,P., Wilkins,T., Rood,J.I., Sonenshein,A.L. and Dupuy,B. (2002) Environmental response and autoregulation of *Clostridium difficile* TxeR, a sigma factor for toxin gene expression. *J. Bacteriol.*, 184, 5971-5978.
10. Antunes,A., Camiade,E., Monot,M., Courtois,E., Barbut,F., Sernova,N.V., Rodionov,D.A., Martin-Verstraete,I. and Dupuy,B. (2012) Global transcriptional control by glucose and carbon regulator CcpA in *Clostridium difficile*. *Nucleic Acids Res.*.
11. Dineen,S.S., Villapakkam,A.C., Nordman,J.T. and Sonenshein,A.L. (2007) Repression of *Clostridium difficile* toxin gene expression by CodY. *Mol. Microbiol.*, 66, 206-219.
12. Hughes,K.T. and Mathee,K. (1998) The anti-sigma factors. *Annu. Rev. Microbiol.*, 52, 231-286.
13. Govind,R., Vediyappan,G., Rolfe,R.D. and Fralick,J.A. (2006) Evidence that *Clostridium difficile* TcdC is a membrane-associated protein. *J. Bacteriol.*, 188, 3716-3720.
14. Bakker,D., Smits,W.K., Kuijper,E.J. and Corver,J. (2012) TcdC Does Not Significantly Repress Toxin Expression in *Clostridium difficile* 630DeltaErm. *PLoS. One.*, 7, e43247.
15. Merrigan,M., Venugopal,A., Mallozzi,M., Roxas,B., Viswanathan,V.K., Johnson,S., Gerding,D.N. and Vedantam,G. (2010) Human hypervirulent *Clostridium difficile* strains exhibit increased sporulation as well as robust toxin production. *J. Bacteriol.*, 192, 4904-4911.
16. Vohra,P. and Poxton,I.R. (2011) Comparison of toxin and spore production in clinically relevant strains of *Clostridium difficile*. *Microbiology*, 157, 1343-1353.
17. Matamouros,S., England,P. and Dupuy,B. (2007) *Clostridium difficile* toxin expression is inhibited by the novel regulator TcdC. *Mol. Microbiol.*, 64, 1274-1288.

18. Carter,G.P., Douce,G.R., Govind,R., Howarth,P.M., Mackin,K.E., Spencer,J., Buckley,A.M., Antunes,A., Kotsanas,D., Jenkin,G.A. *et al.* (2011) The anti-sigma factor TcdC modulates hypervirulence in an epidemic BI/NAP1/027 clinical isolate of *Clostridium difficile*. *PLoS. Pathog.*, 7, e1002317.
19. Cartman,S.T., Kelly,M.L., Heeg,D., Heap,J.T. and Minton,N.P. (2012) Precise manipulation of the *Clostridium difficile* chromosome reveals a lack of association between the *tcdC* genotype and toxin production. *Appl. Environ. Microbiol.*, 78, 4683-4690.
20. Yang,Q., Xiang,J., Yang,S., Li,Q., Zhou,Q., Guan,A., Zhang,X., Zhang,H., Tang,Y. and Xu,G. (2010) Verification of specific G-quadruplex structure by using a novel cyanine dye supramolecular assembly: II. The binding characterization with specific intramolecular G-quadruplex and the recognizing mechanism. *Nucleic Acids Res.*, 38, 1022-1033.
21. Wolf,E., Kim,P.S. and Berger,B. (1997) MultiCoil: a program for predicting two- and three-stranded coiled coils. *Protein Sci.*, 6, 1179-1189.
22. Roy,A., Kucukural,A. and Zhang,Y. (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.*, 5, 725-738.
23. Zhang,Y. (2007) Template-based modeling and free modeling by I-TASSER in CASP7. *Proteins*, 69 Suppl 8, 108-117.
24. Hildebrand,A., Remmert,M., Biegert,A. and Soding,J. (2009) Fast and accurate automatic structure prediction with HHpred. *Proteins*, 77 Suppl 9, 128-132.
25. Arcus,V. (2002) OB-fold domains: a snapshot of the evolution of sequence, structure and function. *Curr. Opin. Struct. Biol.*, 12, 794-801.
26. Theobald,D.L., Mitton-Fry,R.M. and Wuttke,D.S. (2003) Nucleic acid recognition by OB-fold proteins. *Annu. Rev. Biophys. Biomol. Struct.*, 32, 115-133.
27. Soding,J., Biegert,A. and Lupas,A.N. (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.*, 33, W244-W248.
28. Chen,J., Darst,S.A. and Thirumalai,D. (2010) Promoter melting triggered by bacterial RNA polymerase occurs in three steps. *Proc. Natl. Acad. Sci. U. S. A*, 107, 12523-12528.
29. Wigneshweraraj,S.R., Burrows,P.C., Severinov,K. and Buck,M. (2005) Stable DNA opening within open promoter complexes is mediated by the RNA polymerase beta'-jaw domain. *J. Biol. Chem.*, 280, 36176-36184.
30. McGlynn,P., Mahdi,A.A. and Lloyd,R.G. (2000) Characterisation of the catalytically active form of RecG helicase. *Nucleic Acids Res.*, 28, 2324-2332.
31. Rafferty,J.B., Ingleston,S.M., Hargreaves,D., Artymiuk,P.J., Sharples,G.J., Lloyd,R.G. and Rice,D.W. (1998) Structural similarities between Escherichia coli RuvA protein and other DNA-binding proteins and a mutational analysis of its binding to the holliday junction. *J. Mol. Biol.*, 278, 105-116.
32. Svobodova,M., Pinto,A., Nadal,P. and O' Sullivan,C.K. (2012) Comparison of different methods for generation of single-stranded DNA for SELEX processes. *Anal. Bioanal. Chem.*, 404, 835-842.
33. Xu,D., Guo,R., Sobeck,A., Bachrati,C.Z., Yang,J., Enomoto,T., Brown,G.W., Hoatlin,M.E., Hickson,I.D. and Wang,W. (2008) RMI, a new OB-fold complex essential for Bloom syndrome protein to maintain genome stability. *Genes Dev.*, 22, 2843-2855.
34. Murzin,A.G. (1993) OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J.*, 12, 861-867.

35. Bochman,M.L., Paeschke,K. and Zakian,V.A. (2012) DNA secondary structures: stability and function of G-quadruplex structures. *Nat. Rev. Genet.*, 13, 770-780.
36. Scaria,V., Hariharan,M., Arora,A. and Maiti,S. (2006) Quadfinder: server for identification and analysis of quadruplex-forming motifs in nucleotide sequences. *Nucleic Acids Res.*, 34, W683-W685.
37. Lipps,H.J. and Rhodes,D. (2009) G-quadruplex structures: in vivo evidence and function. *Trends Cell Biol.*, 19, 414-422.
38. Qin,Y. and Hurley,L.H. (2008) Structures, folding patterns, and functions of intramolecular DNA G-quadruplexes found in eukaryotic promoter regions. *Biochimie*, 90, 1149-1171.
39. Rawal,P., Kummarasetti,V.B., Ravindran,J., Kumar,N., Halder,K., Sharma,R., Mukerji,M., Das,S.K. and Chowdhury,S. (2006) Genome-wide prediction of G4 DNA as regulatory motifs: role in Escherichia coli global regulation. *Genome Res.*, 16, 644-655.
40. Yadav,V.K., Abraham,J.K., Mani,P., Kulshrestha,R. and Chowdhury,S. (2008) QuadBase: genome-wide database of G4 DNA--occurrence and conservation in human, chimpanzee, mouse and rat promoters and 146 microbes. *Nucleic Acids Res.*, 36, D381-D385.
41. Bugaut,A. and Balasubramanian,S. (2012) 5'-UTR RNA G-quadruplexes: translation regulation and targeting. *Nucleic Acids Res.*, 40, 4727-4741.
42. Wieland,M. and Hartig,J.S. (2007) RNA quadruplex-based modulation of gene expression. *Chem. Biol.*, 14, 757-763.
43. Lei,M., Podell,E.R., Baumann,P. and Cech,T.R. (2003) DNA self-recognition in the structure of Pot1 bound to telomeric single-stranded DNA. *Nature*, 426, 198-203.
44. Paeschke,K., Simonsson,T., Postberg,J., Rhodes,D. and Lipps,H.J. (2005) Telomere end-binding proteins control the formation of G-quadruplex DNA structures in vivo. *Nat. Struct. Mol. Biol.*, 12, 847-854.
45. Salas,T.R., Petruseva,I., Lavrik,O., Bourdoncle,A., Mergny,J.L., Favre,A. and Saintome,C. (2006) Human replication protein A unfolds telomeric G-quadruplexes. *Nucleic Acids Res.*, 34, 4857-4865.
46. Wu,L., Multani,A.S., He,H., Cosme-Blanco,W., Deng,Y., Deng,J.M., Bachilo,O., Pathak,S., Tahara,H., Bailey,S.M. *et al.* (2006) Pot1 deficiency initiates DNA damage checkpoint activation and aberrant homologous recombination at telomeres. *Cell*, 126, 49-62.
47. Zaug,A.J., Podell,E.R. and Cech,T.R. (2005) Human POT1 disrupts telomeric G-quadruplexes allowing telomerase extension in vitro. *Proc. Natl. Acad. Sci. U. S. A*, 102, 10864-10869.
48. Spahr,H., Calero,G., Bushnell,D.A. and Kornberg,R.D. (2009) Schizosacharomyces pombe RNA polymerase II at 3.6-A resolution. *Proc. Natl. Acad. Sci. U. S. A*, 106, 9185-9190.
49. Feklistov,A. and Darst,S.A. (2011) Structural basis for promoter-10 element recognition by the bacterial RNA polymerase sigma subunit. *Cell*, 147, 1257-1269.
50. Braun,V., Hundsberger,T., Leukel,P., Sauerborn,M. and von Eichel-Streiber,C. (1996) Definition of the single integration site of the pathogenicity locus in *Clostridium difficile*. *Gene*, 181, 29-38.

8

The HtrA-like protease CD3284 modulates virulence of *Clostridium difficile*

Dennis Bakker[1], Anthony M. Buckley[2], Anne de Jong[3], Vincent J.C. van Winden[1], Joost P.A. Verhoeks[1], Oscar P. Kuipers[3], Gillian R. Douce[2], Ed J. Kuijper[1], Wiep Klaas Smits[1‡], Jeroen Corver[1‡]

[1] *Department of Medical Microbiology, Section Experimental Bacteriology, Leiden University Medical Center.*
[2] *Institute of Infection, Immunity and Inflammation, University of Glasgow*
[3] *Department of Molecular Genetics, University of Groningen.*
[‡] *Both authors contributed equally*

Infection and Immunity 2014

Abstract

In the past decade, *Clostridium difficile* has emerged as an important gut pathogen. Symptoms of *C. difficile* infection range from mild diarrhea to pseudomembranous colitis. Besides the two main virulence factors Toxin A and B, other virulence factors are likely to play a role in the pathogenesis of the disease. In other Gram positive and Gram negative pathogenic bacteria conserved High temperature requirement A (HtrA)-like proteases have been shown to have a role in protein homeostasis and quality control. This affects the functionality of virulence factors and the resistance of bacteria to (host-induced) environmental stresses.

We found that the *C. difficile* 630 genome encodes a single HtrA-like protease (CD3284; *htrA*) and have analyzed its role *in vivo* and *in vitro* through the creation of an isogenic ClosTron-based mutant of *htrA* in *C. difficile* strain 630Δ*erm* (wild-type). In contrast to the attenuated phenotype seen with *htrA* deletion in other pathogens, this mutant showed enhanced virulence in the Golden Syrian hamster model of acute *C. difficile* infection. Micro-array data analysis showed a pleiotropic effect of *htrA* on the transcriptome of *C. difficile*, including up-regulation of the toxin A gene. In addition, *the htrA* mutant showed reduced spore formation and adherence to colonic cells. Together, our data show that *htrA* can modulate virulence in *C. difficile*.

Introduction

The bacterial pathogen *Clostridium difficile* is the leading cause of infectious nosocomial diarrhoea (1-3). *Clostridium difficile* infection (CDI) can cause disease with a wide variety of symptoms, ranging from mild diarrhea to severe pseudomembranous colitis (1,4,5). Since 2003, the global prevalence of reported CDI cases has escalated (5-8). In addition, the severity of the disease and the mortality has increased (1,4,9).

The main virulence factors of *C. difficile* are the two large clostridial toxins, Toxin A (TcdA) and Toxin B (TcdB) (10-12). These toxins are glycosyltransferases that inactivate Rho, Rac and Cdc42, thereby disrupting the cytoskeleton and tight junctions of the colon epithelial cells, resulting in activation of the inflammasome and cellular apoptosis (10,12). A third toxin, binary toxin, is produced by certain strains (e.g. PCR ribotype (RT) 027 and 078) that have been associated with increased levels of mortality and morbidity (1,9). It has been suggested that binary toxin may contribute to disease in hamsters (13). *In vitro* assays have demonstrated that the binary toxin affects adhesion of *C. difficile* to cells through protrusion formation of the target cells (14).

Besides the three toxins, little is known about other virulence factors and their role in colonization and establishment of an infection in the host. Presently, several surface proteins have been identified or hypothesized to play a role in colonic adhesion. These include the fibronectin-binding protein A (15), S-layer proteins (16), Cwp84 (17), flagellar proteins (18) and CD1581 (19).

Alongside colonization factors, adaptation to stresses in the host (including the antibacterial response, elevated temperatures, extreme pH and osmotic stress) are likely to play a vital role in the establishment of an infection. These stresses can result in the accumulation of (partially) unfolded proteins that are non-functional or form poisonous aggregates (20). The well-conserved family of bacterial High temperature requirement A (HtrA) proteases plays an important role in the protein quality control and homoeostasis by combining proteolytic and chaperone activities in a variety of microorganisms (21-23).

HtrA-like proteases are generally composed of three structurally distinct domains: a trypsin-like serine protease domain, one or two PDZ domains and a trans-membrane domain or a signal peptide (22-24). The PDZ domain(s) are highly flexible domains and are involved in oligomerization, substrate recognition and/or the regulation of protease activity (21,24). Membrane anchored HtrA-like proteases are active as trimers and soluble HtrA-like proteases form larger active oligomers (21,22).

In this study we describe the identification and characterization of a *C. difficile* HtrA-like protease (CD3284; hereafter HtrA). We show that an *htrA* null mutant

has enhanced virulence in the Golden Syrian hamster model of acute *C. difficile* infection. This contrasts data reported for other pathogens, in which mutation of *htrA* results in attenuation (25-27). Furthermore, the *htrA* null mutant displayed a pleiotropic effect in a transcriptomic analysis. Several differentially expressed genes were validated and are discussed in the context of *C. difficile* virulence.

Material & Methods

Bacterial strains and growth conditions

The *C. difficile* and *E. coli* strains and plasmids used in this study are listed in Table 1. *E. coli* strains were grown in Luria Bertani medium (LB, USB cooperation) supplemented with appropriate antibiotics when required. *C. difficile* strains were grown anaerobically in a microaerobic cabinet (Don Whitley DG 250) at 37 ºC in pre-reduced TTY medium (30 g/L Bacto Trypton (BD), 20 g/L Yeast extract (Difco) and 0.1% thioglycolate pH 7.4) or Brain Heart Infusion broth (Oxoid) supplemented with 5 g/L yeast extract and 0.01% L-cysteine (Sigma) (BHIS) (28). Logarithmic growth phase bacteria (OD_{600} 0.4 ± 0.08) from overnight pre-cultures were used to inoculate pre-reduced TTY broth at a starting OD_{600} of 0.05 (±0.01). Optical density readings were taken hourly until the stationary growth phase (8 hour post inoculation) and at 24 and 48 hour post inoculation. Cultures of *C. difficile* strains harboring an expression plasmid were induced with anhydrotetracycline (100 ng/mL) at one hour post inoculation. We routinely monitored the purity of cultures by performing control PCRs to confirm identity of mutant strains and stability of conjugated expression plasmids.

Protein purification of His10-Δ(1-30)-HtrA and His10-Δ(1-30)-HtrA-S217A

To facilitate purification from *Escherichia coli*, CD3284 was cloned without the N-terminal 30 amino acids into pET-16B, resulting in a fusion protein with an N-terminal 10xhis-tag but lacking the predicted transmembrane helix. To generate this clone, the CD3284 open reading frame was amplified using primers CD3284F2 and CD3284R2 using Pfu polymerase (Fermentas) and chromosomal DNA from strain 630Δ*erm* (29) as a template. The resulting product was digested with NdeI and XhoI and ligated into similarly digested pET-16B (Novagen), yielding pVW001. After sequence verification using conventional Sanger sequencing, the plasmid was transformed into *E. coli* C43(DE3), in which expression of toxic proteins/proteases is feasible (30). To induce expression of His10-Δ(1-30)-HtrA, an overnight culture grown in LB with 100 μg/mL ampicillin was diluted 1/100 in 500 mL fresh medium to an optical density at 600 nm of 0.6 after which 1mM of IPTG was added for

3 hour. Hereafter, cells were harvested by centrifugation (10 minutes 10 rpm 4 °C) and the protein was purified essentially as described for Spo0A (31), with the following modifications: an 0.5 mL TALON Superflow column was used, and the stepwise washing/elution was carried out with 8, 4, 4, 4, and 4 column volumes (20mM, 50mM, 100mM, 250mM, 500mM, respectively). After the addition of glycerol to a final concentration of 8%, protein was quantified by measuring the absorption at 280nm on a Nanodrop ND-200 machine (ThermoScientific), using an extinction coefficient (ε) of 14440 M-1 cm-1 and a molecular weight of 36.77 kDa, based on analysis of the protein sequence using the ProtParam tool (http://web.expasy.org/protparam/). Protein not used immediately was stored at -80°C. To generate a catalytic mutant of HtrA, we mutated the conserved Ser217 to an alanine on the basis of published mutants of the *E. coli* DegS protease. To do so, oligo-directed mutagenesis was performed using primers CD3284S217AF and CD3284S217AR on plasmid pVW001, yielding plasmid pJV001. The mutant protein (His10-Δ(1-30)-HtrA-S217A) was purified and quantified as described for His10-Δ(1-30)-HtrA.

Table 1: Strains and plasmids used in this study.

Strains	Description/Genotype	Origin
Escherichia coli		
DH5α		Laboratory stock
CA434		(34)
C43		(30)
Clostridium difficile		
WKS1241	630Δ*erm*	(29)
DB0051	630Δ*erm*, pRPF185	This study
DB0002	630Δ*erm*, c*d3284::*(*ClosTron*, *ermB*)	This study
DB0052	630Δ*erm*, c*d3284::*(*ClosTron*, *ermB*) pRPF185 (*catP*)	This study
DB0047	630Δ*erm*, *cd3284::*(*ClosTron, ermB*) ,pDB0031 (*catP*)	This study
WKS1237	630Δ*erm*, *spo0A::*(*ClosTron*, *ermB*)	(32)
WKS1710	630Δ*erm*, *tcdA::*(*ClosTron*, *ermB*) *tcdB::*(*ClosTron*, *catP*)	(40)
Plasmids		
pMTL007	*catP*, ClosTron GroupII Intron	(32)
pDB0001	pMTL007::*cd3284*-162s, *catP*	This study
pRPF185	*tetR*, *Ptet-gusA*, *catP*,	(35)
pDB0031	*tetR, Ptet-cd3284, catP*	This study
pVW001	*PT7-His10-*Δ(*1-30*)*-htrA, bla*	This study
pJV001	*PT7-His10-*Δ(*1-30*)*-htrA-S217A, bla*	This study

Table 2: Primers and probes used in this study.

Oligos	Sequence (5'- 3')	Description	origin
IBS-cd3284	AAAAAAGCTTATAATTATCCTTAGGTAACAGTCAAGTGCGCCCAGATAGGGTG	ClosTron	This study
EBS2-cd3284	TGAACGCAAGTTTCTAATTTCGATTTTACCTCGATAGAGGAAAGTGTCT	ClosTron	This study
EBS1d-cd3284	CAGATTGTACAAATGTGGTGATAACAGATAAGTCAGTCAAAATAACTTACCTTTCTTTGTGT	ClosTron	This study
EBS universal	CGAAATTAGAAACTTGCGTTCAG TAAA	ClosTron/Control PCR	
oDB0067	CTGAGCTCCTGCAGTAAAGGAGAAAATTTTATGTCAAGAAGAAAGAAAGG	Complementation	This study
oDB0068	TAGGATCCGGTTAGAAATTCACATTTATTGTT	Complementation	This study
oDB0121	CCAGCTTTTTCTGCTGATGA	Control PCR/seq.	This study
oDB0123	GAAGTTCTACGCCGATAGTT	Control PCR/seq.	This study
oDB0124	AATACCTACAACAGATGGCG	Control PCR	This study
NF_794	CACCGACGAGCAAGGCAAGACCG	Control PCR/ seq.	(23)
NF_1323	CTGGACTTCATGAAAAACTAAAAAAAATATTG	sequencing	(23)
oWKS-1177	ATCTAGCTAGCGCCAGGAGAGTTGTTGATTC	Control PCR	This study
rspJ Forward	GATCACAAGTTTCAGGACCTG	qPCR	(4)
rspJ Reverse	GTCTTAGGTGTTGGATTAGC	qPCR	(4)
tcdA Forward	AATCCAATACAAGCCCTGTAG	qPCR	(4)
tcdA Reverse	TATCAGCCCATTGTTTTATGTATTC	qPCR	(4)
tcdA probe	FAM-TCACTGACTTCTCCACCTATCCATACAA-BHQ-1	qPCR	(4)
oDB0117	ACTCAAAGCGCAATAAATCTAGGAGC	qPCR	This study
oDB0118	ACTGGTCTAGGTTTTTGGCTCAACTTGT	qPCR	This study

Protease assay

Protease activity of the purified proteins was assayed as follows. In a 100 **µ**L reaction, 2.5 **µ**M of His10-Δ(1-30)-HtrA or His10-Δ(1-30)-HtrA-S217A was incubated with 20 **µ**M of casein or beta-casein (molecular weight ~24kDa) in 1xHNE buffer (150mM NaCl, 0.1mM, 5mM HEPES, pH7.4, 0.1 mM EDTA), with the pH adjusted to pH5.5 by the addition of 0.1x MES/A buffer (0.01M MES, 0.01M acetic acid, pH5.0). Reactions were incubated on ice (0°C) or at 37°C for 18h with a no-protein control included. After incubation, the reaction was terminated by the addition of 25 **µ**L 5xSDS sample buffer (10% SDS, 10% beta-mercaptoethanol, 50% glycerol, 0.1% bromophenolblue, 250mM Tris-HCl pH6.8, 50mM EDTA) and heating to 96°C for 2 minutes. 20 **µ**L volumes of the samples were loaded on a 12% SDS-PAGE gel and separated by electrophoresis at 80mA. Gels were stained with Coomassie solution (0.1 % Coomassie Brilliant Blue R-250, 20 % (v/v) methanol, 10% (v/v) acetic acid) and destained with 50% methanol -10% (v/v) acetic acid.

Generation of an *htrA* (*cd3284*) mutant strain

An *htrA* mutant was generated by insertional inactivation of the *cd3284* gene in 630Δ*erm* (which will be referred to as wild type in the rest of this paper) using ClosTron technology (32,33). Briefly, the Perutka algorithm on the ClosTron website (http://www.clostron.com) was used to design primers (supplemental Table 2) for retargeting the Group II intron (Sigma; Targetron). The retargeted intron was cloned using the restriction enzymes BsrGI and HindIII into plasmid pMTL007 and the constructs were verified by sequencing (33). The sequenced verified plasmid (pDB001) was transformed to *E. coli* CA434 and transferred via conjugation to *C. difficile* (29,34). The selection of *C. difficile* transconjugants was achieved by sub-culturing on pre-reduced BHIS agar supplemented with thiamphenicol (Sigma; 10 **µ**g/mL) and *C. difficile* selective supplement (Oxoid). After this, several rounds of sub-culturing on pre-reduced BHIS agar supplemented with lincomycin (Sigma; 20 **µ**g/mL) and *C. difficile* selective supplement were used to promote integration of the GroupII intron into the gene of interest. Chromosomal DNA isolated from the transconjugants (QIAamp blood kit, Qiagen) was analyzed by conventional PCR and sequencing of the PCR product to confirm the insertional inactivation of the *cd3284* gene. Primers used for cloning and sequencing are listed (Table 2). In addition, Southern blot analysis with an *ermB* probe was performed to verify a specific single integration into the genome as described previously (28) (data not shown).

Generation of strains carrying a plasmid based inducible *htrA*

Chromosomal DNA from the wild type strain (WKS1241) was used to PCR amplify the open reading frame of *cd3284* with primers oDB0067 and oDB0068. The resulting amplicon was digested with *BamHI* and *SacI* and cloned into pRPF185 thereby replacing the *gusA* gene (35). The sequence of the resulting inducible plasmid (pDB0031) was verified with primers NF_794, NF_1323 (35), oDB0067 and oDB0068. Plasmid pDB0031 was transferred into *E. coli* CA434 and conjugated into the *htrA* null mutant, resulting in a *htrA* mutant that can be complemented by inducible expression of a plasmid located copy of *htrA*. To address potential effects of the conjugated plasmid (pDB0031) in the phenotypic assays we generated control strains by conjugating the pRPF185 plasmid into wild type (DB0051) and *htrA* mutant (DB0052) strains (35). These strains served as controls in the *in-vitro* assays. There was no significant difference between strains with and without pRPF185 (data not shown).

Animal experiments

All procedures were conducted in strict accordance to the Animals (Scientific Procedures) Act 1986 approved by the Home Office, U.K. (project license number PPL60/4218). Female Golden Syrian hamsters that weighed approximately 100g (bred in-house) were housed individually and given water and food *ad libitum*. Telemetry chips (VitalView Emitter) were inserted by laparotomy into the body cavities of the animals at least 3 weeks before infection with *C. difficile*. Once the wounds healed, the animals were placed on receiver pads, and the body temperature and activity were monitored (Vital View software). Each animal received orogastrically 30 mg/kg of clindamycin phosphate in a single dose five days before infection. Six animals per group were inoculated by oral gavage with 10^4 spores of *C. difficile*. Animals were carefully monitored and culled when core body temperature dropped below 35 °C (28).

Statistical analyses were performed using the GraphPad Prism 5.03 (GraphPad Prism Software). A Mantel-Cox Log-rank statistical test was used to determine significant differences in hamster survival times between *C. difficile* strains. P values ≤0.05 were considered significant.

Minimum Inhibitory concentration (MIC) determination

The MIC of *C. difficile* wild type and *htrA* null mutant to erythromycin and clindamycin was determined by the doubling dilution method described previously (26). Briefly, rows of pre-conditioned BHI broth (90 µl) were supplemented with a concentration range of 1024-0.06 µg/mL of either antibiotic. Wells were inoculated with ~$5x10^3$ spores/100 µl and incubated at 37 °C for 48 hour anaerobically.

Wells to allow positive and negative evaluation of growth were included in which either no antibiotic or no spores were added. After incubation, plates were visually inspected and compared to the controls; the MIC value was determined as the lowest concentration of antibiotic at which growth was no longer visible. Results are given as the median MIC from three biological repeats.

RNA extraction

Five mL of the *C. difficile* cultures were diluted with ice cold methanol (1:1) and stored overnight at -80°C. Bacterial pellets, obtained by centrifugation (20 min, 3000 x g, 4°C), were resuspended in 200 µL lysis buffer (100 mM EDTA, 200 mM Tris-HCl pH 7.0, 50 mg/mL lysozyme) and incubated for 1 hour at 37°C. RNA was isolated using Tri-pure reagent (Roche) as previously described (28). The RNA was treated twice with a TurboDNase (Ambion) according to the manufacturer's instruction, followed by another Tri-pure RNA isolation. The Agilent Bioanalyzer (Agilent Technologies Netherlands BV) and a ND-1000 spectrophotometer (NanoDrop Technologies) were used to analyze the quality and purity of the extracted RNA.

DNA Microarray Analysis

Random nonamers were used to convert total RNA into amino allyl-modified cDNA using a Superscript III reverse transcriptase kit (Life Technologies) and purified with Nucleospin Gel and PCR clean-up kit (Machery-Nagel). The synthesized cDNA was labeled with DyLight 550 and DyLight 650 Amine/Reactive Dyess (Pierce Biotech USA) and purified again with the Nucleospin Gel and PCR clean up kit. The Agilent DNA micro-arrays (G2509F Custom Microarray GE 8x15K 60mer) were designed based on the available genome sequence of *C. difficile* 630 (36,37). The custom made micro-array slides consist of 8 sub-arrays each with 15.208 60-mer probes (37). The Agilent microarray slides were competitively hybridized with 300 ng labeled cDNA made from RNA isolated from both wild type and *htrA* mutant (DB0002) during logarithmic phase (OD_{600} 0.5 ±0.1) and stationary phase growth (12 hour post inoculation) . The microarray slides were scanned with a GenePix scanner and the fluorescent intensities were quantified using GenePix 6.1. The raw data files were analyzed with a LimmaR package via the Genome2D pipeline (http://genome 2d.molgenrug.nl) (38) using the following settings; i) the microarray data was normalized using LOESS normalization, followed by quantile normalization between the slides ii) a weight factor was determined based on the quality of each DNA microarray slide. For each time-point (logarithmic or stationary phase) 3 biological replicates were used. The complete data set, experimental raw data and analyzed data, was submitted to the GEO database (accession number GSE55926).

Reverse Transcriptase Quantitative PCR

For Reverse Transcriptase quantitative PCR (RT-qPCR) the RevertAid™ H Minus Reverse Transcriptase kit (Fermentas) was used to synthesize cDNA according to the manufacturer's instruction. Random hexamers were used to convert 750 ng RNA into cDNA. The synthesized cDNA was treated with RNase (Qiagen) for 1 hour at 37°C and stored at -20°C. The software program Molecular Beacon (Premier Biosoft) was used to design primer pairs (Table 2) for the quantitative PCRs (qPCR), based on the available genome sequence of *C. difficile* strain 630 (36).

All primer pairs were first tested by conventional PCR to confirm specificity and amplicon sizes. The RT-qPCR reactions were then performed using a CFX96 real-time PCR detection system (Biorad). The amplification efficiencies of the genes were determined using serially diluted genomic DNA. Expression levels of the *tcdA* and *spo0A* genes were normalized using the amplification efficiencies and the expression level of the reference gene *rpsJ* (39). The qPCR for *tcdA* and *rpsJ* was performed as previously described (28). The qPCR for the *spo0A* gene contained 25 µL Hotstar mastermix (Qiagen), forward (oDB0117) and reverse primers (oDB0118) (80 nm each primer), 2.5 mM $MgCl_2$, 0.06 % SYBRgreen (Sigma) and 2 µL synthesized cDNA. The qPCR protocol included an enzyme activation step for 15 min at 95 °C, followed by 50 cycles of amplification; 95 °C for 30 sec, 52 °C for 30 sec and 72 °C for 30 sec.

To determine the specificity of the fluorescence, the fluorescent data was converted into melt curve peaks. For each biologic replicate (n=3) four technical replicates were tested. An independent student t-test was employed to compare the strains at different time points.

Western blot analysis of TcdA and Spo0A

Quantitative western blots for TcdA and Spo0A were performed as follows. For TcdA, filtered (0.45 µm) bacterial supernatants from 48 hour cultures were analyzed by SDS-PAGE (6% polyacrylamide) and transferred onto nitrocellulose membranes. For Spo0A protein, bacterial pellets from cultures in the logarithmic phase (OD_{600} nm 0.5 ±0.1), and at 8, 12 and 24 hour post inoculation were collected by centrifugation (10 min, 11,000 g, 4 °C). Pellets were resuspended in PBS containing protease inhibitor cocktail (Sigma) and lyzed with a TissueLyser LT bead beater (50 Hz, 10 min, Qiagen) using acid wash glass beads (≤ 106 µm, Sigma). Equal amounts (OD_{600} nm corrected) of the lyzed bacterial pellets were separated on 15 % SDS-PAGE gels and transferred onto nitrocellulose membrane.

For TcdA quantification the membranes were probed with a mouse monoclonal anti-TcdA antibody (TCC8, tgcBiomics, Bingen, Germany) and for Spo0A quantification the membranes were probed with a mouse polyclonal Spo0A serum

(31). The probed membranes were analyzed by using a secondary anti-mouse biotin-labeled antibody (Dako) and a tertiary anti-biotin antibody labeled with Cy3. A Typhoon 9410 scanner (GE Healthcare) was used to measure the fluorescence intensity. *C. difficile* strains WKS1237 (*spo0A* null mutant) and WKS1710 (*tcdA* and *tcdB* null mutant) were used as control strains (32,40). The software package ImageQuant TL (Amersham Biosciences) was used to quantify fluorescence intensity.

Relative quantification of toxin expression

Quantification of the total toxin production was performed on filtered (0.45 μM) 24 and 48 hour post-inoculation supernatants, using a Vero cell based cytotoxicity assay (28). The end-point titer was defined as the first dilution at which Vero cell morphology was indistinguishable from the neutralized 200 fold diluted supernatant and untreated cells (28). *C. difficile* strain WKS1710 (*tcdA* and *tcdB* null mutant) was used as a negative control (40). Each experiment was performed in duplicate on three separate occasions. An independent student t-test was employed to compare the strains at different time points.

Relative quantification of sporulation efficiency

Sporulation efficiency was determined by analyzing culture samples of 24 and 48 hour post inoculation. Serially diluted non-heat shocked and heat shocked (20 min , 65 °C) culture samples were plated onto BHI plates supplemented with 0.1 % taurocholate to stimulate germination and enhance spore recovery (41). Plates were incubated for 48 hour before colony forming units (CFU) were enumerated. The *C. difficile* strain WKS1237 was included as a negative control for sporulation (32). Sporulation efficiency was normalized using the CFU count of the non-treated sample. Three biological replicates were analyzed in duplicate. An independent student t-test was employed to compare the strains at different time points.

Relative quantification of adhesion to Caco-2 Cells

Adhesion of *C. difficile* to Caco-2 cells was performed essentially as described previously with minor modifications (42). Briefly, Caco-2 cells were grown in RPMI (GE healthcare), 10 % fetal calf serum, Penicillin (100 μg/mL) and Streptomycin (100 U/mL) in a humidified 5 % CO_2 atmosphere at 37 °C. Caco-2 cells were seeded into a 24 wells plate at a density of 1x 10^5 cells per well and incubated overnight at 37 °C and 5 % CO_2. Confluent Caco-2 monolayers were washed once with PBS before transfer into the anaerobic cabinet.

Mid-logarithmic phase bacterial cultures (OD_{600} 0.5 ± 0.08) were collected by centrifugation (1 min, 4000 x g, 4 °C) and washed once with pre-reduced PBS. To

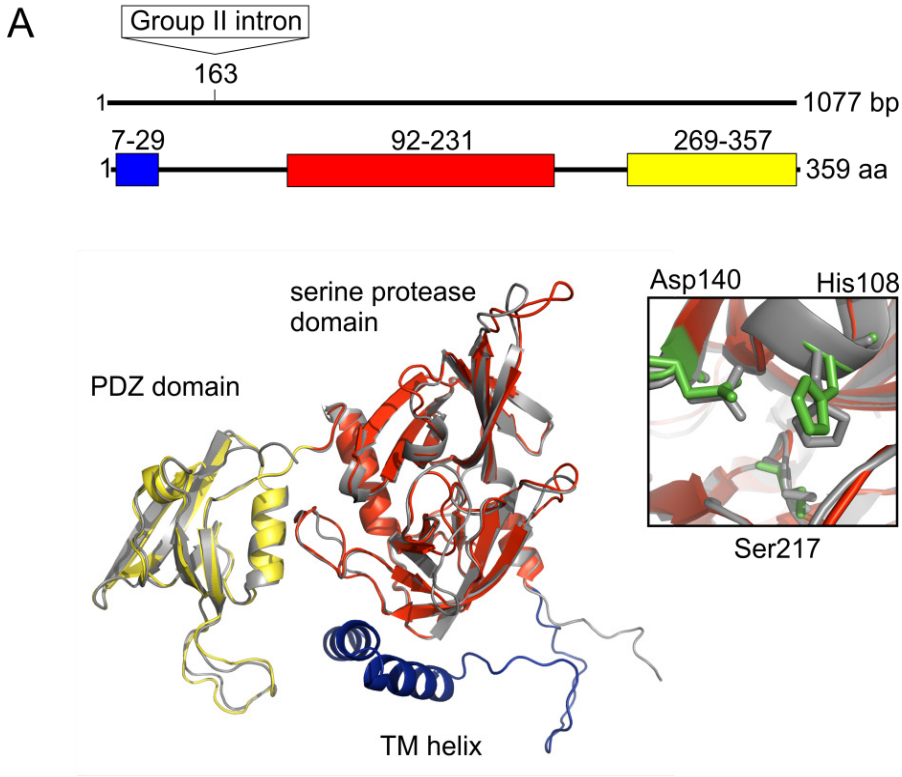


Figure 1: HtrA is a protease. **(A)** Schematic representation of the predicted structural domains of HtrA. The box on the top line indicates the position (nucleotide 163) at which the group II intron has inserted to create the *htrA* null mutant (not on scale). The predicted N-terminal transmembrane helix is shown in blue, the serine protease domain in red and the PDZ domain in yellow. The inset shows the conserved catalytic triad (in green) of the serine protease domain. The predicted structure of *C. difficile* HtrA is overlaid with the structure of *E.coli* DegS (1TE0A) in grey. **(B)** HtrA shows protease activity towards β-casein. 2.5µM of the indicated protein was incubated for 18 hour with 20 µM of beta-casein at 0°C or 37°C. wt = His10-Δ(1-30)-HtrA, S217A is the catalytic mutant His10-Δ(1-30)-HtrA- S217A. The asterisk indicates a proteolytic product of β-casein.

each well of Caco-2 approximately 1x 10^8 bacteria were added. Exact numbers of viable bacteria within the inoculums for each well were calculated by serial dilution and plating of cultures on BHIS plates. After 3 hour of incubation under anaerobic conditions, the Caco-2 cell monolayers with the adhered bacteria were washed 5 times with pre-reduced PBS. Adherent bacteria were released using 100 µL 1x trypsin solution (PAA laboratories, Coelbe, Germany) and serially diluted and plated on pre-reduced BHIS plates. After 48 hour colony forming units were enumerated to determine the percentage of adherent *C. difficile* relative to the initial inoculum. Each experiment was carried out in duplicate and repeated three times. An independent sample t-test was employed to compare the strains at different time points.

Results

Bioinformatics analysis of CD3284

In this study we identify a single HtrA-like protease in *C. difficile* 630Δ*erm*. CD3284 was identified using the BlastP (http://blast.ncbi.nlm.nih.gov/Blast.cgi), Smart (http://smart.embl-heidelberg.de/) and Merops (http://merops.sanger.ac.uk/) databases (36,43-45). Further bioinformatics analyses were performed with a conserved domain prediction webserver (http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml), the trans-membrane prediction webserver TMHMM2.0 (http://www.cbs.dtu.dk/services/TMHMM/) and the signal peptide cleavage site webserver SignalP4.1 (http://www.cbs.dtu.dk/services/SignalP/). These analyses revealed the presence of a trans-membrane domain (Amino Acid (AA) 7-29), a trypsin-like serine protease domain (AA 92-231, E-value $3.96x10^{-29}$) and a PDZ domain (AA 269-357, E-value $2.17x10^{-19}$) (Figure 1A). The N-terminus of CD3284 was predicted by TMHMM to be highly hydrophobic, most likely transmembrane, whereas SignalP analysis revealed no signal sequence, suggesting that the protease and PDZ domains are located extracellularly, but membrane associated (46). Comparison of the CD3284 protein sequence of our wild type strain with the NCBI non-redundant protein sequences database using BlastP, resulted in identical protein sequence hits (Query coverage 100%, E-value 0.0) in other epidemic (e.g. RT027) and non-epidemic (e.g. RT001) *C. difficile* strains.

As protein 3D-structure is more conserved in evolution than its equivalent amino acid sequence, we validated the homology of CD3284 using the I-TASSER (http://zhanglab.ccmb.med.umich.edu/I-TASSER/) and Phyre2 three-dimensional structure prediction server (http://www.sbg.bio.ic.ac.uk/phyre2/) (47,48). These analyses revealed with 100% confidence, over 81% alignment coverage and 40%

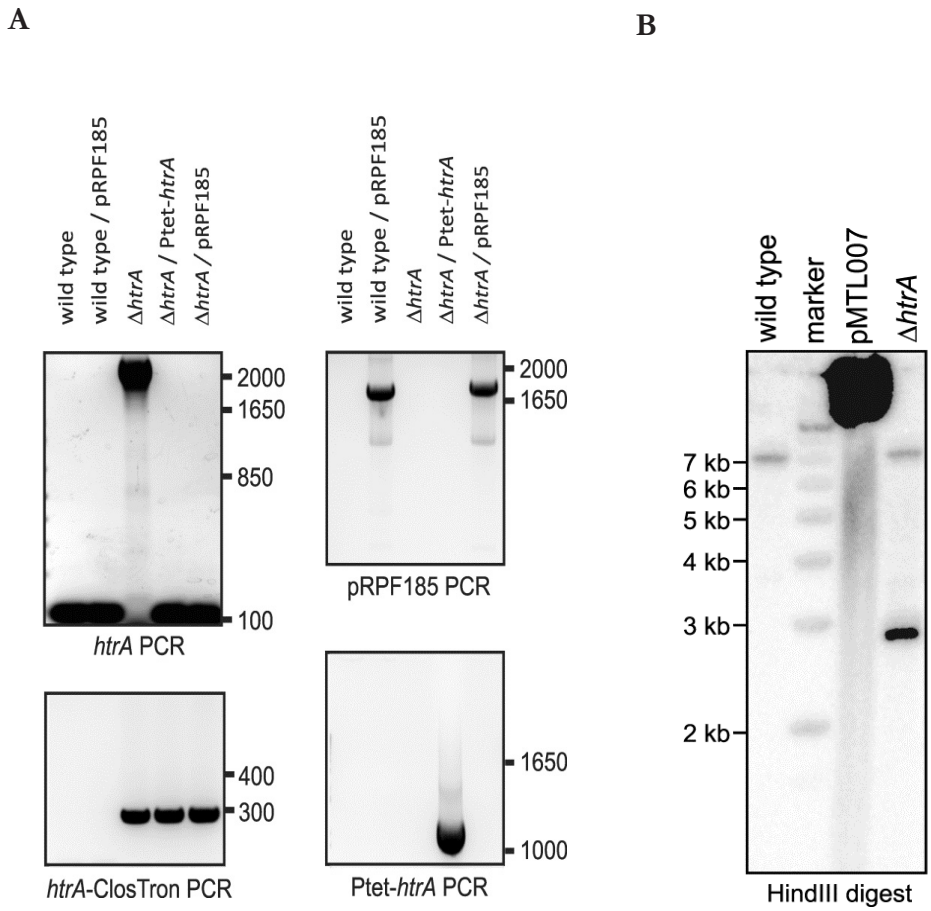


Figure 2: (A) PCR screening of the strains used in this study. *HtrA* PCR: The primers oDB0123 and oDB0124 generated a PCR product of approximately 2200 bp in the *htrA* mutant (Δ*htrA*, DB0002) and a PCR product of 150 bp in the wild type (WKS1241), wild type/pRPF185 (DB0051), complemented *htrA* mutant (Δ*htrA*/*htrA*+, DB0047) and Δ*htrA*/pRPF185 (DB0052). pRPF185 PCR: The primers oWKS-1177 and NF794 generated a PCR product of approximately 1800 bp in the DB0051 and DB0052 strains. *htrA*-ClosTron PCR: The primers oDB0123 and EBSuni generated a PCR product of approximately 280 bp in DB0002, DB0047 and DB0052 strains. Ptet-*htrA* PCR: The primers oDB0123 and NF794 generated a PCR product of approximately 1200 bp in DB0047. **(B)** Southern blot analysis of C. difficile wild type and Δ*htrA*. Chromosomal DNA digested of wild type (WKS1241) and Δ*htrA* (DB0002) with HindIII was hybridized with an *ermB* specific probe. Note that wild type shows a single band, indicative of the *ermB* element that is still present in this strain (29), Δ*htrA* contains an additional element, pMTL007 is shown as a positive control.

amino-acid identity, the structural homology of this protein to the HtrA-like protease DegQ (49,50), and the predicted structure of *C. difficile* HtrA overlaid with the structure of *E.coli* DegS demonstrated a high level of structure homology (Figure 1A). Homologs of DegQ and DegS are generally referred to as HtrA-like proteases in Gram-positive prokaryotes. Considering the above, we will refer hereafter to the CD3284 encoding gene as *htrA*, and the CD3284 protein as HtrA.

C. difficile HtrA (CD3284) shows protease activity towards β-casein

To investigate if *C. difficile* HtrA possesses proteolytic activity similar to that of other HtrA-like proteases (51) we performed a protease assay. Wild type HtrA (His10-Δ(1-30)-HtrA) and a catalytic HtrA mutant (His10-Δ(1-30)-HtrA-S217A) protein were incubated with β-casein, a highly unfolded protein substrate (Figure 1B). After 18 hour of incubation at 37 °C casein was only degraded by wild type HtrA and not by the catalytic HtrA mutant. These results demonstrate the proteolytic activity of *C. difficile* HtrA.

Construction of an *htrA* mutant

To determine the contribution of HtrA to *C. difficile* virulence we constructed an *htrA* mutant using ClosTron technology in the 630Δ*erm* (wild type) (29) strain by insertion of a Group II intron upstream of the sequence encoding the predicted trypsin-like serine protease domain (Figure 1A). Insertional inactivation of genes using the ClosTron technology results in stable mutants and non- functional proteins (32,40). The genotype of the insertional inactivation was confirmed by PCR (Figure 2A), sequence analysis and Southern blot (Figure 2B), which verified that the ClosTron group II intron was inserted at a single site. Initially, the impact of the mutation on growth *in vitro* was considered. This revealed that post inoculation

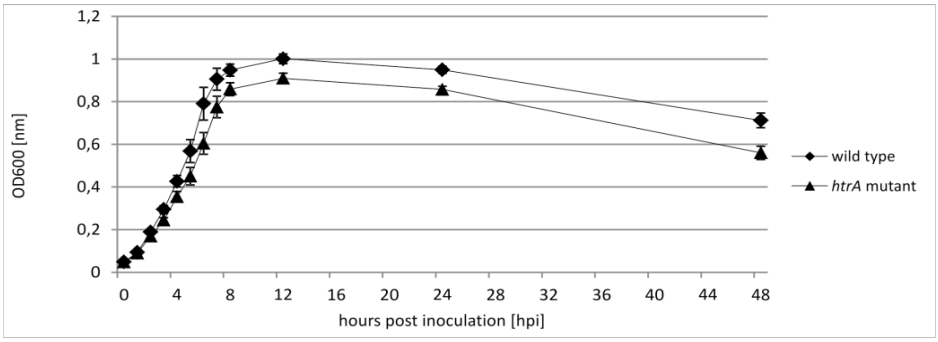


Figure 3: Growth curves of the wild type and the *htrA* mutant strain. The absorbance (OD_{600}) was measured over 48 hours of growth in TTY medium. The error bars indicated the standard error of the mean of three experiments.

both mutant and wild type strains showed logarithmic growth for the first 8 hour and by 12 hour post inoculation both strains had entered stationary growth phase (Figure 3). Though growth kinetics were not significantly different from wild type cells, we noted that the *htrA* mutant strain generally showed a slight growth delay in logarithmic phase and reached lower optical density in stationary phase (Figure 3).

The *htrA* null mutant is more virulent in hamsters

HtrA like proteases have been shown to be important for the virulence of pathogenic bacteria (25,52,53). To determine the contribution of HtrA in *C. difficile* virulence we performed *in vivo* experiments with the *htrA* mutant and wild type in the commonly used Golden Syrian hamster model (54,55). Six clindamycin pre-treated hamsters were orally infected either with wild type or the *htrA* mutant. The mean time to death for hamsters challenged with wild-type was 53 h and 39 min (± *5 h and 3 min; n=6*). Hamsters infected with the *htrA* mutant revealed a significant decreased survival time (31 h 20 min ± 44 min, n=6, p=0.0005) compared to wild type (Figure 4), suggesting that disruption of *htrA* causes enhanced virulence in this strain. Our experience evaluating a number of mutants in the hamster model suggests that animals challenged with other mutants created using ClosTron

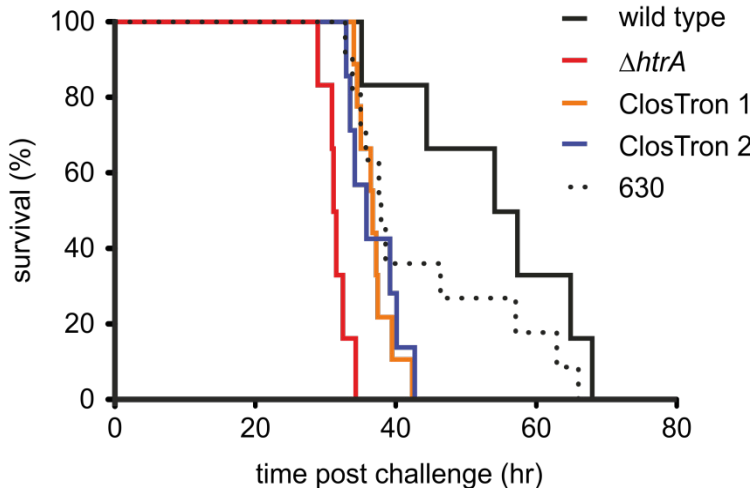


Figure 4: The *htrA* null mutant has increased virulence in a hamster model of infection. Survival curve of hamsters challenged with either 10^4 spores of *C. difficile* 630Δ*erm* (solid black line), *htrA* null mutant (red line), ClosTron mutant 1 & 2 (blue & green lines, respectively) or 630 wild-type (dashed black line). Hamsters challenged with *htrA* mutant (31 h 20 min, n=6) showed a significantly decreased survival time compared to those challenged with 630Δ*erm* (53 h 39 min, n=6, p=0.0005), ClosTron mutant 1 (36 h 41 min, n=7, p=0.049), ClosTron mutant 2 (36 h 47 min, n=9, p<0.0001) or 630 (43 h 43 min, n=11, p<0.0001). Hamsters were culled when body temperature declined below 35°C (clinical end point) and data analysed by Mantel-Cox Log-rank statistical tests.

technology have a shorter survival time compared to wild type (unpublished data). We hypothesize that this observation is due to the greater resistance to clindamycin (as a consequence of ClosTron insertion) and hence greater survival rates *in vivo* following pretreatment of animals with clindamycin to induce disease. Indeed, we find that the *htrA* mutant (and other ClosTron mutants) *in vitro* is more resistant to clindamycin (MIC >256 μg/mL) than the wild type (630Δ*erm*; MIC = 16 μg/mL).

Nevertheless, if we compared the *htrA* mutant to the original parental 630 isolate (Wust *et al*., 1982) (43 h 43 min, n=11, $p<0.0001$) or to other mutants (ClosTron 1, 36 h 41 min, n=7, $p=0.0049$; ClosTron 2, 36 h 47 min, n=9, $p<0.0001$) infected in the same way, the enhanced mortality rate remains significant. We therefore conclude that in the Golden Syrian hamster model the *htrA* null mutant of *C. difficile* demonstrates an enhanced virulence phenotype.

Comparative transcriptional micro-array data analysis of *htrA* mutant and wild type strains

To understand the enhanced virulence of the *htrA* mutant in the Golden Syrian hamster model, we performed comparative transcriptome analysis of the wild type and the *htrA* mutant. The *htrA* mutant displayed a pleiotropic effect in the transcriptomic analysis. The most interesting genes that were differentially regulated in the *htrA* mutant were *tcdA*, *spo0A*, and several genes encoding surface associated proteins.

The transcriptional profiles of the *htrA* mutant strain and wild type strain were studied at two different time points, the logarithmic and stationary growth phase respectively. The DNA micro-array raw data sets and the normalized ratio data are available from the NCBI GEO database (http://www.ncbi.nlm.nih.gov/geo/, accession number GSE55926) (56). In the logarithmic growth phase and in the stationary growth phase 551 genes and 567 genes were significantly differentially expressed, respectively (fold change ≥1.5, $p \leq 0.05$). In total 248 genes in the logarithmic growth phase and 263 genes in the stationary growth phase were up-regulated. Two-hundred and eight genes were up-regulated both in logarithmic and stationary growth phase. Among these, several genes were involved in stress adaptation response, like the *ctsR* and *hrcA* genes (57). The *tcdA* gene encoding the well-characterized virulence factor Toxin A was 3.3-fold and 3.0-fold up-regulated in the logarithmic and stationary growth phase, respectively.

A total of 303 genes in the logarithmic growth phase and 304 genes in the stationary growth phase were down-regulated in the *htrA* mutant. Two-hundred and sixty five genes were down-regulated both in logarithmic and stationary growth phase. Among these are genes involved in cell metabolism, like ribosomal proteins, ABC transporters and PTS transporters. The *spo0A* gene, encoding the

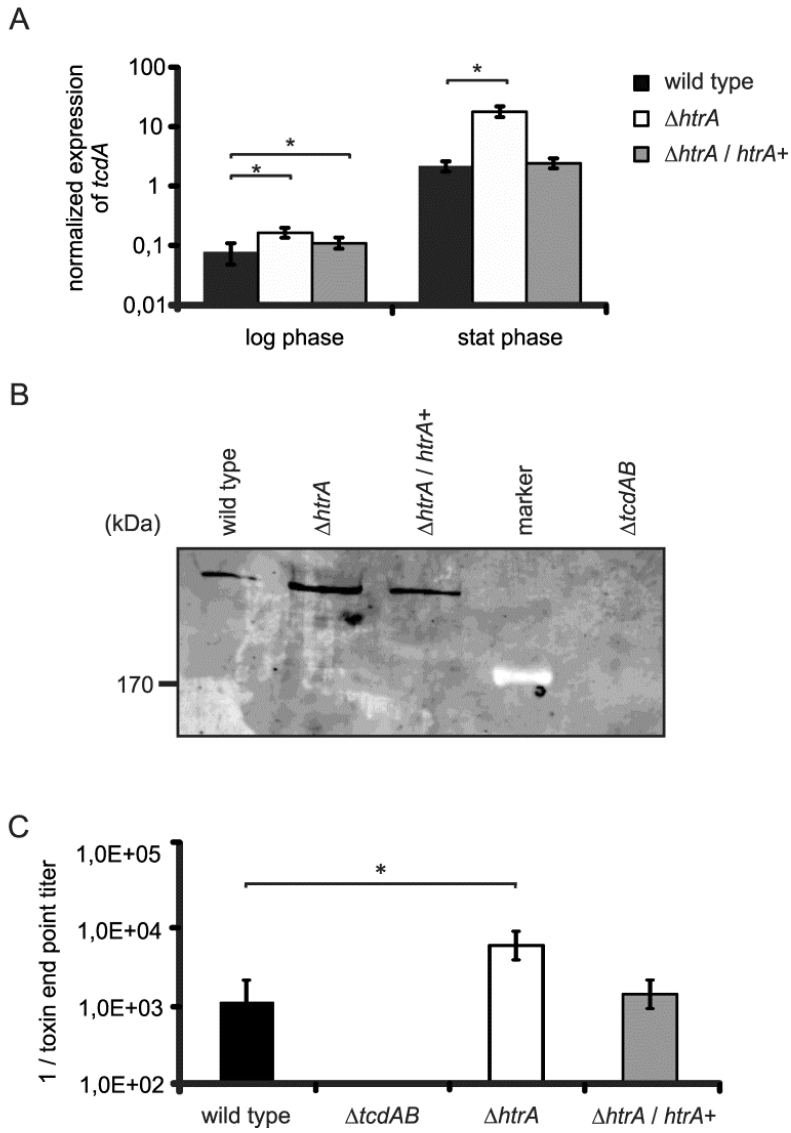


Figure 5: Increased TcdA production of *htrA null* mutant compared to wild type. **(A)** The relative *tcdA* gene expression profiles of *C. difficile* strains in logarithmic and stationary phase. The asterisks (*) indicate a significant difference between wild type and *htrA* mutant ($p<0.000001$) and complemented *htrA* mutant ($p<0.000001$). **(B)** Quantitative Western blot analysis of TcdA in culture supernatant of wild type, *htrA* null mutant and complemented *htrA* mutant at 48 hour post inoculation. **(C)** The toxin end point titers of the various *C. difficile* strains. The asterisk (*) indicate a significant difference between wild type and *htrA* null mutant ($p=0.001$) at the same time point. The error bars indicate the standard error of the mean of three experiments. Wild type is WKS1241, *htrA* mutant is DB0002, complemented *htrA* mutant is DB0047 and *tcdA/B* mutant is WKS1710.

master regulator of sporulation (58,59), was 2.4-fold and 2.3-fold down-regulated in the logarithmic and stationary growth phase, respectively, and many other sporulation genes showed a similar effect. Furthermore, several genes encoding surface-associated proteins (like *slpA* and 5 other paralogues of *slpA*) (16), flagella associated genes (18,60), and CD1581 (a gene with a potential effect on adhesion and intestinal colonization) (19) were significantly down regulated.

Overall, transcriptional profiling of the *htrA* null mutant showed a plethora of differences compared to wild type. The most prominent differentially regulated genes were *tcdA*, *spo0A* and several genes encoding surface associated proteins that were further investigated phenotypically.

Increased TcdA production of *htrA* mutant compared to wild type

Golden Syrian hamsters are exquisitely sensitive to *C. difficile* toxins and increased levels of *tcdA* transcription might explain the observed enhanced virulence. To validate the increased level of *tcdA* transcription in the comparative DNA microarray analysis we determined the *tcdA* transcription by RT-qPCR, quantitative Western blot analysis of TcdA and a cytotoxicity assay.

First, we confirmed the effect on transcription of the *tcdA* gene in the *htrA* mutant. We also complemented the *htrA* mutant (DB0047) to confirm that the observed effect on *tcdA* transcription in the *htrA* mutant was specific. RT-qPCR was used to compare the normalized *tcdA* transcription profiles in wild type, *htrA* mutant and the complemented *htrA* mutant strains in the logarithmic and stationary growth phase. In the logarithmic phase and the stationary phase the normalized *tcdA* transcription of the *htrA* mutant was significantly increased, 2.1-fold ($p<0.000001$) and 8.1-fold ($p<0.000001$), respectively (Figure 5A). Expressing HtrA from a plasmid in the mutant strain (DB0047) partially restored *tcdA* transcript levels in logarithmic phase (1.4-fold, $p=0.0022$) towards wild type level, whilst in stationary phase the *tcdA* transcript level was indistinguishable from wild type levels (Figure 5A). These data suggest that reduction in transcription of *tcdA* in the *htrA* mutant is a consequence of the *htrA* mutation. Second, we determined if the increased level of *tcdA* transcription was reflected in elevated levels of TcdA protein. Quantitative Western blot analysis with a TcdA specific monoclonal antibody on filtered supernatants of 48 hour post inoculation cultures showed a 7-fold increase of TcdA in the *htrA* mutant compared to wild type (Figure 5B). Similar to *tcdA* transcription, complementation of the *htrA* mutant resulted in TcdA levels that were partially restored towards the wild type (2-fold, Figure 5B). The signals detected were specific for TcdA as they were absent in the supernatant of the control strain WKS1710 (40).

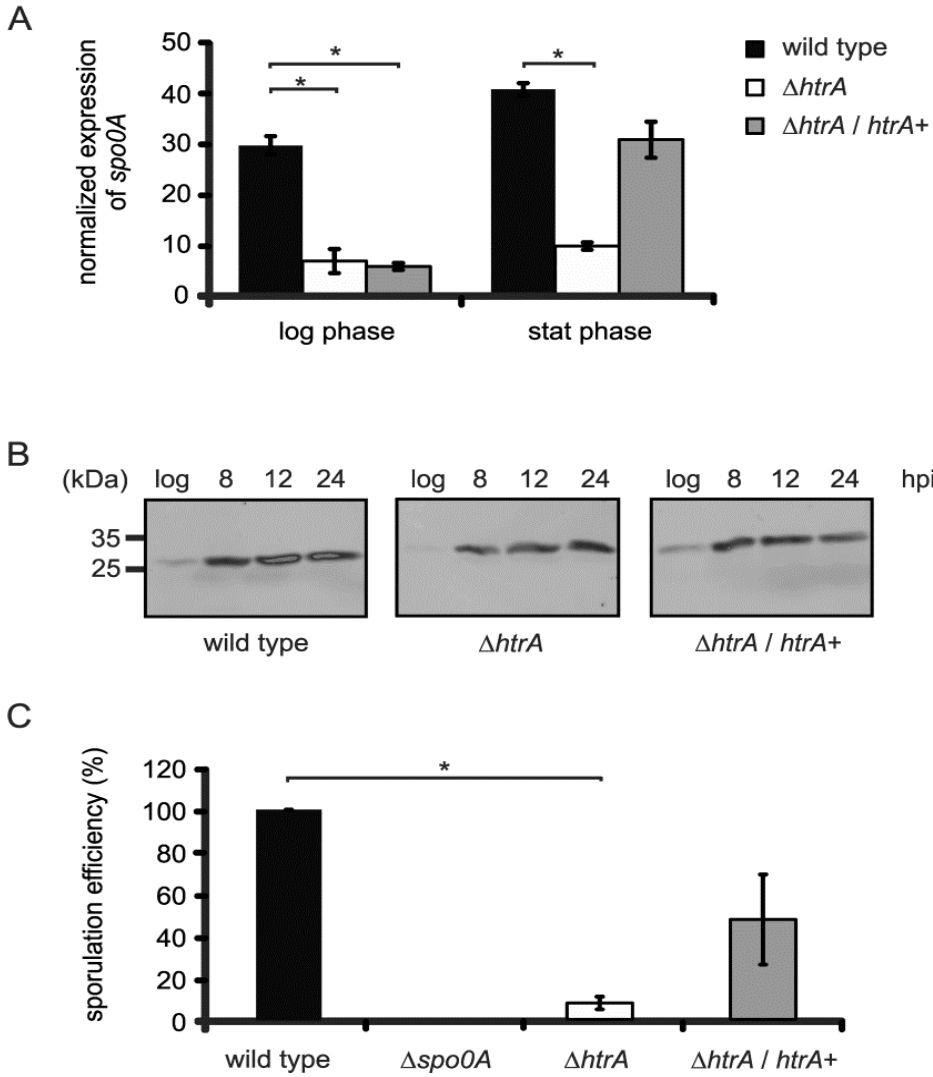


Figure 6: Decreased sporulation of *htrA* mutant compared to wild type. **(A)** The relative *spo0A* gene expression profiles of *C. difficile* strains in logarithmic and stationary phase. The asterisks (*) indicate a significant difference between wild type and *htrA* mutant (*p<0.000001*) and the complemented *htrA* mutant (*p<0.000001*) at the same time point. **(B)** The quantitative Western blot analysis of Spo0A in wild type and *htrA* mutant and the complemented *htrA*mutant in time. **(C)** Normalized sporulation efficiencies (heat resistant CFU compared to wild type) of the different strains at 24h post inoculation. The asterisk (*) indicates a significant difference between wild type and *htrA* mutant (*P=0.0006*). The error bars indicate the standard error of the mean of three experiments. Wild type is WKS1241, *htrA* mutant is DB0002, complemented *htrA* mutant is DB0047 and the *spo0A* mutant is WKS1237.

We also determined the effect of the *htrA* mutant in a cytotoxicity assay on Vero cells, which detects the cumulative effects of both Toxin A and B (28,40) (Figure 5C). We observed a 5-fold increase in toxin end point titer for the *htr*A mutant (p=0.001) compared to the wild type (Figure 5C). Complementing the *htrA* mutant with an inducible copy of *htrA* on a plasmid (DB0047) restored the toxin end point titer to near-wild type levels. To determine these effects as toxin specific, Vero cells were exposed to filter sterilized supernatant of strain WKS1710 (630 ⊿*tcd*A⊿*tcdB*) (40), which as expected did not result in cytotoxic effect. Moreover, pre-incubation of the filter sterilized bacterial supernatants with an anti-toxin (Techlab, Blacksburg, VA, USA) resulted in complete neutralization of the observed cytotoxic effects on the Vero cell monolayers (data not shown). Thus, the effects of the supernatants on Vero cells were Toxin A and/or Toxin B specific.

In summary, we found increased *tcdA* transcription levels, increased TcdA expression levels and elevated toxin end point titers of supernatants derived from *htrA* mutant cells compared to wild type. Taken together, these data contribute to an explanation for the enhanced virulence of the *htrA* mutant strain *in vivo*.

The *htrA* mutant displays decreased sporulation compared to wild type

Our transcriptome analysis revealed pleiotropic effects that included a decreased transcription of *spo0A*, the master regulator of sporulation, in the *htrA* null mutant compared to wild type (58,59). Sporulation is highly relevant for transmission, persistence and biofilm formation of *C. difficile* (61,62). To validate the microarray results, we performed RT-qPCR, quantitative Western blot analysis of Spo0A and a sporulation assay.

We used a RT-qPCR to determine the normalized *spo0A* transcript levels in wild type, *htrA* mutant and the complemented *htrA* mutant strains in the logarithmic and stationary growth phase. The *spo0A* transcript level in the *htrA* mutant was significantly decreased in the logarithmic phase (4.3-fold, p<0.000001) and stationary phase (4.1-fold, p<0.000001) compared to wild type (Figure 6A). The complemented *htrA* mutant failed to reach wild type levels of *spo0A* transcript in the log phase (5.2 fold difference, p<0.000001), but in the stationary phase the complemented strain reached wild type levels of *spo0A* transcript (Figure 6A). Therefore, we believe that the reduced *spo0A* transcript level in the *htrA* mutant is partially, if not completely, due to the *htrA* mutation.

Next, we quantified the decreased Spo0A levels by quantitative Western blot analysis (31). In wild type cells the Spo0A levels increased during growth and signals were absent from the *spo0A* mutant strain as previously reported (Figure 6B) (31). In the *htrA* mutant, Spo0A levels were on average 3-fold lower compared to wild type (Figure 6B). Complementation of the *htrA* mutant restored Spo0A protein levels throughout the growth to near-wild type levels (Figure 6B).

We also investigated the sporulation efficiency of the various strains by determining heat resistant colony forming units (CFU). The efficiency of heat treatment on spore recovery was initially tested using cultures of mutant (32). No colonies could be recovered using this strain, indicating that this is a good way to measure actual spore formation. At 24 hour post inoculation we observed (12-fold) decreased sporulation efficiency for the *htrA* mutant compared to wild type (significant, p=0.0006) (Figure 6C). Expressing *htrA* from a plasmid in the *htrA* mutant strain (DB0047) partly restored sporulation efficiency to wild type levels (Figure 6C). In contrast, sporulation efficiency at 48 hour post inoculation appeared comparable in all strains tested, including the *htrA* mutant (data not shown).

In summary, we show that the *htrA* mutant has a decreased *spo0A* transcript level, a delayed expression of Spo0A protein and a significant reduction in the number of spores formed at 24h post inoculation.

Decreased adherence of *htrA* mutant to Caco2 cell monolayer compared to wild type

In addition to spores, it has been suggested that surface proteins on vegetative cells are important factors for adherence of *C. difficile* to the colonic epithelium and colonization of the intestine (16,19,63). Prompted by the down-regulation of several important cell surface proteins in our transcriptome analyses, we determined the capability of the *htrA* mutant strains to adhere to colon epithelial cells.

We incubated the various strains on a monolayer of Caco-2 cells and determined the fraction of bacterial cells that remained attached after rigorous washing. We observed only 10% adherence of the *htrA* mutant to the Caco-2 cell monolayer (p=0.000001) compared to wild type (Figure 78). Inducing DB0047 resulted in a

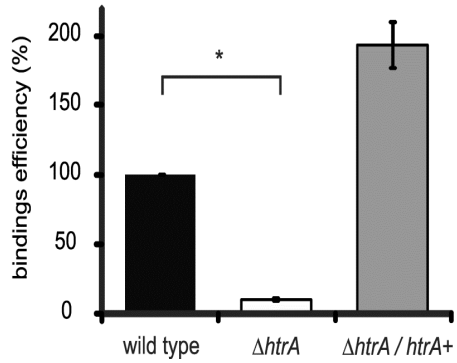


Figure 7: Decreased adhesion of *htrA* null mutant to Caco2 cells. The asterisk (*) indicates a significant difference between wild type and *htrA* mutant (*p=0.000001*). The error bars indicate the standard error of the mean of three experiments. Wild type is WKS1241, Δ*htrA* mutant is DB0002, Δ*htrA*/*htrA*+ is the complemented mutant (DB0047).

slightly increased, but not significant (p=0.17), capability to adhere to a monolayer of Caco-2 cells compared to wild type levels, indicating that the adherence defect of the *htrA* mutant is specific.

Discussion

Role for HtrA in virulence of *C. difficile*

In addition to the prevalence, the severity and mortality of disease caused by *C. difficile* infections has increased significantly in the last decade (1,6,9). It is unclear whether this is a consequence of more diligent screening or a consequence of acquisition of additional virulence traits by the bacteria. The retractable nature of the organism to genetic manipulation has limited opportunities to investigate the role of many genes in the control of virulence traits including the two large clostridial toxins, the binary toxin and other factors that contribute to the virulence of *C. difficile,* such as those affecting colonization and survival of the bacteria in the host (15,16,18,19). In multiple organisms HtrA-like foldase/proteases play a considerable role in virulence by controlling protein homeostasis (21,22,26,52,64). Here, we find that *C. difficile* encodes a single HtrA-like protease. In contrast to observations in other organisms, we have found that a *C. difficile htrA* mutant shows enhanced virulence in the Golden Syrian hamster model of acute *C. difficile* infection (Figure 4).

The observed increase in virulence is likely to be a consequence of elevated toxin levels. Toxin measurements from the hamsters showed no differences in measurement of toxin production in filtered gut samples, however these samples were taken at the endpoint of infection when toxin levels reflect significant symptoms. *In vitro*, we find increased transcription of *tcdA*, increased levels of TcdA protein in the *htrA* mutant, and increased cytotoxicity of the filtered supernatant derived from *htrA* mutant cells, all of which could be substantially reversed by prolonged induction of a plasmid-located copy of *htrA* (Figure 5). Shorter induction did not result in full complementation, most likely because expression levels or the temporal pattern of expression does not match that of wild type.

Our transcriptome analysis did not reveal a significant change in *tcdB* levels in the *htrA* mutant. This has been observed before, and is probably due to lower transcription levels of *tcdB* compared to *tcdA* in general though our data cannot exclude differential regulation of the two toxin genes (28,65,66).

We did notice changes in transcription of several other virulence-associated genes. We validated a number of phenotypes that could be related to these genes experimentally, including a reduction in sporulation frequency and decrease in

colon epithelial cell adherence. In agreement with the observation for TcdA, these effects could be reversed by complementation (Figures 6 and 7). Whilst the most significant impact of this mutation in the hamster model appears to be the change of toxin expression, this does not reflect the potential role of HtrA more globally on gene expression.

The delay in sporulation, as measured by enumeration of colonies on taurocholate containing plates after heat shock, does not necessarily reflect a block in sporulation, but could also be the result of a defect in germination. However, the complementary observation of decrease in *spo0A* transcript level in the mutant strain would suggest the impact to be on sporulation. In addition, this delay in sporulation could result in extending the time spentin the vegetative state during which more toxin than wild type cells could be produced.

Structure of *C. difficile* HtrA

The HtrA family of proteases can be distinguished from other serine proteases by the presence of at least one PDZ domain (21,23). PDZ domains are involved in the regulation of the activity of the protein through protein-protein interactions or binding of specific substrates (22,50). HtrA proteins in other organisms can either be membrane associated or secreted (49,50). A detailed bioinformatics analysis of *C. difficile* HtrA suggests that the protein is a membrane-associated protein with a single PDZ domain (Figure 1A).

HtrA-like proteases of *E. coli* are among the best-characterized members of the family. DegP and DegQ consist of single protease domains with two associated PDZ domains (22,49). Both proteins are foldase/proteases that have relatively broad substrate specificity. DegS is a protease with a single PDZ domain, that acts specifically as a site-1 protease in a regulatory pathway for the extracytoplasmic stress response (49). Considering the domain structure, one might expect HtrA of *C. difficile* to be a site-1 protease. However, we would argue against this based on three findings; first, our protease assay with a non *C. difficile* related unfolded protein suggests a non-specific proteolytic activity. Second, our transcriptome analyses suggest a pleiotropic effect, rather than a specific pathway to be affected. Third, homology modeling of the predicted structure of *C. difficile* HtrA using I-TASSER and Phyre2 suggests that the overall fold of the catalytic domain and PDZ domain is more similar to the DegQ/P subfamily of HtrA-like proteases (22,24,49,50). Finally, site-1 proteases in Gram-positive bacteria generally belong to the PrsW family (67-69). Indeed, in *C. difficile* PrsW (CD0552) is involved in the activation of the extracytoplasmic function sigma factor CsfT (70).

HtrA and virulence

In the majority of pathogenic bacteria, HtrA is directly or indirectly implicated in virulence. In many bacteria *htrA* null mutants show a complete loss of virulence, or at least a significant level of attenuation (21,22,25,26,52). To our knowledge, our study is the first example of increased virulence for an *htrA* null mutant.

Only a few studies have found little or no effect of *htrA* on virulence in animal models. In *Staphylococcus aureus* COL, the deletion of *htrA* led to minimal changes in surface protein expression and virulence (71). In *Brucella abortis* the *htrA* mutant was not found to be attenuated in a BALB/c mouse model (72). For *Porphyromonas gingivalis* only late effects were observed in a competition assay in mice (73). For *Streptococcus pyogenes* no effects were seen in a murine model of subcutaneous infection (74). In these cases, *htrA* mutant strains did display an increased sensitivity to stresses, or an altered expression of virulence factors. However, this may reflect the choice of model and environmental niche tested.

It should be noted that the effects of HtrA can be strain dependent (26,71), and may be masked by additional *htrA* like genes that can compensate in site specific deletion mutants (75). In the case of *C. difficile*, we have not identified any additional HtrA-like genes (data not shown), but we cannot exclude the possibility that other (serine) protease(s) might be able to substitute for certain functions of HtrA. The contribution of the sporulation and adhesion phenotypes of the *htrA* mutant of *C. difficile* to virulence therefore awaits validation in a model other than the hamster, in which toxin production is so dominant that observation of changes to other traits can be limited. For example, analysis of this mutant in mice, which can be colonized but are less susceptible to toxin production, may be a good alternative to study the effects of *htrA* on colonization, transmission and persistence (15,35).

Mode of action of HtrA

In most pathogens, attenuation as a consequence of mutation in *htrA* appears to be the result of a reduced capacity of the cells to deal with environmental or host-dependent stresses (e.g. heat or oxidative stress), a reduced level of virulence factors that depend on HtrA for processing or activation, and/or a reduced invasiveness of the bacterium (21,24,50). In some cases, the effects of HtrA have not been traced to a specific protein or pathway (76,77).

Our study revealed broad effects of the mutation of *htrA* on multiple pathways. Interestingly, several of these pathways have been linked to each other in previous work. For instance, it was noted that the asporogenic mutant of *spo0A*, showed elevated levels of toxin production in PCR ribotype 027, but not in 630Δ*erm* (61,78). Similarly, there is a clear link between motility and toxin production as the alternative sigma factor SigD is required for both (79). Interestingly, mutants

in the flagellar operon showed changes in both toxin levels (80) and their capacity to adhere to Caco-2 cells (60). Finally, more general effects of stress (nutrient depletion, temperature) on toxin levels were noted (81,82).

Our study shows that *htrA* of *C. difficile* can modulate toxin levels, the formation of spores and the adhesion to colonic cells, all of which are highly relevant for virulence. However, the high connectivity of the networks regulating these processes does not allow us to draw any conclusions as to the direct targets of HtrA in *C. difficile*. Future research will be aimed at identifying which proteins are direct targets of the foldase/protease activity of *C. difficile* HtrA.

Acknowledgments

Wiep Klaas Smits was supported by Veni grant 016.116.043 from NWO-ZonMW and an LUMC Gisela Thier fellowship. Anthony M. Buckley is funded by a Wellcome Trust grant 086418. We thank N. Minton and S. Kuehne for providing the ClosTron system and the toxin A/B mutant. We thank R. Fagan and N. Fairweather for providing the pRF185 plasmid.

Reference List

1. Loo, V. G., L. Poirier, M. A. Miller, M. Oughton, M. D. Libman, S. Michaud, A. M. Bourgault, T. Nguyen, C. Frenette, M. Kelly, A. Vibien, P. Brassard, S. Fenn, K. Dewar, T. J. Hudson, R. Horn, P. Rene, Y. Monczak, and A. Dascal. 2005. A predominantly clonal multi-institutional outbreak of *Clostridium difficile*-associated diarrhea with high morbidity and mortality. N.Engl.J.Med. 353:2442-2449.
2. Kyne, L., M. B. Hamel, R. Polavaram, and C. P. Kelly. 2002. Health care costs and mortality associated with nosocomial diarrhea due to *Clostridium difficile*. Clin.Infect.Dis. 34:346-353.
3. Polage, C. R., J. V. Solnick, and S. H. Cohen. 2012. Nosocomial diarrhea: evaluation and treatment of causes other than *Clostridium difficile*. Clin.Infect.Dis. 55:982-989.
4. Goorhuis, A., K. T. van der, N. Vaessen, F. W. Dekker, B. R. van den, C. Harmanus, H. S. van den, D. W. Notermans, and E. J. Kuijper. 2007. Spread and epidemiology of *Clostridium difficile* polymerase chain reaction ribotype 027/toxinotype III in The Netherlands. Clin.Infect.Dis. 45:695-703.
5. Kuijper, E. J., B. Coignard, and P. Tull. 2006. Emergence of *Clostridium difficile*-associated disease in North America and Europe. Clin.Microbiol.Infect. 12 Suppl 6:2-18.
6. Kuijper, E. J., F. Barbut, J. S. Brazier, N. Kleinkauf, T. Eckmanns, M. L. Lambert, D. Drudy, F. Fitzpatrick, C. Wiuff, D. J. Brown, J. E. Coia, H. Pituch, P. Reichert, J. Even, J. Mossong, A. F. Widmer, K. E. Olsen, F. Allerberger, D. W. Notermans, M. Delmee, B. Coignard, M. Wilcox, B. Patel, R. Frei, E. Nagy, E. Bouza, M. Marin, T. Akerlund, A. Virolainen-Julkunen, O. Lyytikainen, S. Kotila, A. Ingebretsen, B. Smyth, P. Rooney, I. R. Poxton, and D. L. Monnet. 2008. Update of *Clostridium difficile* infection due to PCR ribotype 027 in Europe, 2008. Euro.Surveill 13.
7. Paltansing, S., R. J. van den Berg, R. A. Guseinova, C. E. Visser, E. R. van der Vorm, and E. J. Kuijper. 2007. Characteristics and incidence of *Clostridium difficile*-associated disease in The Netherlands, 2005. Clin.Microbiol.Infect. 13:1058-1064.
8. Pepin, J., L. Valiquette, M. E. Alary, P. Villemure, A. Pelletier, K. Forget, K. Pepin, and D. Chouinard. 2004. *Clostridium difficile*-associated diarrhea in a region of Quebec from 1991 to 2003: a changing pattern of disease severity. CMAJ. 171:466-472.
9. Goorhuis, A., D. Bakker, J. Corver, S. B. Debast, C. Harmanus, D. W. Notermans, A. A. Bergwerff, F. W. Dekker, and E. J. Kuijper. 2008. Emergence of *Clostridium difficile* infection due to a new hypervirulent strain, polymerase chain reaction ribotype 078. Clin.Infect.Dis. 47:1162-1170.
10. Just, I. and R. Gerhard. 2004. Large clostridial cytotoxins. Rev.Physiol Biochem. Pharmacol. 152:23-47.
11. Voth, D. E. and J. D. Ballard. 2005. *Clostridium difficile* toxins: mechanism of action and role in disease. Clin.Microbiol.Rev. 18:247-263.
12. Jank, T., T. Giesemann, and K. Aktories. 2007. Rho-glucosylating *Clostridium difficile* toxins A and B: new insights into structure and function. Glycobiology 17:15R-22R.
13. Kuehne, S. A., M. M. Collery, M. L. Kelly, S. T. Cartman, A. Cockayne, and N. P. Minton. 2013. Importance of Toxin A, Toxin B, and CDT in Virulence of an Epidemic *Clostridium difficile* Strain. J.Infect.Dis.
14. Schwan, C., B. Stecher, T. Tzivelekidis, H. M. van, M. Rohde, W. D. Hardt, J. Wehland, and K. Aktories. 2009. *Clostridium difficile* toxin CDT induces formation of microtubule-based protrusions and increases adherence of bacteria. PLoS.Pathog. 5:e1000626.

15. Barketi-Klai, A., S. Hoys, S. Lambert-Bordes, A. Collignon, and I. Kansau. 2011. Role of fibronectin-binding protein A in *Clostridium difficile* intestinal colonization. J.Med. Microbiol. 60:1155-1161.
16. Calabi, E., F. Calabi, A. D. Phillips, and N. F. Fairweather. 2002. Binding of *Clostridium difficile* surface layer proteins to gastrointestinal tissues. Infect.Immun. 70:5770-5778.
17. Janoir, C., S. Pechine, C. Grosdidier, and A. Collignon. 2007. Cwp84, a surface-associated protein of *Clostridium difficile*, is a cysteine protease with degrading activity on extracellular matrix proteins. J.Bacteriol. 189:7174-7180.
18. Tasteyre, A., M. C. Barc, A. Collignon, H. Boureau, and T. Karjalainen. 2001. Role of FliC and FliD flagellar proteins of *Clostridium difficile* in adherence and gut colonization. Infect.Immun. 69:7937-7940.
19. Janoir, C., C. Deneve, S. Bouttier, F. Barbut, S. Hoys, L. Caleechum, D. Chapeton-Montes, F. C. Pereira, A. O. Henriques, A. Collignon, M. Monot, and B. Dupuy. 2013. Adaptive Strategies and Pathogenesis of *Clostridium difficile* from In Vivo Transcriptomics. Infect. Immun. 81:3757-3769.
20. Fang, F. C. 2004. Antimicrobial reactive oxygen and nitrogen species: concepts and controversies. Nat.Rev.Microbiol. 2:820-832.
21. Clausen, T., M. Kaiser, R. Huber, and M. Ehrmann. 2011. HTRA proteases: regulated proteolysis in protein quality control. Nat.Rev.Mol.Cell Biol. 12:152-162.
22. Malet, H., F. Canellas, J. Sawa, J. Yan, K. Thalassinos, M. Ehrmann, T. Clausen, and H. R. Saibil. 2012. Newly folded substrates inside the molecular cage of the HtrA chaperone DegQ. Nat.Struct.Mol.Biol. 19:152-157.
23. Hansen, G. and R. Hilgenfeld. 2013. Architecture and regulation of HtrA-family proteins involved in protein quality control and stress response. Cell Mol.Life Sci. 70:761-775.
24. Dalbey, R. E., P. Wang, and J. M. van Dijl. 2012. Membrane proteases in the bacterial protein secretion and quality control pathway. Microbiol.Mol.Biol.Rev. 76:311-330.
25. Chitlaru, T., G. Zaide, S. Ehrlich, I. Inbar, O. Cohen, and A. Shafferman. 2011. HtrA is a major virulence determinant of Bacillus anthracis. Mol.Microbiol. 81:1542-1559.
26. Ibrahim, Y. M., A. R. Kerr, J. McCluskey, and T. J. Mitchell. 2004. Role of HtrA in the virulence and competence of Streptococcus pneumoniae. Infect.Immun. 72:3584-3591.
27. Ingmer, H. and L. Brondsted. 2009. Proteases in bacterial pathogenesis. Res.Microbiol. 160:704-710.
28. Bakker, D., W. K. Smits, E. J. Kuijper, and J. Corver. 2012. TcdC Does Not Significantly Repress Toxin Expression in *Clostridium difficile* 630DeltaErm. PLoS.One. 7:e43247.
29. Hussain, H. A., A. P. Roberts, and P. Mullany. 2005. Generation of an erythromycin-sensitive derivative of *Clostridium difficile* strain 630 (630Deltaerm) and demonstration that the conjugative transposon Tn*916*DeltaE enters the genome of this strain at multiple sites. J.Med.Microbiol. 54:137-141.
30. Miroux, B. and J. E. Walker. 1996. Over-production of proteins in Escherichia coli: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. J.Mol.Biol. 260:289-298.
31. Rosenbusch, K. E., D. Bakker, E. J. Kuijper, and W. K. Smits. 2012. C. difficile 630Deltaerm Spo0A regulates sporulation, but does not contribute to toxin production, by direct high-affinity binding to target DNA. PLoS.One. 7:e48608.

32. Heap, J. T., O. J. Pennington, S. T. Cartman, G. P. Carter, and N. P. Minton. 2007. The ClosTron: a universal gene knock-out system for the genus *Clostridium* . J.Microbiol. Methods 70:452-464.
33. Heap, J. T., S. A. Kuehne, M. Ehsaan, S. T. Cartman, C. M. Cooksley, J. C. Scott, and N. P. Minton. 2010. The ClosTron: Mutagenesis in *Clostridium* refined and streamlined. J.Microbiol.Methods 80:49-55.
34. Purdy, D., T. A. O'Keeffe, M. Elmore, M. Herbert, A. McLeod, M. Bokori-Brown, A. Ostrowski, and N. P. Minton. 2002. Conjugative transfer of clostridial shuttle vectors from Escherichia coli to *Clostridium difficile* through circumvention of the restriction barrier. Mol.Microbiol. 46:439-452.
35. Fagan, R. P. and N. F. Fairweather. 2011. *Clostridium difficile* has two parallel and essential Sec secretion systems. J.Biol.Chem. 286:27483-27493.
36. Sebaihia, M., B. W. Wren, P. Mullany, N. F. Fairweather, N. Minton, R. Stabler, N. R. Thomson, A. P. Roberts, A. M. Cerdeno-Tarraga, H. Wang, M. T. Holden, A. Wright, C. Churcher, M. A. Quail, S. Baker, N. Bason, K. Brooks, T. Chillingworth, A. Cronin, P. Davis, L. Dowd, A. Fraser, T. Feltwell, Z. Hance, S. Holroyd, K. Jagels, S. Moule, K. Mungall, C. Price, E. Rabbinowitsch, S. Sharp, M. Simmonds, K. Stevens, L. Unwin, S. Whithead, B. Dupuy, G. Dougan, B. Barrell, and J. Parkhill. 2006. The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome. Nat. Genet. 38:779-786.
37. Wolber, P. K., P. J. Collins, A. B. Lucas, W. A. De, and K. W. Shannon. 2006. The Agilent in situ-synthesized microarray platform. Methods Enzymol. 410:28-57.:28-57.
38. Smyth, G. K. 2004. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat.Appl.Genet.Mol.Biol. 3:Article3. Epub;2004 Feb 12..
39. Metcalf, D., S. Sharif, and J. S. Weese. 2010. Evaluation of candidate reference genes in *Clostridium difficile* for gene expression normalization. Anaerobe. 16:439-443.
40. Kuehne, S. A., S. T. Cartman, J. T. Heap, M. L. Kelly, A. Cockayne, and N. P. Minton. 2010. The role of toxin A and toxin B in *Clostridium difficile* infection. Nature 467:711-713.
41. Burns, D. A., J. T. Heap, and N. P. Minton. 2010. SleC is essential for germination of *Clostridium difficile* spores in nutrient-rich medium supplemented with the bile salt taurocholate. J.Bacteriol. 192:657-664.
42. Cerquetti, M., A. Serafino, A. Sebastianelli, and P. Mastrantonio. 2002. Binding of *Clostridium difficile* to Caco-2 epithelial cell line and to extracellular matrix proteins. FEMS Immunol.Med.Microbiol. 32:211-218.
43. Ponting, C. P., J. Schultz, F. Milpetz, and P. Bork. 1999. SMART: identification and annotation of domains from signalling and extracellular protein sequences. Nucleic Acids Res. 27:229-232.
44. Schultz, J., F. Milpetz, P. Bork, and C. P. Ponting. 1998. SMART, a simple modular architecture research tool: identification of signaling domains. Proc.Natl.Acad.Sci.U.S.A. 95:5857-5864.
45. Rawlings, N. D., A. J. Barrett, and A. Bateman. 2012. MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Res. 40:D343-D350.
46. Petersen, T. N., S. Brunak, H. G. von, and H. Nielsen. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat.Methods. 8:785-786.

47. Roy, A., A. Kucukural, and Y. Zhang. 2010. I-TASSER: a unified platform for automated protein structure and function prediction. Nat.Protoc. 5:725-738.
48. Kelley, L. A. and M. J. Sternberg. 2009. Protein structure prediction on the Web: a case study using the Phyre server. Nat.Protoc. 4:363-371.
49. Meltzer, M., S. Hasenbein, N. Mamant, M. Merdanovic, S. Poepsel, P. Hauske, M. Kaiser, R. Huber, T. Krojer, T. Clausen, and M. Ehrmann. 2009. Structure, function and regulation of the conserved serine proteases DegP and DegS of Escherichia coli. Res.Microbiol. 160:660-666.
50. Kim, D. Y. and K. K. Kim. 2005. Structure and function of HtrA family proteins, the key players in protein quality control. J.Biochem.Mol.Biol. 38:266-274.
51. Sawa, J., H. Malet, T. Krojer, F. Canellas, M. Ehrmann, and T. Clausen. 2011. Molecular adaptation of the DegQ protease to exert protein quality control in the bacterial cell envelope. J.Biol.Chem. 286:30680-30690.
52. Wilson, R. L., L. L. Brown, D. Kirkwood-Watts, T. K. Warren, S. A. Lund, D. S. King, K. F. Jones, and D. E. Hruby. 2006. Listeria monocytogenes 10403S HtrA is necessary for resistance to cellular stress and virulence. Infect.Immun. 74:765-768.
53. Chatfield, S. N., K. Strahan, D. Pickard, I. G. Charles, C. E. Hormaeche, and G. Dougan. 1992. Evaluation of Salmonella typhimurium strains harbouring defined mutations in htrA and aroA in the murine salmonellosis model. Microb.Pathog. 12:145-151.
54. Buckley, A. M., J. Spencer, D. Candlish, J. J. Irvine, and G. R. Douce. 2011. Infection of hamsters with the UK *Clostridium difficile* ribotype 027 outbreak strain R20291. J.Med. Microbiol. 60:1174-1180.
55. Goulding, D., H. Thompson, J. Emerson, N. F. Fairweather, G. Dougan, and G. R. Douce. 2009. Distinctive profiles of infection and pathology in hamsters infected with *Clostridium difficile* strains 630 and B1. Infect.Immun. 77:5478-5485.
56. Barrett, T., D. B. Troup, S. E. Wilhite, P. Ledoux, D. Rudnev, C. Evangelista, I. F. Kim, A. Soboleva, M. Tomashevsky, K. A. Marshall, K. H. Phillippy, P. M. Sherman, R. N. Muertter, and R. Edgar. 2009. NCBI GEO: archive for high-throughput functional genomic data. Nucleic Acids Res. 37:D885-D890.
57. Alsaker, K. V. and E. T. Papoutsakis. 2005. Transcriptional program of early sporulation and stationary-phase events in *Clostridium* acetobutylicum. J.Bacteriol. 187:7103-7118.
58. Errington, J. 2003. Regulation of endospore formation in Bacillus subtilis. Nat.Rev. Microbiol. 1:117-126.
59. Paredes, C. J., K. V. Alsaker, and E. T. Papoutsakis. 2005. A comparative genomic view of clostridial sporulation and physiology. Nat.Rev.Microbiol. 3:969-978.
60. Baban, S. T., S. A. Kuehne, A. Barketi-Klai, S. T. Cartman, M. L. Kelly, K. R. Hardie, I. Kansau, A. Collignon, and N. P. Minton. 2013. The Role of Flagella in *Clostridium difficile* Pathogenesis: Comparison between a Non-Epidemic and an Epidemic Strain. PLoS. One. 8:e73026.
61. Deakin, L. J., S. Clare, R. P. Fagan, L. F. Dawson, D. J. Pickard, M. R. West, B. W. Wren, N. F. Fairweather, G. Dougan, and T. D. Lawley. 2012. The *Clostridium difficile* spo0A gene is a persistence and transmission factor. Infect.Immun. 80:2704-2711.
62. Ethapa, T., R. Leuzzi, Y. K. Ng, S. T. Baban, R. Adamo, S. A. Kuehne, M. Scarselli, N. P. Minton, D. Serruto, and M. Unnikrishnan. 2013. Multiple factors modulate biofilm formation by the anaerobic pathogen *Clostridium difficile*. J.Bacteriol. 195:545-555.

63. Deneve, C., C. Janoir, I. Poilane, C. Fantinato, and A. Collignon. 2009. New trends in *Clostridium difficile* virulence and pathogenesis. Int.J.Antimicrob.Agents 33 Suppl 1:S24-S28.
64. Biswas, S. and I. Biswas. 2005. Role of HtrA in surface protein expression and biofilm formation by Streptococcus mutans. Infect.Immun. 73:6923-6934.
65. Mani, N., D. Lyras, L. Barroso, P. Howarth, T. Wilkins, J. I. Rood, A. L. Sonenshein, and B. Dupuy. 2002. Environmental response and autoregulation of *Clostridium difficile* TxeR, a sigma factor for toxin gene expression. J.Bacteriol. 184:5971-5978.
66. Mani, N. and B. Dupuy. 2001. Regulation of toxin synthesis in *Clostridium difficile* by an alternative RNA polymerase sigma factor. Proc.Natl.Acad.Sci.U.S.A 98:5844-5849.
67. Schobel, S., S. Zellmeier, W. Schumann, and T. Wiegert. 2004. The Bacillus subtilis sigmaW anti-sigma factor RsiW is degraded by intramembrane proteolysis through YluC. Mol.Microbiol. 52:1091-1105.
68. Chen, G. and X. Zhang. 2010. New insights into S2P signaling cascades: regulation, variation, and conservation. Protein Sci. 19:2015-2030.
69. Ellermeier, C. D. and R. Losick. 2006. Evidence for a novel protease governing regulated intramembrane proteolysis and resistance to antimicrobial peptides in Bacillus subtilis. Genes Dev. 20:1911-1922.
70. Ho, T. D. and C. D. Ellermeier. 2011. PrsW is required for colonization, resistance to antimicrobial peptides, and expression of extracytoplasmic function sigma factors in *Clostridium difficile*. Infect.Immun. 79:3229-3238.
71. Rigoulay, C., J. M. Entenza, D. Halpern, E. Widmer, P. Moreillon, I. Poquet, and A. Gruss. 2005. Comparative analysis of the roles of HtrA-like surface proteases in two virulent Staphylococcus aureus strains. Infect.Immun. 73:563-572.
72. Phillips, R. W. and R. M. Roop. 2001. Brucella abortus HtrA functions as an authentic stress response protease but is not required for wild-type virulence in BALB/c mice. Infect.Immun. 69:5911-5913.
73. Yuan, L., P. H. Rodrigues, M. Belanger, W. A. Dunn, Jr., and A. Progulske-Fox. 2008. Porphyromonas gingivalis htrA is involved in cellular invasion and in vivo survival. Microbiology. 154:1161-1169.
74. Lyon, W. R. and M. G. Caparon. 2004. Role for serine protease HtrA (DegP) of Streptococcus pyogenes in the biogenesis of virulence factors SpeB and the hemolysin streptolysin S. Infect.Immun. 72:1618-1625.
75. Farn, J. and M. Roberts. 2004. Effect of inactivation of the HtrA-like serine protease DegQ on the virulence of Salmonella enterica serovar Typhimurium in mice. Infect. Immun. 72:7357-7359.
76. Sebert, M. E., K. P. Patel, M. Plotnick, and J. N. Weiser. 2005. Pneumococcal HtrA protease mediates inhibition of competence by the CiaRH two-component signaling system. J.Bacteriol. 187:3969-3979.
77. Tsui, H. C., S. K. Keen, L. T. Sham, K. J. Wayne, and M. E. Winkler. 2011. Dynamic distribution of the SecA and SecY translocase subunits and septal localization of the HtrA surface chaperone/protease during Streptococcus pneumoniae D39 cell division. MBio. 2:e00202-e00211.
78. Mackin, K. E., G. P. Carter, P. Howarth, J. I. Rood, and D. Lyras. 2013. Spo0A differentially regulates toxin production in evolutionarily diverse strains of *Clostridium difficile*. PLoS.One. 8:e79666.

79. McKee, R. W., M. R. Mangalea, E. B. Purcell, E. K. Borchardt, and R. Tamayo. 2013. The second messenger c-di-GMP regulates *Clostridium difficile* toxin production by controlling expression of sigD. J.Bacteriol.
80. Aubry, A., G. Hussack, W. Chen, R. KuoLee, S. M. Twine, K. M. Fulton, S. Foote, C. D. Carrillo, J. Tanha, and S. M. Logan. 2012. Modulation of toxin production by the flagellar regulon in *Clostridium difficile*. Infect.Immun. 80:3521-3532.
81. Dupuy, B. and A. L. Sonenshein. 1998. Regulated transcription of *Clostridium difficile* toxin genes. Mol.Microbiol. 27:107-120.
82. Karlsson, S., A. Lindberg, E. Norin, L. G. Burman, and T. Akerlund. 2000. Toxins, butyric acid, and other short-chain fatty acids are coordinately expressed and down-regulated by cysteine in *Clostridium difficile*. Infect.Immun. 68:5881-5888.

9

General discussion

General discussion

Clostridium difficile can colonize the intestines of humans and animals. Up to 3% of healthy individuals and almost 40% of hospitalized patients are colonized by toxinogenic *C. difficile* strains (13,76). Antibiotic treatment can result in overgrowth of the commensal flora by *C. difficile,* thereby causing disease (13,64,72). Clinical symptoms of *Clostridium difficile* infection (CDI) can range from mild diarrhoea to life-threating pseudomembranous colitis (3,4). CDI is currently the leading cause of infectious nosocomial diarrhoea (3,62). *C. difficile* is also found in the community, in animals and in meat, which expands the possibilities for acquiring CDI outside healthcare facilities (23,33,35,36,77).

The increase in the incidence, morbidity and mortality of CDI in the last decade is mainly associated with two virulent PCR Ribotypes (RT), i.e. RT 027 and RT 078 (35,51,58,59,73) and is related to more debilitating and excessive diarrhoea, higher relapse rates and a higher chance of complications such as pseudomembranous colitis and colectomy (7,24,35,51,59). Based on incidence rates and management of CDI the financial burden is estimated at approximately $8000 million for the USA and €3000 million in Europe (7,24). As especially the elderly (≥65 years of age) are at risk for getting CDI and the proportion of elderly is expected to increase (75.3 million in 2004 to 134.5 million in 2050 in the European population), it is expected that the costs related to CDI will rise even further (http://epp.eurostat.ec.europa.eu).

To reduce the medical and financial burden of CDI it is of utmost importance to find better options to curtail the spread and treat the disease. This objective can be attained by intervening at multiple levels. First, rapid and accurate diagnosis of CDI patients and molecular typing are essential tools to prevent, monitor and control infections and outbreaks. Secondly, studying and understanding the mechanisms of *C. difficile* fitness and colonization may lead to insights to new therapeutic options aimed at preventing colonization. Finally, understanding the mechanisms of toxin regulation may result in the development of effective medical counter measures to reduce the levels of produced toxin and the clinical symptoms.

Molecular methods for diagnosis and typing *C. difficile*

Measurements to reduce the medical and financial burden on the healthcare system include awareness of CDI by the hospital staff, hygiene precautions, early diagnosis and surveillance of CDI (74). The pillars for diagnosis of CDI are still the typical clinical features in combination with laboratory assays, such as enzyme immuno assays (EIAs) (17). However, the currently used laboratory assays are time-consuming or lack optimal sensitivities and specificities for accurate and rapid

diagnosis of CDI (17,78). The lack of accurate (stand-alone) assays has persuaded clinicians to introduce two-or three step diagnostic algorithms, which results in increasing costs and time for diagnosis of CDI.

Molecular assays are potentially rapid (26,65) and have been shown to perform better than the traditionally used assays (17,20,26,51,60,65). In **Chapter 2** we describe the evaluation of three in-house developed molecular assays for the diagnosis of CDI, in comparison with the labour-intensive gold standards, i.e. the cytotoxicity assay and the toxigenic culture. Due to the high prevalence of CDI (20%) in the sample collection, the evaluated molecular assays nearly fulfilled the criteria defined for sensitivity (≥ 96.0%) but not for the specificity (≤ 90.7%) to be applicable as a stand-alone test (60). The calculated positive predicted values decreased substantially during a more realistic prevalence of CDI in a clinical setting (10%) and in a community setting (2%). Therefore, we concluded that the evaluated molecular assays are not applicable as a stand-alone test for diagnosis of CDI. However, the reported high negative predicted values are valuable in the detection of CDI cases both in an endemic and outbreak situation. Thus, we concluded that the molecular based assays are applicable as a first screening assay in a two-step screening algorithm for diagnosis of CDI.

A major drawback of the gold standards is the long turn-around time (≥ 48 hrs). Moreover, improper transport and storage conditions for long-time storage of faecal samples could lead to degradation of the toxins and to decreased culturability of *Clostridium difficile* (2,34). Together, these drawbacks could lead to false negative results that could blur the performance of the highly sensitive molecular based assays when comparing to the gold standards. The major advantages of molecular based assays compared to the gold standard are the short turnaround time and independence of phenotypic detection of toxins and culturability of *C. difficile*. Therefore, with the development of the molecular tests the gold standard for CDI diagnosis may be reconsidered.

A recently published study demonstrated that only toxin detection using a cytotoxicity assay positively correlates with clinical outcome of the disease and not the presence of *C. difficile* per se (61). Therefore, the molecular tests cannot be applied in a stand-alone test. Nevertheless, asymptomatic carriership could be an important source of transmission within hospitals (28,30,61). These observations underline the necessity and the importance to distinguish both *C. difficile* carriers and CDI patients. In order to make this distinction, we will still need to use a two-step algorithm for diagnosis of CDI.

In a two-step algorithm, the first assay should be a rapid and accurate assay capable of detecting both asymptomatic carriers and CDI patients. Ideally, this assay should be a multiplex molecular based assay which is able to identify *C. difficile*

and detect both toxin genes directly in faecal samples. This should be followed by an assay which detects produced toxins, to distinguish between CDI patients and asymptomatic carriers of *C. difficile*. To date, several assays are described which are able to detect toxins, such as the cytotoxicity assay and EIAs (17). At present, the cytotoxicity assay has the overall best performance. Preferably the second assay should also have a short turnaround time. Therefore, the applicability of new promising highly sensitive techniques (such as mass spectrometry in combination with micro-extraction methods (66) to detect the presence of *C. difficile* toxins should be further explored. Until that time, it is recommended to use a two-step algorithm to distinguish between CDI patients and asymptomatic carriers of *C. difficile*. The first screening assay should be a molecular based (multiplex) assay, followed by an EIA. The great benefit of this two-step algorithm over the commonly used two-step algorithms is the more rapid diagnosis of CDI with comparable accuracy.

Another essential tool for the control of CDI is to monitor outbreaks within healthcare facilities with molecular (sub-)typing methods (47). Multi Locus Variable number of tandem repeat Analysis (MLVA) is currently one of the most discriminatory and reproducible typing methods to study outbreaks and identify transmission routes of *C. difficile* (29,46). In **Chapter 3** we describe the application of an optimized MLVA for the virulent RT 078 on a collection of *C. difficile* strains from human and porcine origin. Despite the lack of a direct epidemiological link, we conclude that there is a high relatedness between human and porcine strains. This high relatedness could be an indication of a common source or transmission. Additional evidence for this suggested common source or transmission is the shared tetracycline resistance determinant, encoded on Tn*916* between the high related human and porcine strains. Tetracycline is frequently used in the veterinary industry but not in the (Dutch) healthcare system (Nethmap 2013, 240 http://www.swab.nl (56)). This suggests that tetracycline resistant *C. difficile* potentially arises in the veterinary industry and is being passed on into the human population. Although we can not exclude horizontal transfer of Tn*916* due to its wide distribution in Gram-positive bacteria (63), frequent use of tetracycline in the veterinary industry has also been implicated in the wide dissemination of *Staphylococcus aureus* Sequence Type 398 and enterococci in humans and pigs (1,19,79,80). For *C. difficile*, our observations on tetracyclin resistance in combination with a 5 times higher carriage rate of farmers compared to the community (45) and the rise of RT 078 in the (Dutch) community (42) indicate that farm animals may represent an important reservoir for CDI. Further research combining human and veterinary medicine is needed to prove interspecies transmission of CDI.

Automated analysis, high reproducibility, discriminatory power and the ability to easily transfer data world-wide are the main advantages of MLVA over other

sub-typing methods, such as REA and PFGE (46,47). However, in **Chapter 3** we also expose some shortcomings of MLVA as sub-typing method. We found that the PCRs for the loci $B7_{Cd}$, $C6_{Cd}$ and $G8_{Cd}$ had to be optimized due to the lack of specific PCR products and that locus $A6_{Cd}$ was absent in RT 078. This, together with the observation of invariability of loci $H9_{Cd}$ and $F3_{Cd}$ in RT 027 (9,71) and the low variability of loci $H9_{Cd}$ and $F3_{Cd}$ in RT 078, demonstrates the potential loss of discriminative power of the MLVA to distinguish between strains of certain RTs. Epidemiological links may be additionally obscured by the potential presence of multiple MLVA types within one faecal sample (67). This stresses that better and more accurate typing methods must be explored to prevent the spread and transmission of CDI.

Whole genome sequencing (WGS) as typing method can address all the above mentioned weaknesses of MLVA as a (sub)typing method. The introduction of automated analysis pipelines has opened the way for WGS as a high throughput typing method in the future (47). It has also been shown that in a single sequence run WGS can reliably detect mixed infections (27), which are estimated to occur in 7-13% of the CDI cases (31,70). In contrast, most current typing methods rely on a single colony for typing and thereby potentially miss or misclassify transmission routes. We expect that WGS will become the preferred typing method in the future, as high quality WGS data can also be used to extract additional genomic information, such as Single Nucleotide Polymorphisms (SNPs), Multi Locus Sequence Type (MLST), and sequence variants of virulence genes (37,47), and with third generation sequencing (Pacific Biosystems) potentially even PCR Ribotypes.

Non-toxin virulence factors of *C. difficile*

Preventing colonization is one of the best ways to intervene in the spread and transmission of *C. difficile*. Spores are the infectious vehicle of CDI (72) and the ingestion of spores is the first step in colonization of a host by *C. difficile*. Subsequently, spores can germinate, and the newly formed vegetative cells can outgrow and adhere to colon epithelial cells. During this process successful colonization depends on the ability of *C. difficile* to cope with environmental and host-induced stresses, such as antibiotics and reactive oxygen species. Importantly, this is frequently influenced by the presence of mobile elements that contain resistance genes or make novel metabolic pathways possible.

The endospores formed by *Clostridium spp* and Bacillus *spp* are produced to survive hostile or unfavourable environments and sporulation of *C. difficile* is recognized as a persistence and transmission factor for CDI (21,64). The sporulation pathway is largely conserved between *Clostridium spp* and Bacillus *subtilis* (57). Both in *Clostridia* and *Bacilli*, spore formation is strictly dependent on Spo0A, the master regulator

of sporulation (32,40,57). Phosphorylation of Spo0A, at least in *Bacilli*, induces binding of Spo0A to target DNA (32,57), which leads to activation or repression of downstream genes. In **Chapter 7** we describe that *C. difficile* Spo0A binds to similar conserved DNA sequences as *B. subtilis* Spo0A. Although we demonstrate that Spo0A binds to the promoter region of the Toxin B gene, we do not find reduced toxin levels in an isogenic Spo0A mutant as previously published (69). Our data, however, are in agreement with another report where no positive effect of Spo0A on toxin production was found (21). The apparent discrepancy between the data may in part be explained by differences in experimental conditions such as medium and the methods used for toxin quantification (21,69). Recently, another layer of complexity has been added to the function of Spo0A. *In vivo* toxin production may be regulated differently by Spo0A in phylogenetically distinct strain types (52). These differences clearly demonstrate that *in vitro* properties need to be corroborated in an *in vivo* situation and should be interpreted with the strain background in mind. To address the differences between the various studies on Spo0A in *C. difficile*, future work should include a comprehensive *in vivo* analysis of Spo0A binding (Chromatin immunoprecipitation-sequencing) combined with an extensive transcriptome (RNA-Sequencing) analysis, to determine the effect on the transcription of genes in various strains. These analyses may shed light on the role of Spo0A in pathogenesis and virulence in the various phylogenetically distinct *C. difficile* strains.

Adhesion to the mucosal surfaces is an essential step in the successful colonization of the host by *C. difficile* (72). Relatively little is known about adhesion of *C. difficile* to the mucosal surfaces in the host, compared to the overwhelming information on toxins and sporulation and their role in pathogenesis. To date, a few cell surface proteins have been described to play a role in adhesion and colonization, such as SlpA (11) and flagellar proteins (68). Denève and colleagues hypothesized that gene expression of adhesion and colonization factors is enhanced under stress conditions (22).

In **Chapter 8** we knocked-out the gene encoding an HtrA-like protease which we expected to be involved in adaptation and response to stress. The micro-array analysis of the *htrA* mutant showed decreased gene expression of well-known characterized adhesion and colonization factors, such as SlpA and flagellar proteins (11,68). Knocking-out another hypothesized stress-related protease also resulted in a reduced expression of some adhesion and colonization factors, but not of the *slpA* and flagellar proteins (**our unpublished results**). *In vitro* characterization of the two protease mutants showed decreased adhesion to Caco-2 cells compared to the wild-type strain (**Chapter 8 and our unpublished results**). These observations indicate that both proteases are involved in the regulation of adhesion and

colonization factors. The difference between the two protease mutant strains in the gene expression patterns is suggestive for the existence of various specific stress related pathways in *C. difficile*. Further research is needed to unravel the various stress-related pathways and their role in colonization and development of infection in the host.

Besides the decreased *in vitro* sporulation and adhesion properties, expected to result in decreased virulence, the *C. difficile htrA* mutant also showed increased toxin production (**Chapter 8**), which could result in enhanced virulence. Indeed, we demonstrate enhanced virulence of the *htrA* mutant in a hamster model, likely as a result of the increased TcdA levels in combination with exquisite sensitivity of the hamsters for toxin production. This raises the question which of the phenotypes (i.e. decreased sporulation, decreased adhesion or increased toxin A production) is more relevant in case of a human infection. Our study on the *in vivo* effects of HtrA is limited by the choice of animal model for the following reasons. First, to get reproducible infections in hamsters, the use of relatively large inocula is necessary (10). This may mask subtle effects as a result of differences in stress resistance. Second, the hamster model is not suitable to study persistence and transmission of *C. difficile*, due to the acute nature of the disease in hamsters. To overcome these limitations, several strategies can be employed.

HtrA is involved in different stress responses in various organisms (5,16,53). To determine if the *C. difficile htrA* mutant strains are more sensitive for environmental stresses, stress factors (such as heat, hydrogen peroxide, salt) can be added to bacterial cultures of wild type and the *htrA* mutant strains, while growth kinetics and/or the number of viable cells are compared. These experiments can also be performed in a mixed culture of *htrA* mutant and wild type strains to determine the relative fitness of the mutant under stress conditions. An *in vivo* competition experiment, even in the hamster model, may reveal subtle differences between wild type and *htrA* mutant strains that go undetected during a single strain infection. To determine if the *in vitro* decreased sporulation and adhesion properties observed *in vitro* play an important role in persistence and transmission, these properties should be investigated in an *in vivo* animal model which is less sensitive to the produced toxins, such as the mouse (21). We expect that an *htrA* knockout in such a model will show a decrease in virulence similar to other pathogenic bacteria (15,16,64).

HtrA-like proteases of *E. coli* are the best-characterized members of the HtrA protease family and they have been shown to possess a dual function as a re-foldase and protease (53,55). Our studies do not address whether the re-folding and/or degradation function of HtrA is required for the observed effect (53,55). It would be interesting to test the importance of both functions by complementing the *htrA* mutant with a proteolytically inactive HtrA protein, rather than wild type, followed

by a transcriptome analysis. Variances in gene expression profiles between both mutants could give insight in which function of HtrA is of importance in which response.

In many micro-organisms, antibiotic resistance can be acquired via incorporation of mobile elements (8). Incorporation of antibiotic resistance genes could result in a growth advantage during colonization and the ability to outcompete the commensal flora after antibiotic treatment. Besides antibiotic resistance determinants, mobile elements also frequently carry other virulence determinants. Therefore, the spread of mobile elements greatly impacts the fitness of *C. difficile*. Moreover, analysis of the presence of mobile elements could also provide valuable information for identification of possible transmission routes. In **Chapter 3 and 4** we describe the presence of conjugative transposons, encoding antibiotic resistance markers, in the emerging RT 078 strains. In **Chapter 3** we describe that the high relatedness of the RT078 strains and the shared tetracycline resistance, encoded by Tn*916*, indicates a possible common source or transmission, as discussed above. The second mobile element, Tn*6164*, encodes besides several antibiotic resistance genes also other (putative) virulence genes of various origins (**Chapter 4)**. The increased mortality rates of patients infected with *C. difficile*::Tn*6164* (29% vs 3%) suggests that the presence of this element may render RT 078 strain more virulent. The observed trend could be further explored by comparing isogenic RT 078 and RT 078::Tn*6164* strains in an *in vivo* model. Additional epidemiological research is also needed to confirm the possible link between the presence of Tn*6164* and increased virulence.

Our work shows that mobile elements can be used as an indication of a possible common source (**Chapter 3** and C. W. Knetsch personal communication) and may be linked to virulence of *C. difficile* (**Chapter 4**). Together, these observations underline the importance to detect the presence of mobile elements. However, the presence of mobile elements is not recognized by PCR Ribotyping, which is currently the most frequently used method (47). Due to their unstable nature and variable location, mobile elements are also not recognized by the WGS SNP typing method which focuses on the core, rather than the accessory, genome. However, WGS can be a powerful tool for recognition of antibiotic resistance and potential virulence genes encoded on mobile elements (**Chapter 4** and C. W. Knetsch personal communication). Therefore, we argue that WGS should not only focus on the core genome, but also should take the accessory genome into account. Applying WGS as the typing method in a good surveillance system will result in an increased understanding of the role mobile elements could play in the spread of *C. difficile* and their importance for virulence.

Toxin regulation

Toxin A and Toxin B are the two major virulence factors of *C. difficile* and primarily responsible for causing the symptoms of CDI (48,49,75). Understanding the mechanisms behind regulation of toxin production is necessary for the development of effective medical countermeasures to moderate the clinical course and outcome of disease. Besides both toxin genes, the PaLoc also contains the gene encoding the proposed negative regulator (*tcdC*) of toxin gene transcription (44,75). Both hypervirulent RTs 027 and 078 contain point-mutations in *tcdC*, which has been proposed as a possible explanation for their increased virulence (18,35,76). In **Chapters 5 and 6**, we investigated toxin regulatory capacities and biochemical properties of TcdC.

In **Chapter 5** we evaluated the role of TcdC in toxin production by comparing the transcription levels of the toxin genes and toxin expression levels of an isogenic *tcdC* mutant and the parental strain (wild type). We concluded that TcdC is not a major negative regulator of toxin expression. Our work and that of an independent study by Cartman *et al* (14) contradicts previous reports that demonstrate that TcdC is a major repressor for toxin production (12,54). We cannot exclude that TcdC has a more profound effect in another strain on toxin production under other conditions than ours. Preliminary data have shown that overexpression of TcdC in *C. difficile* can result in decreased toxin production (**our unpublished observations**). This is in line with a report of decreased toxin production upon expression of TcdC from a multi-copy plasmid (12). As of yet unidentified conditions leading to the overexpression of TcdC could therefore also reconcile our findings with the *in vitro* property of TcdC as a repressor of toxin production (54).

In **Chapter 6** we demonstrate a so far uncharacterized function for TcdC. Bioinformatic analysis of the TcdC sequence predicted that the C-terminal domain of TcdC is an Oligonucleotide Binding (OB)-fold which is able to bind to DNA. Biochemical analyses showed that TcdC binds to G3-quadruplex structures (QS). It has been shown that QS can influence several *in vivo* processes, like transcription (6). Though the C-terminal domain was identified as important for toxin repression (25,54), the role of TcdC binding to QS in toxin regulation is unclear since no QS are predicted to be formed in the PaLoc region. Two interesting experiments could shed more light on the function of TcdC and the relevance of QS structure binding by this protein. First, TcdC *in vitro* binding to DNA-QS *in vitro* might mimic binding to other (unknown) secondary structures *in vivo*. Chromatin immunoprecipitation experiments could reveal the DNA sequences that are bound by TcdC *in vivo,* thereby providing a better understanding of the *in vivo* function of TcdC. Secondly, to clarify the role of the DNA binding domain, knowledge about the topology of TcdC is essential. The predictions of the location of the TcdC DNA binding

domain are ambiguous. Topology experiments should reveal the subcellular location of the DNA binding domain. If the DNA binding domain is localized extracellularly, it is not likely that TcdC plays a role in the *in vivo* transcriptional regulation of toxin production. Together these experiments may provide a better understanding of the function of TcdC.

Genetic tools

Over the last years major advances have been made in the genetics and molecular biology of *C. difficile*. The introduction of (replicative) plasmids by conjugation and the Group II intron based mutagenesis system (ClosTron) has spurred the genetic characterization and initial understanding of the virulence of *C. difficile* (12,39,41,43). However, the use of these 1st generation genetic tools has several limitations. First of all, phenotypic complementation of knocked-out genes using multi-copy (replicative) plasmids is not ideal. It is unclear how the protein expression level in the complemented strain relates to the physiological level in the wild type strain. This may result in over-expression of the introduced gene and thereby over-interpretation of the effect (12). Secondly, the use of ClosTron to generate mutants may lead to polar effects on downstream genes (50). Another disadvantage of the ClosTron system is the impossibility to introduce (point) mutations and reporter genes into the *C. difficile* genome.

The development of the 2nd generation genetic manipulation tools, the Allelic-Coupled Exchange (ACE) systems, is a major advance in studying *C. difficile* virulence genes (14,38). First of all, the use of ACE does not necessarily result in polar effects on non-inactivated genes and because of the fact that genes can be introduced in single copy on the chromosome, protein levels in the complemented strain are more likely to occur at physiologically relevant levels. Furthermore, the ACE system is capable of introducing and/or restoring possibly important (point) mutations in genes which are linked to increased virulence (14,18). Finally, the ACE systems do not rely on the introduction of antibiotic selection markers, allowing a fair comparison of mutant versus wild type in *in vivo* studies.

The use of different strains has introduced another layer of complexity in studying the influences of genes on virulence. Mackin and colleagues demonstrated that *in vivo* toxin production may be regulated differently by Spo0A in phylogenetically distinct strain types (52). Ideally, there should be a wide agreement which strain(s) and genetic tools should be used to study *C. difficile* virulence factors. As this is unlikely to happen due to the important differences between phylogenetically distinct *C. difficile* strains, each manuscript should mention that the presented data is representative only for the used strain and conditions.

Concluding remarks

Over the last years major advances have been made in the field of *C. difficile* research. Despite the continuous progress of research in *C. difficile* epidemiology and molecular biology, the medical and financial burden on the healthcare system is still a serious problem. This thesis shows that the development of molecular-based techniques in the detection and typing of *C. difficile* could be very valuable in combating CDI. This also applies for the genetics and molecular biology of *C. difficile* in understanding the virulence of *C. difficile*. However, the continuing development of new techniques, both typing and genetic manipulation tools, has also led to the recognition of the potential weaknesses of the currently frequent used techniques. Therefore, it is of utmost importance to keep up with the technical progress in the field of *C. difficile* research to enable further elucidation of the mechanisms that determine the virulence of *C. difficile*.

Reference List

1. Agerso, Y., A. G. Pedersen, and F. M. Aarestrup. 2006. Identification of Tn*5397*-like and Tn*916*-like transposons and diversity of the tetracycline resistance gene *tet*(M) in enterococci from humans, pigs and poultry. J.Antimicrob.Chemother. 57:832-839.
2. Arroyo, L. G., J. Rousseau, B. M. Willey, D. E. Low, H. Staempfli, A. McGeer, and J. S. Weese. 2005. Use of a selective enrichment broth to recover *Clostridium difficile* from stool swabs stored under different conditions. J.Clin.Microbiol. 43:5341-5343.
3. Bartlett, J. G. 2002. Clinical practice. Antibiotic-associated diarrhea. N.Engl.J.Med. 346:334-339.
4. Bartlett, J. G. and D. N. Gerding. 2008. Clinical recognition and diagnosis of *Clostridium difficile* infection. Clin.Infect.Dis. 46 Suppl 1:S12-8. doi: 10.1086/521863.:S12-S18.
5. Biswas, S. and I. Biswas. 2005. Role of HtrA in surface protein expression and biofilm formation by Streptococcus mutans. Infect.Immun. 73:6923-6934.
6. Bochman, M. L., K. Paeschke, and V. A. Zakian. 2012. DNA secondary structures: stability and function of G-quadruplex structures. Nat.Rev.Genet. 13:770-780.
7. Bouza, E. 2012. Consequences of *Clostridium difficile* infection: understanding the healthcare burden. Clin.Microbiol.Infect. 18 Suppl 6:5-12. doi: 10.1111/1469-0691.12064.:5-12.
8. Broaders, E., C. G. Gahan, and J. R. Marchesi. 2013. Mobile genetic elements of the human gastrointestinal tract: potential for spread of antibiotic resistance genes. Gut Microbes. 4:271-280.
9. Broukhanski, G., A. Simor, and D. R. Pillai. 2011. Defining criteria to interpret multilocus variable-number tandem repeat analysis to aid *Clostridium difficile* outbreak investigation. J.Med.Microbiol. 60:1095-1100.
10. Buckley, A. M., J. Spencer, D. Candlish, J. J. Irvine, and G. R. Douce. 2011. Infection of hamsters with the UK *Clostridium difficile* Ribotype 027 outbreak strain R20291. J.Med. Microbiol. 60:1174-1180.
11. Calabi, E., F. Calabi, A. D. Phillips, and N. F. Fairweather. 2002. Binding of *Clostridium difficile* surface layer proteins to gastrointestinal tissues. Infect.Immun. 70:5770-5778.
12. Carter, G. P., G. R. Douce, R. Govind, P. M. Howarth, K. E. Mackin, J. Spencer, A. M. Buckley, A. Antunes, D. Kotsanas, G. A. Jenkin, B. Dupuy, J. I. Rood, and D. Lyras. 2011. The Anti-Sigma Factor TcdC Modulates Hypervirulence in an Epidemic BI/NAP1/027 Clinical Isolate of *Clostridium difficile*. PLoS.Pathog. 7:e1002317.
13. Cartman, S. T., J. T. Heap, S. A. Kuehne, A. Cockayne, and N. P. Minton. 2010. The emergence of 'hypervirulence' in *Clostridium difficile*. Int.J.Med.Microbiol. 300:387-395.
14. Cartman, S. T., M. L. Kelly, D. Heeg, J. T. Heap, and N. P. Minton. 2012. Precise Manipulation of the *Clostridium difficile* Chromosome Reveals a Lack of Association between the *tcdC* Genotype and Toxin Production. Appl.Environ.Microbiol. 78:4683-4690.
15. Chitlaru, T., G. Zaide, S. Ehrlich, I. Inbar, O. Cohen, and A. Shafferman. 2011. HtrA is a major virulence determinant of Bacillus anthracis. Mol.Microbiol. 81:1542-1559.
16. Clausen, T., M. Kaiser, R. Huber, and M. Ehrmann. 2011. HTRA proteases: regulated proteolysis in protein quality control. Nat.Rev.Mol.Cell Biol. 12:152-162.

17. Crobach, M. J., O. M. Dekkers, M. H. Wilcox, and E. J. Kuijper. 2009. European Society of Clinical Microbiology and Infectious Diseases (ESCMID): data review and recommendations for diagnosing *Clostridium difficile*-infection (CDI). Clin.Microbiol. Infect. 15:1053-1066.
18. Curry, S. R., J. W. Marsh, C. A. Muto, M. M. O'Leary, A. W. Pasculle, and L. H. Harrison. 2007. *tcdC* genotypes associated with severe TcdC truncation in an epidemic clone and other strains of *Clostridium difficile*. J.Clin.Microbiol. 45:215-221.
19. de Neeling, A. J., M. J. van den Broek, E. C. Spalburg, M. G. van Santen-Verheuvel, W. D. Dam-Deisz, H. C. Boshuizen, A. W. van de Giessen, D. E. van, and X. W. Huijsdens. 2007. High prevalence of methicillin resistant Staphylococcus aureus in pigs. Vet. Microbiol. 122:366-372.
20. Deak, E., S. A. Miller, and R. M. Humphries. 2013. Comparison of the Illumigene, Simplexa, and AmpliVue *Clostridium difficile* Molecular Assays for Diagnosis of C. difficile Infection. J.Clin.Microbiol.
21. Deakin, L. J., S. Clare, R. P. Fagan, L. F. Dawson, D. J. Pickard, M. R. West, B. W. Wren, N. F. Fairweather, G. Dougan, and T. D. Lawley. 2012. The *Clostridium difficile* spo0A gene is a persistence and transmission factor. Infect.Immun. 80:2704-2711.
22. Denève, C., C. Delomenie, M. C. Barc, A. Collignon, and C. Janoir. 2008. Antibiotics involved in *Clostridium difficile*-associated disease increase colonization factor gene expression. J.Med.Microbiol. 57:732-738.
23. Dial, S., J. A. Delaney, A. N. Barkun, and S. Suissa. 2005. Use of gastric acid-suppressive agents and the risk of community-acquired *Clostridium difficile*-associated disease. JAMA. 294:2989-2995.
24. Dubberke, E. R. and M. A. Olsen. 2012. Burden of *Clostridium difficile* on the healthcare system. Clin.Infect.Dis. 55 Suppl 2:S88-92. doi: 10.1093/cid/cis335.:S88-S92.
25. Dupuy, B. and S. Matamouros. 2006. Regulation of toxin and bacteriocin synthesis in *Clostridium* species by a new subgroup of RNA polymerase sigma-factors. Res.Microbiol. 157:201-205.
26. Eastwood, K., P. Else, A. Charlett, and M. Wilcox. 2009. Comparison of nine commercially available *Clostridium difficile* toxin detection assays, a real-time PCR assay for *C. difficile tcdB*, and a glutamate dehydrogenase detection assay to cytotoxin testing and cytotoxigenic culture methods. J.Clin.Microbiol. 47:3211-3217.
27. Eyre, D. W., M. L. Cule, D. Griffiths, D. W. Crook, T. E. Peto, A. S. Walker, and D. J. Wilson. 2013. Detection of mixed infection from bacterial whole genome sequence data allows assessment of its role in *Clostridium difficile* transmission. PLoS.Comput.Biol. 9:e1003059.
28. Eyre, D. W., M. L. Cule, D. J. Wilson, D. Griffiths, A. Vaughan, L. O'Connor, C. L. Ip, T. Golubchik, E. M. Batty, J. M. Finney, D. H. Wyllie, X. Didelot, P. Piazza, R. Bowden, K. E. Dingle, R. M. Harding, D. W. Crook, M. H. Wilcox, T. E. Peto, and A. S. Walker. 2013. Diverse sources of C. difficile infection identified on whole-genome sequencing. N.Engl.J.Med. 369:1195-1205.
29. Eyre, D. W., W. N. Fawley, E. L. Best, D. Griffiths, N. E. Stoesser, D. W. Crook, T. Peto, A. S. Walker, and M. H. Wilcox. 2013. Comparison of multilocus variable number tandem repeat analysis and whole genome sequencing for investigation of *Clostridium difficile* transmission. J.Clin.Microbiol.

30. Eyre, D. W., D. Griffiths, A. Vaughan, T. Golubchik, M. Acharya, L. O'Connor, D. W. Crook, A. S. Walker, and T. E. Peto. 2013. Asymptomatic *Clostridium difficile* colonisation and onward transmission. PLoS.One. 8:e78445.
31. Eyre, D. W., A. S. Walker, D. Griffiths, M. H. Wilcox, D. H. Wyllie, K. E. Dingle, D. W. Crook, and T. E. Peto. 2012. *Clostridium difficile* mixed infection and reinfection. J.Clin. Microbiol. 50:142-144.
32. Ferrari, F. A., K. Trach, D. LeCoq, J. Spence, E. Ferrari, and J. A. Hoch. 1985. Characterization of the spo0A locus and its deduced product. Proc.Natl.Acad.Sci.U.S.A. 82:2647-2651.
33. Freeman, J., M. P. Bauer, S. D. Baines, J. Corver, W. N. Fawley, B. Goorhuis, E. J. Kuijper, and M. H. Wilcox. 2010. The changing epidemiology of *Clostridium difficile* infections. Clin.Microbiol.Rev. 23:529-549.
34. Freeman, J. and M. H. Wilcox. 2003. The effects of storage conditions on viability of *Clostridium difficile* vegetative cells and spores and toxin activity in human faeces. J.Clin. Pathol. 56:126-128.
35. Goorhuis, A., D. Bakker, J. Corver, S. B. Debast, C. Harmanus, D. W. Notermans, A. A. Bergwerff, F. W. Dekker, and E. J. Kuijper. 2008. Emergence of *Clostridium difficile* infection due to a new hypervirulent strain, polymerase chain reaction Ribotype 078. Clin.Infect.Dis. 47:1162-1170.
36. Goorhuis, A., S. B. Debast, L. A. van Leengoed, C. Harmanus, D. W. Notermans, A. A. Bergwerff, and E. J. Kuijper. 2008. *Clostridium difficile* PCR Ribotype 078: an emerging strain in humans and in pigs? J.Clin.Microbiol. 46:1157-07.
37. Gurtler, V. and D. Grando. 2013. Reprint of New opportunities for improved ribotyping of C. difficile clinical isolates by exploring their genomes. J.Microbiol.Methods.10.
38. Heap, J. T., M. Ehsaan, C. M. Cooksley, Y. K. Ng, S. T. Cartman, K. Winzer, and N. P. Minton. 2012. Integration of DNA into bacterial chromosomes from plasmids without a counter-selection marker. Nucleic Acids Res. 40:e59.
39. Heap, J. T., S. A. Kuehne, M. Ehsaan, S. T. Cartman, C. M. Cooksley, J. C. Scott, and N. P. Minton. 2010. The ClosTron: Mutagenesis in *Clostridium* refined and streamlined. J.Microbiol.Methods 80:49-55.
40. Heap, J. T., O. J. Pennington, S. T. Cartman, G. P. Carter, and N. P. Minton. 2007. The ClosTron: a universal gene knock-out system for the genus *Clostridium* . J.Microbiol. Methods 70:452-464.
41. Heap, J. T., O. J. Pennington, S. T. Cartman, and N. P. Minton. 2009. A modular system for *Clostridium* shuttle plasmids. J.Microbiol.Methods. 78:79-85.
42. Hensgens, M. P., E. C. Keessen, M. M. Squire, T. V. Riley, M. G. Koene, B. E. de, L. J. Lipman, and E. J. Kuijper. 2012. *Clostridium difficile* infection in the community: a zoonotic disease? Clin.Microbiol.Infect. 18:635-645.
43. Ho, T. D. and C. D. Ellermeier. 2011. PrsW is required for colonization, resistance to antimicrobial peptides, and expression of extracytoplasmic function sigma factors in *Clostridium difficile*. Infect.Immun. 79:3229-3238.
44. Hundsberger, T., V. Braun, M. Weidmann, P. Leukel, M. Sauerborn, and C. von Eichel-Streiber. 1997. Transcription analysis of the genes *tcdA-E* of the pathogenicity locus of *Clostridium difficile*. Eur.J.Biochem. 244:735-742.
45. Keessen, E. C., W. Gaastra, and L. J. Lipman. 2011. *Clostridium difficile* infection in humans and animals, differences and similarities. Vet.Microbiol. 153:205-217.

46. Killgore, G., A. Thompson, S. Johnson, J. Brazier, E. Kuijper, J. Pepin, E. H. Frost, P. Savelkoul, B. Nicholson, R. J. van den Berg, H. Kato, S. P. Sambol, W. Zukowski, C. Woods, B. Limbago, D. N. Gerding, and L. C. McDonald. 2008. Comparison of seven techniques for typing international epidemic strains of *Clostridium difficile*: restriction endonuclease analysis, pulsed-field gel electrophoresis, PCR-ribotyping, multilocus sequence typing, multilocus variable-number tandem-repeat analysis, amplified fragment length polymorphism, and surface layer protein A gene sequence typing. J.Clin.Microbiol. 46:431-437.
47. Knetsch, C. W., T. D. Lawley, M. P. Hensgens, J. Corver, M. W. Wilcox, and E. J. Kuijper. 2013. Current application and future perspectives of molecular typing methods to study *Clostridium difficile* infections. Euro.Surveill. 18:20381.
48. Kuehne, S. A., S. T. Cartman, J. T. Heap, M. L. Kelly, A. Cockayne, and N. P. Minton. 2010. The role of toxin A and toxin B in *Clostridium difficile* infection. Nature 467:711-713.
49. Kuehne, S. A., M. M. Collery, M. L. Kelly, S. T. Cartman, A. Cockayne, and N. P. Minton. 2013. Importance of Toxin A, Toxin B, and CDT in Virulence of an Epidemic *Clostridium difficile* Strain. J.Infect.Dis.
50. Kuehne, S. A. and N. P. Minton. 2012. ClosTron-mediated engineering of *Clostridium* . Bioengineered. 3:247-254.
51. Loo, V. G., L. Poirier, M. A. Miller, M. Oughton, M. D. Libman, S. Michaud, A. M. Bourgault, T. Nguyen, C. Frenette, M. Kelly, A. Vibien, P. Brassard, S. Fenn, K. Dewar, T. J. Hudson, R. Horn, P. Rene, Y. Monczak, and A. Dascal. 2005. A predominantly clonal multi-institutional outbreak of *Clostridium difficile*-associated diarrhea with high morbidity and mortality. N.Engl.J.Med. 353:2442-2449.
52. Mackin, K. E., G. P. Carter, P. Howarth, J. I. Rood, and D. Lyras. 2013. Spo0A Differentially Regulates Toxin Production in Evolutionarily Diverse Strains of *Clostridium difficile*. PLoS.One. 8:e79666.
53. Malet, H., F. Canellas, J. Sawa, J. Yan, K. Thalassinos, M. Ehrmann, T. Clausen, and H. R. Saibil. 2012. Newly folded substrates inside the molecular cage of the HtrA chaperone DegQ. Nat.Struct.Mol.Biol. 19:152-157.
54. Matamouros, S., P. England, and B. Dupuy. 2007. *Clostridium difficile* toxin expression is inhibited by the novel regulator TcdC. Mol.Microbiol. 64:1274-1288.
55. Meltzer, M., S. Hasenbein, N. Mamant, M. Merdanovic, S. Poepsel, P. Hauske, M. Kaiser, R. Huber, T. Krojer, T. Clausen, and M. Ehrmann. 2009. Structure, function and regulation of the conserved serine proteases DegP and DegS of Escherichia coli. Res.Microbiol. 160:660-666.
56. Mevius D.J., Wit B., and van Pelt W. Monitoring of Antimicrobial Resistance and Antibiotic Usage in Animals in the Netherlands in 2006/2007 available at: http://www.cvi.wur.nl/NR/rdonlyres/DDA15856-1179-4CAB-BAC6-28C4728ACA03/83791/MARAN_2007_def2.pdf. 2007.
57. Paredes, C. J., K. V. Alsaker, and E. T. Papoutsakis. 2005. A comparative genomic view of clostridial sporulation and physiology. Nat.Rev.Microbiol. 3:969-978.
58. Pepin, J., L. Valiquette, M. E. Alary, P. Villemure, A. Pelletier, K. Forget, K. Pepin, and D. Chouinard. 2004. *Clostridium difficile*-associated diarrhea in a region of Quebec from 1991 to 2003: a changing pattern of disease severity. CMAJ. 171:466-472.

59. Pepin, J., L. Valiquette, and B. Cossette. 2005. Mortality attributable to nosocomial *Clostridium difficile*-associated disease during an epidemic caused by a hypervirulent strain in Quebec. CMAJ. 173:1037-1042.
60. Planche, T., A. Aghaizu, R. Holliman, P. Riley, J. Poloniecki, A. Breathnach, and S. Krishna. 2008. Diagnosis of *Clostridium difficile* infection by toxin detection kits: a systematic review. Lancet Infect.Dis. 8:777-784. doi:S1473-3099(08)70233-0 (pii);10.1016/S1473-3099(08)70233-0 (doi).
61. Planche, T. D., K. A. Davies, P. G. Coen, J. M. Finney, I. M. Monahan, K. A. Morris, L. O'Connor, S. J. Oakley, C. F. Pope, M. W. Wren, N. P. Shetty, D. W. Crook, and M. H. Wilcox. 2013. Differences in outcome according to *Clostridium difficile* testing method: a prospective multicentre diagnostic validation study of C difficile infection. Lancet Infect.Dis. 13:936-945.
62. Polage, C. R., J. V. Solnick, and S. H. Cohen. 2012. Nosocomial diarrhea: evaluation and treatment of causes other than *Clostridium difficile*. Clin.Infect.Dis. 55:982-989.
63. Roberts, A. P. and P. Mullany. 2009. A modular master on the move: the Tn*916* family of mobile genetic elements. Trends Microbiol. 17:251-258.
64. Rupnik, M., M. H. Wilcox, and D. N. Gerding. 2009. *Clostridium difficile* infection: new developments in epidemiology and pathogenesis. Nat.Rev.Microbiol. 7:526-536.
65. Sloan, L. M., B. J. Duresko, D. R. Gustafson, and J. E. Rosenblatt. 2008. Comparison of real-time PCR for detection of the *tcdC* gene with four toxin immunoassays and culture in diagnosis of *Clostridium difficile* infection. J.Clin.Microbiol. 46:1996-2001.
66. Tait, E., K. A. Hill, J. D. Perry, S. P. Stanforth, and J. R. Dean. 2013. Development of a novel method for detection of *Clostridium difficile* using HS-SPME-GC-MS. J.Appl. Microbiol.10.
67. Tanner, H. E., K. J. Hardy, and P. M. Hawkey. 2010. Co-existence of multiple MLVA sub-types of *Clostridium difficile* PCR-Ribotype 027 strains within faecal specimens. J.Clin.Microbiol.
68. Tasteyre, A., M. C. Barc, A. Collignon, H. Boureau, and T. Karjalainen. 2001. Role of FliC and FliD flagellar proteins of *Clostridium difficile* in adherence and gut colonization. Infect.Immun. 69:7937-7940.
69. Underwood, S., S. Guan, V. Vijayasubhash, S. D. Baines, L. Graham, R. J. Lewis, M. H. Wilcox, and K. Stephenson. 2009. Characterization of the sporulation initiation pathway of *Clostridium difficile* and its role in toxin production. J.Bacteriol. 191:7296-7305.
70. van den Berg, R. J., H. A. Ameen, T. Furusawa, E. C. Claas, d. van, V, and E. J. Kuijper. 2005. Coexistence of multiple PCR-Ribotype strains of *Clostridium difficile* in faecal samples limits epidemiological studies. J.Med.Microbiol. 54:173-179.
71. van den Berg, R. J., I. Schaap, K. E. Templeton, C. H. Klaassen, and E. J. Kuijper. 2007. Typing and subtyping of *Clostridium difficile* isolates by using multiple-locus variable-number tandem-repeat analysis. J.Clin.Microbiol. 45:1024-1028.
72. Vedantam, G., A. Clark, M. Chu, R. McQuade, M. Mallozzi, and V. K. Viswanathan. 2012. *Clostridium difficile* infection: toxins and non-toxin virulence factors, and their contributions to disease establishment and host response. Gut Microbes. 3:121-134.
73. Viswanathan, V. K., M. J. Mallozzi, and G. Vedantam. 2010. *Clostridium difficile* infection: An overview of the disease and its pathogenesis, epidemiology and interventions. Gut Microbes. 1:234-242.

74. Vonberg, R. P., E. J. Kuijper, M. H. Wilcox, F. Barbut, P. Tull, P. Gastmeier, P. J. van den Broek, A. Colville, B. Coignard, T. Daha, S. Debast, B. I. Duerden, S. van den Hof, T. van der Kooi, H. J. Maarleveld, E. Nagy, D. W. Notermans, J. O'Driscoll, B. Patel, S. Stone, and C. Wiuff. 2008. Infection control measures to limit the spread of *Clostridium difficile*. Clin.Microbiol.Infect. 14 Suppl 5:2-20. doi: 10.1111/j.1469-0691.2008.01992.x.:2-20.
75. Voth, D. E. and J. D. Ballard. 2005. *Clostridium difficile* toxins: mechanism of action and role in disease. Clin.Microbiol.Rev. 18:247-263.
76. Warny, M., J. Pepin, A. Fang, G. Killgore, A. Thompson, J. Brazier, E. Frost, and L. C. McDonald. 2005. Toxin production by an emerging strain of *Clostridium difficile* associated with outbreaks of severe disease in North America and Europe. Lancet 366:1079-1084.
77. Weese, J. S., B. P. Avery, J. Rousseau, and R. J. Reid-Smith. 2009. Detection and enumeration of *Clostridium difficile* spores in retail beef and pork. Appl.Environ. Microbiol. 75:5009-5011.
78. Wilcox, M. H. 2012. Overcoming barriers to effective recognition and diagnosis of *Clostridium difficile* infection. Clin.Microbiol.Infect. 18 Suppl 6:13-20. doi: 10.1111/1469-0691.12057.:13-20.
79. Wulf, M. W., A. Markestein, F. T. van der Linden, A. Voss, C. Klaassen, and C. M. Verduin. 2008. First outbreak of methicillin-resistant Staphylococcus aureus ST398 in a Dutch hospital, June 2007. Euro.Surveill 13.
80. Wulf, M. W., M. Sorum, N. A. van, R. Skov, W. J. Melchers, C. H. Klaassen, and A. Voss. 2008. Prevalence of methicillin-resistant Staphylococcus aureus among veterinarians: an international study. Clin.Microbiol.Infect. 14:29-34.

Dutch summary
Nederlandse samenvatting

Dit proefschrift beslaat vele facetten van onderzoek aan *Clostridium difficile*, van toegepast tot fundamenteel. Deze Gram-positieve, sporevormende bacterie kan een darminfectie veroorzaken, ook wel *Clostridium difficile* infectie (CDI) genoemd. De symptomen van CDI worden veroorzaakt door toxinen die de darm beschadigen. CDI wordt tegenwoordig erkend als de belangrijkste oorzaak van infectieuze nosocomiale diarree. In het laatste decennium is de incidentie en ernst van CDI toegenomen. Deze toename heeft wereldwijd belangrijke financiële en medische gevolgen voor de gezondheidszorg en patiënten. Dit proefschrift kan onderverdeeld worden in 4 onderdelen: i) een algemene inleiding over het toegepaste en fundamentele onderzoek naar CDI; ii) een beschrijving van de ontwikkeling van diagnostische en typering methoden voor *Clostridium difficile*; iii) een deel over onderzoek naar de regulatie van (mogelijke) virulentie factoren; en iv) een discussie over de bevindingen van dit proefschrift.

Inleiding

Clostridium difficile werd bijna tachtig jaar geleden voor het eerst beschreven door Hall en O'Toole. De bacterie werd voor het eerst geïsoleerd uit faeces van pasgeborenen en beschreven als een zeer beweeglijke, Gram-positieve staaf met sub-terminale sporen. Gezonde individuen zijn meestal beschermd door de commensale flora in het colon. Echter, een antibiotica behandeling kan leiden tot de overgroei van de commensale flora door *C. difficile,* wat uiteindelijk kan leiden tot CDI. De klinische symptomen van CDI kunnen variëren van diarree tot aan ernstige colitis. Naast het gebruik van antibiotica zijn gevorderde leeftijd (>60 jaar), onderliggende ziekten en infectiedruk andere risicofactoren voor CDI.

In het laatste decennium is de epidemiologie van CDI radicaal veranderd. Sinds 2002 worden grote uitbraken in Noord-Amerika en Europa van CDI gerapporteerd die gekarakteriseerd worden door een ernstiger verloop van de ziekte en verhoogde sterftecijfers. Deze toename heeft grote implicaties voor de patiënten en de gezondheidszorg, en heeft daarnaast ook financiële gevolgen voor het zorgstelsel. Op basis van de huidige incidentie cijfers en de verwachte vergrijzing zal deze financiële last in de toekomst alleen maar verder toenemen.

Om de bovenstaande gevolgen van CDI te verminderen is het van het grootste belang de verspreiding van CDI te beperken en nieuwe behandelopties voor de ziekte te onderzoeken. Deze doelen kunnen bereikt worden door in te grijpen op meerdere niveaus. Ten eerste zijn een snelle en accurate diagnose van CDI patiënten en moleculaire typering essentiële instrumenten om infecties en uitbraken te herkennen en onder controle te krijgen. Ten tweede kan het bestuderen en begrijpen van de mechanismen die belangrijk zijn voor kolonisatie leiden tot nieuwe therapeutische opties, die gericht zijn op het voorkomen van kolonisatie.

Tot slot kan het begrijpen van mechanismen van toxine-regulatie leiden tot de ontwikkeling van effectieve behandelmethodes om de klinische symptomen van CDI te verminderen.

Moleculaire methoden voor de diagnose en typering van *C. difficile*

Om de verspreiding van *C. difficile* te beperken is het van belang dat het ziekenhuispersoneel op de hoogte is van het voorkomen van CDI. Verder zijn (goede) hygiëne voorzorgsmaatregelen, vroegtijdige diagnose en het monitoren van CDI essentieel. De diagnose van CDI is nog steeds gebaseerd op de typische klinische kenmerken, in combinatie met laboratoriumtesten. Echter, de testen die momenteel het meest gebruikt worden zijn tijdrovend of hebben een gebrek aan sensitiviteit en specificiteit voor nauwkeurige en snelle diagnose van CDI. Deze beperkingen hebben ertoe geleid dat meerdere monsters per patiënt getest worden en dat voor de diagnose van CDI twee-stap algoritmes gebruikt worden. Echter, voor preventie van de verspreiding van CDI zijn snellere en meer accurate testen essentieel.

In **hoofdstuk 2** beschrijven we de evaluatie van drie snelle, binnen het instituut ontwikkelde, moleculaire testen voor de diagnose van CDI. Alle drie geëvalueerde moleculaire testen voldeden door de hoge prevalentie van CDI in de gebruikte monster collectie (20 %) bijna aan de criteria die gedefinieerd zijn voor het gebruik als zelfstandige test. Echter, de berekende positieve voorspellende waarde daalde aanzienlijk bij een meer realistische prevalentie van CDI in een klinische omgeving (10 %) en in de gemeenschap (2 %). Daarom hebben wij geconcludeerd dat de moleculaire testen niet gebruikt kunnen worden als een snelle zelfstandige test voor de diagnose van CDI.

Een nadeel van een moleculaire test is het onvermogen om onderscheid te kunnen maken tussen asymptomatisch dragerschap en CDI. Een onlangs gepubliceerde studie heeft aangetoond dat alleen de detectie van toxinen positief correleert met klinisch verloop van de ziekte en niet de aanwezigheid van *C. difficile* per se. Niettemin kan asymptomatisch dragerschap een belangrijke bron van verspreiding binnen ziekenhuizen zijn. Door de hoge negatieve voorspellende waarden van de geëvalueerde moleculaire testen (**hoofdstuk 2**), zullen deze zeer waardevol zijn als een eerste screening test in een twee-stap algoritme, gevolgd door een snelle fenotypische toxine-detectie methode (bijvoorbeeld een Enzyme Immuno-Assay) voor de diagnose van CDI. Het grootste voordeel van dit tweestaps algoritme boven de tot nu toe gebruikte tweestaps algoritmes is de snelle diagnose van CDI met bijbehorende gevoeligheid en specificiteit.

Een andere belangrijke methode in de bestrijding van CDI is om uitbraken binnen zorginstellingen in kaart te brengen met moleculaire (sub-) typeringsmethoden. De

Multi Locus Variable number of tandem repeat Analysis (MLVA) is momenteel een van de meest onderscheidende en reproduceerbare typeringsmethoden. De MLVA kan onderscheid maken tussen stammen van hetzelfde PCR-Ribotype (RT) en hierdoor is het mogelijk om transmissie routes te identificeren. In **hoofdstuk 3** beschrijven we de toepassing van een geoptimaliseerde MLVA op een collectie van RT 078 *C. difficile* stammen, geïsoleerd van mensen en varkens. Ondanks het ontbreken van een rechtstreeks epidemiologisch verband, concludeerden we dat er een grote verwantschap bestaat tussen de menselijke en varkensstammen. Deze hoge verwantschap kan een indicatie zijn van een gemeenschappelijke bron of transmissie. In **hoofdstuk 3** beschrijven we ook dat de MLVA enkele zwakke punten heeft. Vooral de lage variabiliteit van enkele loci in sommige RTs, de afwezigheid van een locus in RT 078 en de benodigde optimalisatie voor de PCRs voor drie loci in RT 078 zijn de tekortkomingen van de MLVA als (sub-)typerings methode. Samen met de mogelijke aanwezigheid van meerdere MLVA types binnen een fecaal monster zou dit kunnen leiden tot het missen van epidemiologische verbanden. Het gebruik van Whole Genome Sequencing (WGS) als typerings methode zou op termijn wellicht een alternatief kunnen zijn dat veel van deze nadelen ondervangt.

Non-toxinogene virulentiefactoren van *C. difficile*

De transmissie van CDI verloopt via de fecaal-orale route. Het voorkomen van kolonisatie is een van de beste manieren om in te grijpen in de verspreiding en overdracht van *C. difficile*. Sporen spelen een belangrijke rol in de transmissie en kolonisatie van een gastheer. Na inname ontkiemen de sporen en groeien uit tot vegetatieve cellen die zich kunnen hechten aan de epitheelcellen in de dikke darm. Succesvolle kolonisatie is naast sporulatie afhankelijk van het vermogen van *C. difficile* om te kunnen omgaan met het veranderende milieu en gastheer-geïnduceerde stressfactoren, zoals antibiotica en zuurstofradicalen. De aanwezigheid van mobiele elementen die resistentiegenen bevatten of nieuwe metabole routes mogelijk maken is hierop van grote invloed.

De sporen van *Clostridium spp* en Bacillus *spp* worden geproduceerd om ongunstige omstandigheden te overleven. Het sporulatie proces is in grote lijnen hetzelfde in *Clostridium spp* en Bacillus *subtilis*. Sporulatie van *C. difficile* wordt erkend als een persistentie- en transmissie-factor voor CDI. In **hoofdstuk 7** beschrijven we dat *C. difficile* Spo0A bindt aan DNA sequenties die lijken op de herkenningssequentie van Bacillus *subtilis* Spo0A. Daarnaast hebben we ook aangetoond dat Spo0A bindt aan de promoter van het toxine B gen. We vonden echter geen verlaagd toxineniveau in de Spo0A mutant, in tegenstelling tot een eerdere publicatie.

Naast sporulatie is hechting aan de epitheelcellen van de darm mogelijk een belangrijke stap in de succesvolle kolonisatie van de gastheer. Er is echter relatief

weinig bekend over de hechting van *C. difficile* aan de epitheelcellen en de rol van hechting in de pathogenese van CDI. Tot op heden zijn er slechts een aantal oppervlakte-eiwitten beschreven die een rol spelen in de hechting en kolonisatie. Van de expressie van hechting- en kolonisatie-factoren word verondersteld dat deze verhoogd is onder stressomstandigheden.

In **hoofdstuk 8** hebben we het gen dat codeert voor het HtrA protease, dat mogelijk betrokken is bij de reactie op stress, gemuteerd. De micro-array analyse van deze *htrA* mutant toonde verlaagde genexpressie aan van een aantal bekende hechting- en kolonisatie-factoren, zoals SlpA en flagellaire eiwitten. Verdere experimenten gaven aan dat de mutant minder sporuleerde en minder goed hechtte aan epitheelcellen in vergelijking met wild-type. Daarnaast vertoonde de *htrA* mutant ook een verhoogde toxine A productie. Dit, in combinatie met de hoge gevoeligheid van hamsters voor toxinen, verklaart mogelijk de verhoogde virulentie van de *htrA* mutant in het hamstermodel van CDI. Om te onderzoeken welke van de aangetoonde fenotypen (verminderde sporulatie, verminderde hechting of verhoogde toxine A productie) meer relevant is voor de ontwikkeling van de ziekte in mensen zullen *in vivo* experimenten gedaan moeten worden in een diermodel dat minder gevoelig is voor toxinen.

In veel micro-organismen kan resistentie tegen antibiotica verworven worden via incorporatie van mobiele elementen. Antibiotica resistentie kan resulteren in een groeivoordeel tijdens de kolonisatie en geeft de mogelijkheid om de commensale flora te overgroeien na een antibiotica behandeling. Naast resistentiegenen dragen mobiele elementen ook vaak virulentie factoren met zich mee. Daarom kan de verspreiding van mobiele elementen grote invloed hebben op de fitness van *C. difficile*. Daarnaast kan de aanwezigheid van mobiele elementen aanvullende informatie geven over mogelijke transmissie routes.

In **hoofdstuk 3 en 4** beschrijven we de aanwezigheid van twee mobiele elementen, die coderen voor antibiotica resistentie. In **hoofdstuk 3** laten we zien dat stammen geïsoleerd uit varkens en mensen in hoge mate verwant zijn, en een gemeenschappelijke resistentie tegen tetracycline hebben die afkomstig is van een Tn*916*-achtig element. Interessant is dat tetracycline veelvuldig gebruikt wordt in de veterinaire sector, maar niet in de (Nederlandse) gezondheidszorg. Dit zou kunnen betekenen dat tetracycline resistentie wordt doorgegeven aan de bevolking vanuit de veterinaire industrie. De aanwezigheid van een Tn*916*-achtig element en tetracycline resistentie zou dus mogelijk extra informatie kunnen geven over transmissie routes tussen de veterinaire industrie en mensen.

In **hoofdstuk 4** beschrijven we het mobiele element Tn*6164* dat naast diverse resistentie genen ook codeert voor (potentiële) virulentie factoren. De verhoogde mortaliteit van patiënten geïnfecteerd met RT 078 met Tn*6164* suggereert

dat de aanwezigheid van dit element ervoor zorgt dat deze stammen virulenter worden. Echter, verder onderzoek is nodig om het mogelijke verband tussen de aanwezigheid van Tn*6164* en verhoogde virulentie te bevestigen. Bij elkaar genomen onderstrepen onze resultaten het belang om resistentie tegen antibiotica te monitoren in *C. difficile* stammen, afkomstig van zowel mensen als dieren.

De aanwezigheid van mobiele elementen wordt op dit moment niet herkend door de vaak gebruikte "PCR ribotypering"-methode en wordt daarom niet meegenomen in de (moleculaire) epidemiologie van CDI. Bovendien worden mobiele elementen vanwege hun instabiele aard en locatie in het "aanvullende" genoom, niet herkend door de WGS Single Nucleotide Polymorphisme (SNP) typeringsmethode. Toch heeft WGS bewezen dat het een krachtig instrument kan zijn om antibioticaresistentie en mogelijke virulentie genen, gecodeerd op mobiele elementen, te herkennen (**hoofdstuk 4** en C.W. Knetsch, persoonlijke communicatie). Daarom zou WGS niet alleen gericht moeten zijn op de verschillen in het "kern" genoom, maar ook op de aanwezigheid van de mobiele elementen in het "aanvullende" genoom. Toepassing van WGS als typeringsmethode, in combinatie met een goed monitoring systeem, zou kunnen resulteren in een beter begrip van de rol die mobiele elementen spelen bij de verspreiding en virulentie van *C. difficile*.

Regulatie van toxine productie

Toxine A en toxine B zijn de twee belangrijkste virulentiefactoren van *C. difficile* en primair verantwoordelijk voor de symptomen van CDI. Het bestuderen van de regulatiemechanismen van toxineproductie zou kunnen leiden tot de ontwikkeling van effectieve maatregelen om het klinische verloop en de uitkomst van de ziekte te kunnen verbeteren. Beide toxinen genen zijn gelegen op het Pathogenicity Locus (PaLoc), wat daarnaast ook een gen (*tcdC*) bevat dat codeert voor een veronderstelde negatieve regulator (TcdC) van toxineproductie. In de **hoofdstukken 5** en **6** hebben we verschillende aspecten van TcdC onderzocht.

In **hoofdstuk 5** onderzochten we de transcriptie niveaus van de toxine genen en de daadwerkelijke productie van toxines in een *tcdC* mutant en een wild type stam. We concludeerden dat TcdC geen belangrijke negatieve regulator van toxine productie is. Dit is in tegenspraak met eerdere rapporten die aantonen dat TcdC een belangrijke negatieve regulator is van toxinen productie. We kunnen echter niet uitsluiten dat TcdC onder andere omstandigheden en/of in fylogenetisch verschillende stammen toch een effect heeft op de productie van toxinen.

In **hoofdstuk 6** beschrijven we een tot nu toe onbekende functie voor TcdC. Bioinformatische analyse voorspelt dat de C-terminus van TcdC een Oligonucleotide Binding vouwing (OB-fold) heeft en mogelijk kan binden aan

DNA. Een biochemische analyse laat zien dat TcdC kan binden aan zogenaamde quadruplex structuren (QS) waarvan is aangetoond dat ze verscheidende *in vivo* processen (zoals transcriptie) kunnen beïnvloeden. Omdat er echter geen QS voorspeld worden in de PaLoc-regio, is het verband tussen QS-binding en toxine-regulatie vooralsnog onduidelijk. Verder onderzoek is nodig om meer inzicht te kunnen krijgen in de functie van TcdC.

Conclusie

In de afgelopen jaren is er grote voortuitgang geboekt op het gebied van moleculair en epidemiologisch onderzoek naar *C. difficile*. Ondanks deze vooruitgang, zijn de medische en financiële gevolgen voor patiënten en voor de gezondheidszorg nog niet afgenomen. Dit proefschrift toont aan dat de ontwikkeling van moleculaire technieken, gebaseerd op de detectie en typering van *C. difficile,* waardevol kunnen zijn in de signalering en preventie van CDI. Hetzelfde geldt ook voor de rol van de moleculaire biologie met betrekking tot het begrijpen van de virulentie van *C. difficile*. Echter, de voortdurende ontwikkeling van nieuwe technieken, voor zowel typeringen als genetische manipulatie, heeft ook geleid tot inzicht in de potentiële zwakheden van de momenteel veel gebruikte technieken. Daarom is het erg belangrijk om op de hoogte te blijven van de technische vooruitgang op het gebied van *C. difficile* onderzoek om de mechanismes van virulentie verder te kunnen ontrafelen.

Bibliography

De Jonge R, Bakker D, Vliet AH, Kuipers EJ, Vandenbroucke-Grauls CM, Kusters JG. Direct random insertion mutagenesis of Helicobacter pylori. J Microbiol Methods. **2003** Jan;52(1):93-100

Drudy D, Goorhuis B, Bakker D, Kyne L, van den Berg R, Fenelon L, Fanning S, Kuijper EJ. Clindamycin Resistant Clone of *Clostridium difficile* PCR Ribotype 027, Europe. Infect Dis. **2008** Nov;14(11):1825.

Goorhuis A, Bakker D, Corver J, Debast SB, Harmanus C, Notermans DW, Bergwerff AA, Dekker FW, Kuijper EJ. Emergence of *Clostridium difficile* infection due to a new hypervirulent strain, PCR-ribotype 078. Clin Infect Dis. **2008** Nov 1;47(9):1162-70.

Pituch H, Bakker D, Kuijper E, Obuch-Woszczatyński P, Wultańska D, Nurzyńska G, Bielec A, Bar-Andziak E, Łuczak M. First isolation of *Clostridium difficile* PCR-ribotype 027/toxinotype III in Poland. Pol J Microbiol. **2008**;57(3):267-8.

Ergen EK, Akalin H, Yilmaz E, Sinirtaş M, Alver O, Heper Y, Ozakin C, Bakker D, Ener B, Mistik R, Helvaci S, Kuijper EJ. Nosocomial diarrhea and *Clostridium difficile* associated diarrhea in a Turkish University Hospital. Med Mal Infect. **2009** Jun;39(6):382-7.

Keessen EC, Leengoed LA, Bakker D, van den Brink KM, Kuijper EJ, Lipman LJ. Prevalence of *Clostridium difficile* in swine thought to have *Clostridium difficile* infections (CDI) in eleven swine operations in the Netherlands. Tijdschr Diergeneeskd. **2010** Feb 15;135(4):134-7.

Svraka S, Kuijper E, Duizer E, Bakker D, Koopmans M. *Clostridium difficile* is not associated with outbreaks of viral gastroenteritis in the elderly in the Netherlands. Eur J Clin Microbiol Infect Dis. **2010** Jun;29(6):677-82.

Bakker D, Corver J, Harmanus C, Goorhuis A, Keessen EC, Fawley WN, Wilcox MH, Kuijper EJ. Relatedness of human and animal *Clostridium difficile* PCR Ribotype 078 isolates based on Multi Locus Variable number of tandem repeat Analysis and tetracycline resistance. J Clin Microbiol. **2010** Oct;48(10):3744-9.

Knetsch CW†, Bakker D ‡, de Boer RF, Sanders I, Hofs S, Kooistra-Smid AM, Corver J, Eastwood K, Wilcox MH, Kuijper EJ. Comparison of Real-Time PCR Techniques to Cytotoxigenic Culture Methods for Diagnosing *Clostridium difficile* Infection. J Clin Microbiol. **2011** Jan;49(1):227-31. † These two authors contributed equally to the study.

Corver J, Bakker D, Brouwer MS, Harmanus C, Hensgens MP, Roberts AP, Lipman LJ, Kuijper EJ, van Leeuwen HC. Analysis of a *Clostridium difficile* PCR Ribotype 078 100 kilobase island reveals the presence of a Novel transposon, Tn*6164*. BMC Microbiol. **2012** Jul 2;12:130.

Bakker D, Smits WK, Kuijper EJ, Corver J. TcdC does not significantly repress toxin expression in *Clostridium difficile* 630**Δ***erm*. Plos One. **2012**;7(8):e43247.

Rosenbusch KE, Bakker D, Kuijper EJ, Smits WK. *Clostridium difficile* 630Δ*erm* SpoOA regulates sporulation, but does not contribute to toxin production, by direct high-affinity to target DNA. Plos One **2012** ;7(10):e48608.

van Leeuwen HC, Bakker D, Steindel P, Kuijper EJ, Corver J. *Clostridium difficile* TcdC protein binds four-stranded G-quadruplex structures. Nucleic Acids Res. **2013** Feb 1;41(4):2382-93.

van den Berg RJ, Bakker D, Kuijper EJ.Diagnosis of *Clostridium difficile* infection using real-time PCR. Methods Mol Biol. **2013**;943:247-56.

Hensbergen PJ, Klychnikov OI, Bakker D, van Winden VJ, Ras N, Cordfunke RA, Dragan I, Deelder AM, Kuijper EJ, Corver J, Drijfhout JW, van Leeuwen HC. A novel secreted metalloprotease (CD2830) from *Clostridium difficile* cleaves specific proline sequences in LPTXG cell surface proteins. Molecular Cell Proteomics **2014** May;13(5):1231-44.

Bakker D, Buckley AM, de Jong A, Kuipers OP, Douce G, Kuijper EJ, Smits WK, Corver J. The HtrA-like protease CD3284 modulates virulence of *Clostridium difficile*. Infection and Immunity **2014** Jul 21

Curriculum Vitae

Dennis Bakker, zoon van Pieter Willem Bakker en Carolina Liduina Maria Bergsteijn, kwam op 7 november 1975 te Gorinchem ter wereld. In 1992 behaalde hij zijn MAVO diploma aan OSG De Randijk te Nieuwegein. In datzelfde jaar starte hij met de MBO studie Laboratorium technieken in de richting Medische Microbiologie aan het Utrecht College, waar hij in 1996 zijn diploma behaalde. In 1996 werd deze opleiding vervolgd met de studie Hoger Laboratorium Onderwijs (HLO) aan de Hogeschool van Utrecht. Direct na het behalen van zijn HLO diploma in 2000 starte hij als Medisch Microbiologisch analist in het VuMC. In het jaar 2003 begon hij als Medisch Microbiologisch analist in het LUMC. In 2006 maakte hij de overstap naar het bedrijfsleven (Crucell) als research analist. Op deze plek werd zijn nieuwsgierigheid en interesse in het wetenschappelijk onderzoek gewekt. De gewekte nieuwsgierigheid en interesse werd verder uitgebouwd door als research analist in het LUMC bij de Prof. Dr. E. Kuijper aan de slag te gaan. De ambitie om verder te gaan in het wetenschappelijk onderzoek heeft geleid tot het starten van zijn promotie-onderzoek bij het LUMC te leiden onder de supervisie Prof. Dr. E. Kuijper, Dr. Ir. J. Corver en Dr. W.K. Smits. Dit promotie-onderzoek heeft geleid tot dit proefschrift. Sinds april 2014 werkt hij bij Janssen BV Biologics als Team leider van de Bio Assay team op de Quality control afdeling Final Product testing.

Stellingen behorende bij het proefschrift

Molecular characterization of pathogenic *Clostridium difficile* strains

1. De hoge negatieve voorspellende waarde en de snelheid van real-time PCR maakt deze detectietechniek zeer geschikt als eerste screeningstest in een twee-staps algoritme voor de diagnose van *Clostridium difficile* infecties (dit proefschrift).

2. Het frequente gebruik van antibiotica in de veterinaire industrie kan leiden tot de verspreiding van *Clostridium difficile* bij consumptiedieren en vormt een potentiële bron voor humane infecties met (multi)resistente en virulentere *C. difficile* stammen (dit proefschrift).

3. Het ontkrachten van een paradigma op grond van *in vivo* experimenten kunnen de basis vormen voor nieuwe inzichten in de functie van TcdC (dit proefschrift).

4. Het uitschakelen van een eiwit (HtrA) in de stressresponse cascade hoeft niet altijd te leiden tot een minder virulente stam (dit proefschrift).

5. Het gebruik van verschillende genetische manipulatie technieken en verschillende *Clostridium difficile* stammen leidt tot verschillende uitkomsten (PLoS One. 2013 Nov 13;8(11), Nature. 2009 Apr 30;458(7242), Nature. 2010 Oct 7;467(7316) en dit proefschrift).

6. Asymptomatisch *Clostridium difficile* dragerschap is een belangrijke bron van ziekenhuis-infecties (N.Engl.J.Med. 2013; 369:1195-1205.

7. Een goede samenwerking tussen fundamentele onderzoekers en epidemiologen is een uitstekende mogelijkheid om potentiële virulentie genen te identificeren en hiermee de verdere verspreiding van virulente *Clostridium difficile* stammen te beperken (Microbiology. 2011; 157:3113-23).

8. Fecale transplantatie ("poep transplantatie") is een veilige en efficiënte behandeling tegen hardnekkige en terugkerende *Clostridium difficile* infecties (N.Engl.J.Med 2013;368:407-15).

9. "Een investering in kennis betaalt zich terug met de hoogste rente." (Benjamin Franklin, 1750, Poor Richard's Almanac).

Dennis Bakker, Leiden 2014