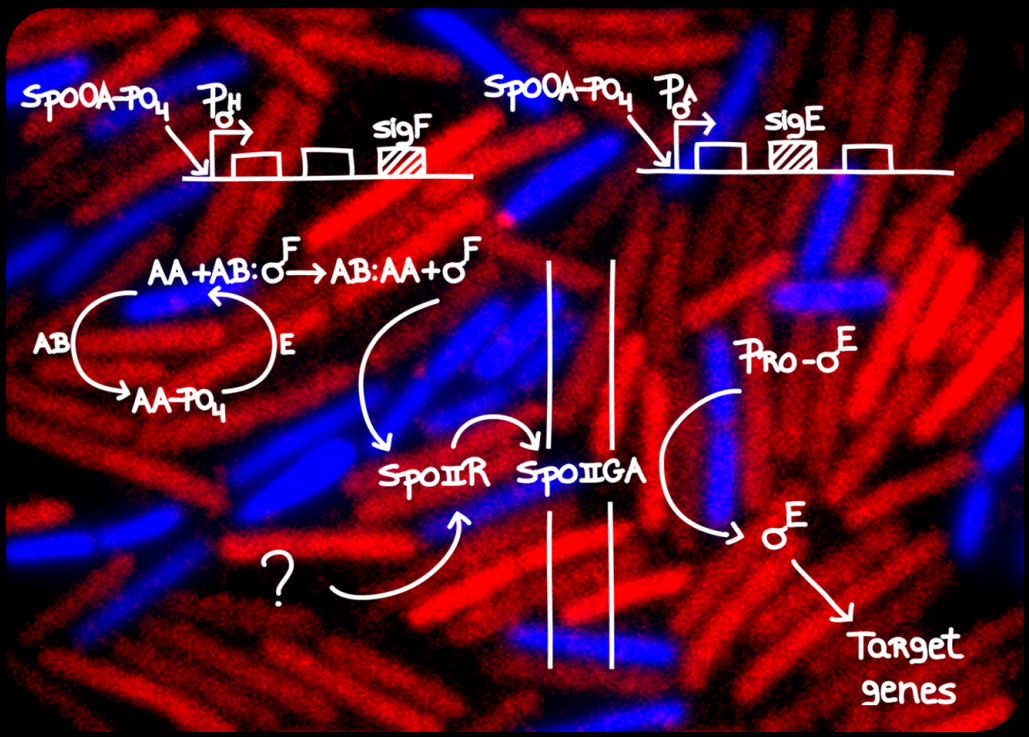
Spore differentiation in relation to the infectious cycle of the enteric pathogen *Clostridium difficile*

Fátima C. Pereira



Dissertation presented to obtain the Ph.D degree in Biology

Instituto de Tecnologia Química e Biológica António Xavier | Universidade Nova de Lisboa

Oeiras,
May, 2014

INSTITUTO
DE TECNOLOGIA
QUÍMICA E BIOLÓGICA
ANTÓNIO XAVIER /UNL
Knowledge Creation

Spore differentiation in relation to the infectious cycle of the enteric pathogen *Clostridium difficile*

Fátima C. Pereira

Dissertation presented to obtain the Ph.D degree in Biology

Instituto de Tecnologia Química e Biológica António Xavier | Universidade Nova de Lisboa

Oeiras, May, 2014

Financial support from Fundação para a Ciência e Tecnologia
through grant SFRH/BD/45459/2008 awarded to Fátima Pereira

Nas partes desta Tese que se encontram escritas em Português não foi aplicado o novo acordo ortográfico

ACKNOWLEDGEMENTS

To Instituto de Tecnologia Química e Biológica António Xavier of Universidade Nova de Lisboa, for receiving me as a PhD student, for providing all the working conditions and the scientific environment for the execution of this work.

To Fundação para a Ciência e Tecnologia (FCT) and European Molecular Biology Organization (EMBO), for the financial support.

To Instituto Gulbenkian de Ciência, for my initial training in the PhD Programme (PDIGC), and to all my colleagues at the time.

To my supervisor Dr. Adriano O. Henriques, for the opportunity to work in his lab and for his continuous help and support throughout this journey. For all I learned from him and for sharing with me his passion for science.

To Dr. Bruno Dupuy, Prof. Isabelle Martin-Verstraete and Dr. Claire Janoir for welcoming me in their labs, providing me a rich scientific environment as well as excellent working conditions. A special *merci* to Dr. Bruno Dupuy for his constant interest and support during the development of this work.

To Mónica Serrano for all the helpful discussions, enthusiasm and sense of humour. Thanks for your assistance and positive attitude.

To all the present and former colleagues and lab neighbours: Gonçalo Real, Lia Domingues, Pedro Matos, Magda Atilano, Ana Jorge, Mafalda Henriques, Helena Veiga, Bruno Fonseca, Luísa Côrte, Catarina Fernandes, João Bota, Ana Paiva, Patrícia Amaral, Wilson Antunes, Teresa Maio, Nejka Planinc, Tiago dos Vultos, Joana Rodrigues, Gabriela Dabrek, Carolina Freitas and Carolina Cassona, for the great work atmosphere. To Teresa Silva, always willing to help. To Filipa Nunes, Teresa Costa, Anabela Isidro, Cláudia Serra and Rita Tomé, who more than colleagues, are my friends. To Anabela, for guiding me during my first steps in this lab and her good advises. To Cláudia,

for her company during the nights spent at the lab and her support in the times when this journey seemed endless. To Teresa, for her friendship and relaxed moments, having lunch at the beach or watching a movie. To Filipa, for showing me how to always laugh of the misfortune. To Rita, for the free hugs, encouragement and contagious joy.

To Laure Saujet, for our excellent scientific interaction that resulted in joint papers.

To Marc Monot, Laure Saujet, Emilie Camiade, Evelyne Couture-Tosi, Ana Antunes, Rita Tomé and Sandra Hoys, for the lab support and scientific help during my visits to Institut Pasteur and to Faculté de Pharmacie, Paris.

To Patrícia Amaral, for her valuable help with the image on the cover.

To all of my friends who have supported me during the dark times and celebrated with me through the good. A special thank to Guida, for her unconditional friendship and loyalty.

To my parents for their unconditional love, they will always be an inspiration to me. Finally, to my brother, for his constant support and encouragement throughout my life.

ABSTRACT

Clostridium difficile is a gram positive, spore former, anaerobic bacterium that is able to cause infection and disease, with symptoms ranging from mild diarrhea to pseudomembranous colitis, toxic megacolon, sepsis and death. In the last decade new strains have emerged that caused outbreaks of increased disease severity and higher recurrence, morbidity and mortality rates, and *C. difficile* is now considered both a main nosocomial pathogen associated with antibiotic therapy as well as a major concern in the community. Spores produced by *C. difficile* are extremely difficult to eradicate, can accumulate in the environment and in the host for long periods of time, and are the main vehicle for transmission and persistence of the organism. Although spores are critical to the pathogenesis of *C. difficile*, their composition and formation remain poorly characterized, especially in comparison with the model organism *Bacillus subtilis*. Sporulation in *B. subtilis* is controlled by the sequential activation of four RNA-polymerase sigma subunits that regulate gene expression during spore formation. σ^F in the forespore and σ^E in the mother cell control the initial stages of spore morphogenesis, and are replaced at later stages by σ^G and σ^K, respectively. A comprehensive description of the sporulation pathway in *B. subtilis* relied on the employment of GFP as a fluorescent reporter. However, the use of GFP as a reporter requires the presence of oxygen. To overcome this limitation, we have implemented a fluorescent reporter for *C. difficile*, which is based on the SNAP-tag technology. The SNAP coding sequence was optimized for expression in *C. difficile*, and we demonstrate here that it can be used as a quantitative reporter to measure gene expression at the single cell level as well as a tag for protein localization in *C. difficile*. We have characterized the mutants for the *C. difficile* sigma factors of sporulation at the morphological level, and by employing the SNAP-tag reporter we have established the main periods and the cell type for the transcription and activity of each of these

sigma factors. Our results indicate that while the main periods of activity of the sigma factors are conserved in *C. difficile,* a reduced temporal segregation between the activities of early and late sigma factors is observed when in comparison with the model organism *B. subtilis*. We further observed that the activity of σ^E is partially independent of σ^F, despite the fact that the forespore product SpoIIR is required for pro-σ^E processing. In addition, σ^G activity is not dependent on σ^E, and the activity of σ^K does not require σ^G. In agreement, our genome-wide analysis of gene expression during spore formation confirmed a weaker connection between the mother cell and the forespore in *C. difficile*. This global transcriptional analysis allowed the identification of 225 genes controlled by each of these sporulation sigma factors: 25 in the σ^F regulon, 97 σ^E-controlled genes, 50 σ^G-dependent genes and 56 genes under σ^K control. Our analysis of a mutant for the transcriptional regulator SpoIIID also demonstrates that many of the σ^E target genes here identified are either repressed or activated by this transcriptional regulator, including *sigK*, which is activated. While most of the genes known to be essential for sporulation in *B. subtilis* are conserved, many other genes, especially in the σ^E and σ^K regulons, are novel. Our analysis supports the view that the top level in the control of the sporulation network, defined by the cell type-specific sigma factors is conserved in evolution, together with a core of target genes possibly representing the conserved machinery for endosporulation, while others are group specific.

One of the genes expressed under the control of σ^K is *CD1581*, coding for a previously uncharacterized 17 kDa cysteine-rich protein. Here we show that this protein, termed Sp17, is the most abundant component of the spore surface layers and is required for efficient colonization in a mouse axenic model. Spores of an *sp17* mutant fail to assemble a spore polar structure, observed in the wild type spores, which may serve a role in adhesion to the colonic epithelium. Thus, our study unveils a link between spore surface and colonization. Furthermore, Sp17 is unique of *C. difficile*, suggesting that the

outer spore coat of this organism has distinctive structural and functional properties that allowed it to adapt to its host and to efficiently colonize it.

SUMÁRIO

Clostridium difficile é uma bactéria anaeróbia estrita, de reacção de Gram-positiva, formadora de esporos, e capaz de causar infecção e doença, com sintomas variáveis que podem ir desde diarreia leve até colite pseudomembranosa, megacólon tóxico, sépsis e morte. Na última década surgiram várias estirpes epidémicas de *C. difficile* associadas a surtos de maior severidade da doença, com maiores taxas de incidência, morbidez e mortalidade, que tornaram este organismo não só na causa mais comum de infecções nosocomiais associadas a tratamento por antibiótico em adultos, mas também numa ameaça séria a nível da comunidade. Os esporos produzidos por *C. difficile* são extremamente difíceis de erradicar, podem acumular no meio ambiente ou no hospedeiro por períodos longos de tempo, e são o principal meio de transmissão e persistência deste organismo. Apesar da importância dos esporos na patogénese de *C. difficile*, pouco se sabe sobre a sua constituição ou processo de formação, principalmente em comparação com o organismo modelo *Bacillus subtilis*. Em *B. subtilis* o processo de esporulação é controlado através da activação sequencial de quatro subunidades sigma da polimerase do RNA, que regulam a expressão genética no decorrer do processo de formação do esporo. Os factores σ^F no pré-esporo e σ^E na célula-mãe controlam os estadios iniciais da morfogénese sendo substituídos, nos estadios mais tardios, pelos factores σ^G e σ^K, respectivamente. A descrição pormenorizada da via de esporulação em *B. subtilis* baseou-se em larga medida no utilização da GFP como repórter de expressão génica e de localização sub-celular de proteínas. No entanto, a reacção de formação do cromóforo da GFP requer oxigénio. De forma a ultrapassar esta limitação foi desenvolvido neste trabalho um repórter fluorescente para *C. difficile* com base na tecnologia SNAP-*tag*. A sequência SNAP foi optimizada para expressão em *C. difficile*, e demonstrou-se a sua utilização para monitorizar quantitativamente a expressão génica e a

localização sub-celular de proteínas neste organismo. Procedeu-se então à caracterização dos mutantes dos factores sigma de esporulação de *C. difficile*, tendo sido estabelecidos a janela temporal e a célula em que ocorre a sua transcrição e actividade. Em geral, os resultados aqui apresentados indicam que, enquanto que os principais períodos de actividade dos factores sigma são conservados em *C. difficile*, existe uma redução na segregação temporal entre a actividade dos factores sigma iniciais ou tardios quando em comparação com o organismo modelo *B. subtilis*. Observou-se ainda que a actividade de σ^E é parcialmente independente de σ^F, embora o papel da proteína SpoIIR do pré-esporo na sinalização da activação proteolítica de pro-σ^E seja mantido. Para além disso, a actividade de σ^G também não depende de σ^E, assim como a actividade de σ^K não requer σ^G. Em consonância com estes resultados, uma análise global da expressão genética de *C. difficile* durante a esporulação confirmou uma ligação mais ténue entre os programas genéticos do pré-esporo e da célula-mãe. Esta análise transcripcional global permitiu ainda a identificação de 225 genes sob o controlo dos factores sigma de esporulação: 25 destes pertencem ao regulão σ^F, 97 são controlados por σ^E, 50 são dependentes de σ^G e 56 estão sob controlo de σ^K. Uma análise do mutante para o regulador transcripcional auxiliar SpoIIID demonstrou ainda que muitos dos genes pertencentes ao regulão σ^E são reprimidos ou ativados por este regulador, como é o caso de *sigK*, activado por SpoIIID. Enquanto que a maioria dos genes essenciais para a esporulação em *B. subtilis* são conservados, muitos outros, sobretudo nos regulões σ^E e σ^K, são novos. A nossa análise sugere que o nível superior de controlo da esporulação, definido pelos 4 factores sigma específicos do pré-esporo ou da célula mãe é conservado, juntamente com um grupo de genes alvo que possivelmente define a maquinaria mínima para a esporulação, enquanto que outros são específicos de determinados grupos ou organismos.

Um dos genes expresso sob o controlo de σ^K é *CD1581*, que codifica para uma proteína de 17 kDa, rica em cisteínas. Mostramos aqui que esta

proteína, denominada de Sp17, é o componente mais abundante da superfície do esporo, e é necessária para a eficiente colonização de ratinhos axénicos por *C. difficile*. Os esporos de um mutante *sp17* não são capazes de montar uma estrutura polar que é observada nos esporos selvagens, e que provavelmente desempenha um papel na aderência do esporo ao epitélio do cólon. O trabalho aqui desenvolvido estabelece assim uma ligação directa entre a superfície do esporo e a colonização. Além disso, Sp17 é exclusivo de *C. difficile*, o que sugere que a superfície do esporo neste organismo tem propriedades estruturais e funcionais distintas que representam, provavelmente adaptações à colonização dos seus hospedeiros.

x

THESIS OUTLINE

This Thesis is divided into six Chapters.

Chapter 1 provides an introduction to the enteric pathogen and spore former *Clostridium difficile*, with emphasis on sporulation, a determinant factor of pathogenesis. Since our knowledge about spore formation and composition is limited in *C. difficile*, the overview of sporulation provided by this Chapter is based on the model organism *Bacillus subtilis*. The text is however supplemented with available information about sporulation in *C. difficile* and other Clostridial spore formers.

Chapter 2 reports the development of a fluorescent reporter for single cell analysis of gene expression in *C. difficile*. We provide evidence that this fluorescent reporter, which is based on the SNAP-tag technology, can be employed as a quantitative reporter to monitor gene expression in this organism. Our results further demonstrate its applicability at the translational level, to determine the cellular localization of proteins.

Chapters 3 and **4** provide a detailed description of the sporulation regulatory network of *C. difficile*. The role played by the main regulators σ^F, σ^E, σ^G and σ^K during sporulation in *C. difficile* is presented on Chapter 3, together with the main periods of transcription and activity for these sigma factors and the co-dependencies for their expression and activity. Chapter 4 describes the identification of genes that are expressed during sporulation under the control of each of these sigma factors, and we further describe the role played by two additional regulators of sporulation, SpoIIID and SpoVT. Together, these Chapters provide evidence that in what concerns the genetic control, sporulation in *C. difficile* seems closer to the model organism *B. subtilis* than to other Clostridial spore formers studied so far. However, some deviations from the *B. subtilis* paradigm are observed: *i*) σ^K is not strictly required for heat resistant spore formation; *ii*) the periods of activity of each

sigma factor in *C. difficile* are less segregated in time than in the model organism *B. subtilis*; *iii*) the cell-cell signaling pathways that operate in *B. subtilis* are simpler or absent in *C. difficile*, resulting in a less tight control of the spore morphogenesis; *iv*) most members of the σ^E and σ^K regulons lack homologues in *B. subtilis*, suggesting a distinct composition for the surface layers of the *C. difficile* spore.

Chapter 5 reports the identification and characterization of Sp17, an abundant component of the spore coat, identified on Chapter 4 as a previously uncharacterized member of the σ^K regulon. The Chapter evidences a role for Sp17 in the assembly of the spore surface layers and in host colonization. The chapter establishes a direct link between the spore surface and host colonization in *C. difficile*.

Lastly, **Chapter 6** presents a general discussion of the results and points to future lines of work.

The results presented in **Chapters 2**, **3** and **4** are based on published material, as indicated in the beginning of each Chapter. Note however that some unpublished material was added when appropriate. Some of the results presented on **Chapter 5** are also based on published material, but most of the results presented in this Chapter constitute the basis for a future manuscript.

TABLE OF CONTENTS

Chapter 1

Introduction

THE ENTERIC PATHOGEN *CLOSTRIDIUM DIFFICILE*

We are not alone. A vast array of microorganisms lives on or inside the human body, with our intestine carrying the highest density and diversity of microbial cells that inhabit us. Our gut harbours more than 1000 of bacteria phylotypes and this microbial diversity is highly variable over time and across populations. Most (90%) of our gut microbiota belongs to only two phyla: the Bacteriodetes and the Firmicutes. Actinobacteria, Proteobacteria and Verrucomicrobia are also represented, but as minor constituents (Lozupone *et al.*, 2012). The trillions of bacteria that live in the human gut form a complex ecological community that helps in food digestion, protects against enteropathogens, and stimulate the immune system (Backhed *et al.*, 2004; Stecher and Hardt, 2011; Littman and Pamer, 2011). The host, in turn, provides habitat and a constant influx of nutrients (Berry *et al.*, 2013).

Besides sustaining a large number and diversity of commensal microbes, the human gut also provides habitat to a panel of harmful pathogenic microorganisms. The enteropathogen *Clostridium difficile* is a gram positive spore forming anaerobe that inhabits the mammalian gut. This organism is able to produce two potent cytotoxins and cause infection (Carter *et al.*, 2010; Carroll and Bartlett, 2011; Rupnik *et al.*, 2009; Dawson *et al.*, 2009; Shen, 2012).

C. difficile was first identified and described in 1935 by Hall and O´Toole as part of the intestinal microflora in neonates (Hall and O´Toole, 1935). Later, in 1977, Bartlett *et al*. isolated *C. difficile* from the faeces of hamsters with clindamycin-induced colitis, confirming this pathogen as the cause of antibiotic-induced disease in animals (Bartlett *et al.*, 1977). One year later, in 1978, multiple studies provided further evidence that *C. difficile* was

the causative agent of antibiotic-associated diarrhoea and colitis not only in animals, but also in humans (Larson *et al.*, 1978; Bartlett *et al.*, 1978). Symptoms of the disease caused by *C. difficile* infection (CDI) can range from mild diarrhoea, abdominal pain, fever and leucocytosis, to more severe symptoms such as pseudomembranous colitis, toxic megacolon, bowel perforation, sepsis and death (Rupnik *et al.*, 2009; Carroll and Bartlett, 2011; Dawson *et al.*, 2009).

Antibiotics and CDI

C. difficile is an enteric pathogen that relies on the disturbance of the normal gut microbiota to expand in the gut and cause infection (Britton and Young, 2012; Carroll and Bartlett, 2011; Rupnik *et al.*, 2009). Individuals with a normal, balanced microbiota are usually resistant to infection by *C. difficile.* This is because the commensal bacteria that inhabit the gut are able to directly or indirectly prevent colonization by this pathogen (Figure 1.1) (Britton and Young, 2012). Direct mechanisms for the inhibition of *C. difficile* growth involve the production of antimicrobial compounds, such as Thuricin CD, a bacteriocin produced by the intestinal bacterium *Bacillus thuringiensis* with narrow activity against *C. difficile* (Rea *et al.*, 2010). Another direct mechanism by which the gut microbiota prevents *C. difficile* colonization is via the transformation of bile acids, which have profound effects on *C. difficile* spore germination and vegetative growth (see also the *Spore germination* section) (Sorg and Sonenshein, 2008; Sorg and Sonenshein, 2010; Francis *et al.*, 2013; Theriot *et al.*, 2014).

Commensal bacteria may also contribute to inhibition of *C. difficile* growth in the intestine through indirect mechanisms, such as nutrient exhaustion or the stimulation of the host immune system (Figure 1.1). As an example of the first, the access to sialic acid, which is liberated upon depletion of the commensal microbiota, is determinant for the expansion of *C. difficile* in the gut (Ng *et al.*, 2013). Supporting this view, the pre-

colonization of the gut by non-toxicogenic *C. difficile* strains prevents the colonization by most toxinogenic strains, in a process referred to as "niche exclusion" (Nagaro *et al.*, 2013). Regarding the second mechanism, the commensal microbiota produces microbial associated molecular patterns (MAMPs) that stimulate the production of adaptive or innate effectors by the host immune system, preventing *C. difficile* expansion in the gut (Littman and Pamer, 2011; Britton and Young, 2012).

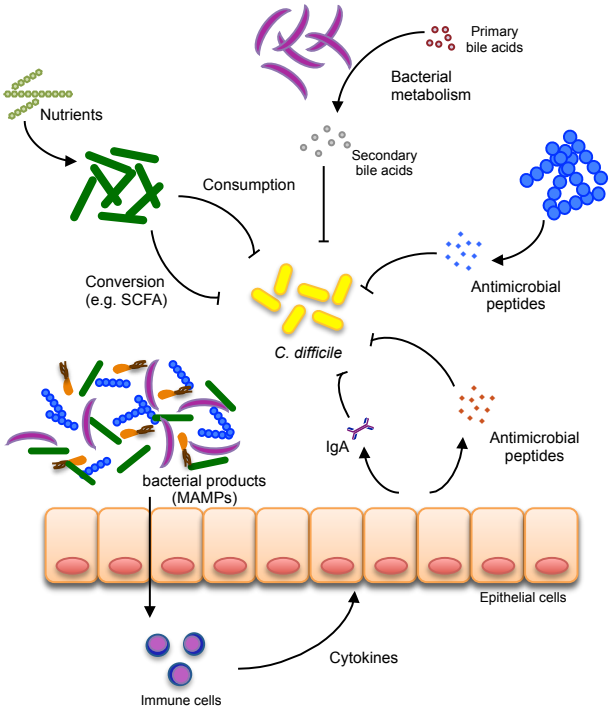


Figure 1.1. Potential mechanisms involved in colonization resistance against *C. difficile* in the gut. Direct inhibition of *C. difficile* by the gut microbiota can occur via competition for nutrients, conversion of host or diet compounds into secondary metabolites that inhibit *C. difficile* growth (*e.g.*, bile acids) or production of primary compounds with bacteriostatic activity against *C. difficile* (*e.g.*, bacteriocins). Indirect control of *C. difficile* expansion can occur via the stimulation of the host immune system through the production of microbial associated molecular patterns (MAMPs). Detection of MAMPs by the host induces the activation of an host immune cascade and the production of innate (*e.g.*, antimicrobial peptides) or adaptive (*e.g.*, IgA) immune effectors. Adapted from Britton and Young, 2012.

During antibiotic treatment, *C. difficile* can colonize the gut because, unlike most of the commensals, this pathogen resists to a wide range of

antibiotics. Resistance to antibiotics such as erythromycin, chloramphenicol or tetracycline is largely mediated by transposons that are present in the *C. difficile* genome (Lyras *et al.*, 2004; Hussain *et al.*, 2005; Sebaihia *et al.*, 2006). In contrast, resistance to fluoroquinolones is mediated by mutations in specific genes, such as point mutations in the A or B subunits of DNA gyrase (*gyrA* or *gyrB*)(He *et al.*, 2013).

The changing epidemiology

CDI is mainly a nosocomial infection, associated with elderly and debilitated patients, long hospital stays, prolonged periods of antibiotic therapy and the uptake of immunosuppressors (Kuipers and Surawicz, 2008; Rupnik *et al.*, 2009; Carroll and Bartlett, 2011). It is estimated that 7-17% of hospitalized patients carry toxinogenic strains of *C. difficile* (Poutanen and Simor, 2004; Liao *et al.*, 2012). The incidence of CDI has risen dramatically in the turn of this century, and it is nowadays the most common cause of antibiotic associated-diarrhoea in the industrialised world. This was because new, highly virulent *C. difficile* strains have emerged worldwide to cause outbreaks of increased disease severity and higher recurrence, morbidity and mortality rates (Kelly and LaMont, 2008; Brazier, 2008; Cartman *et al.*, 2010; Bartlett, 2010; Dawson *et al.*, 2009; Deneve *et al.*, 2009). Most of these strains belong to a particular PCR-ribotype, the 027 ribotype (Clements *et al.*, 2010; O'Connor *et al.*, 2009; Freeman *et al.*, 2010). A factor that contributed to the increased virulence of 027 strains was the acquisition of fluoroquinolone resistance (He *et al.*, 2013).

In the UK the appearance of these virulent 027 strains led to a drastic raise in the number of CDI cases, which have reached historical highs in 2006, with 55000 cases being detected in a single year (Figure 1.2A) (Brazier, 2012; Kuijper *et al.*, 2008; Bauer *et al.*, 2011). Fortunately, the implementation of prevention measures and the activity of infection control teams successfully conducted to a decrease in the number of deaths caused

by CDI in the UK after 2007 (Figure 1.2A) (Jones *et al.*, 2013) (Source: Health Protection Agency, http://www.hpa.org.uk). Similarly, in the USA the *C. difficile* related deaths also drastically rose from 5.7 deaths per billion inhabitants in 2001 to 23.7 in 2006 (Figure 1.2B) (Source: Centre for Disease Control, National Centre for Health Statistics, http://www.cdc.gov/nchs/). The medical costs associated with this disease also represent a major economic burden, with annual health-care costs of hospital-onset CDI estimated in 1.1 billion dollars in the USA and in 3 billion euros in Europe (data from the European Centre for Disease Prevention and Control, available at: http://www.ecdc.europa.eu/en/healthtopics/).

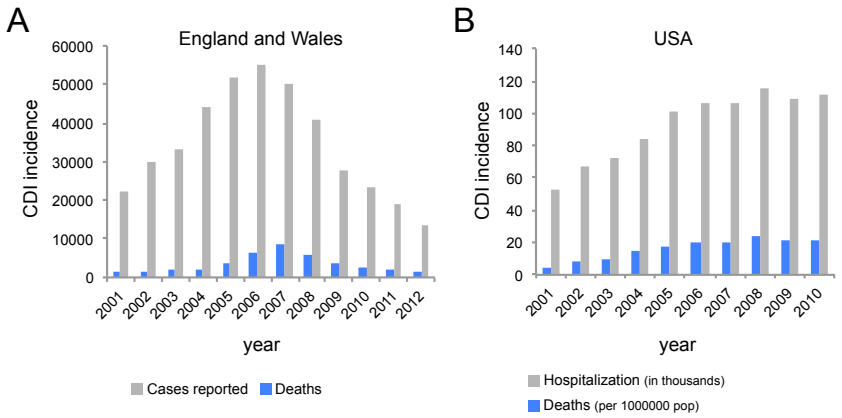


Figure 1.2. CDI incidence. (A) Total number of CDI reported cases (grey) and related deaths (blue) in England and Wales during the indicated time period. Source: Health Protection Agency, http://www.hpa.org.uk. **(B)** Hospitalization associated with CDI (in grey; numbers in thousands) and mortality rates (in blue, age-adjusted numbers per 1.000.000 population) in the USA during the indicated time period. Source for mortality data: Centre for Disease Control, National Centre for Health Statistics, http://www.cdc.gov/nchs/; source for hospitalization associated with CDI: Lucado *et al.*, 2009.

In Portugal the first detection of *C. difficile* 027 strains implicated in a hospital outbreak dates from January 2012, involving 12 patients, with a crude mortality rate of 50% (Antunes *et al.*, 2012). The CDI incidence in Portugal is estimated at 13 cases per 10000 hospital admissions (Bauer *et al.*, 2011).

The appearance of new virulent strains of *C. difficile* was followed by an increase in the number of community-acquired CDI cases (Freeman *et al.*, 2010; Hensgens *et al.*, 2012). Compared with hospitalized patients, the infected members of the community are younger, with fewer co-morbidities, and in some cases they have not received any antibiotic therapy in the 6 weeks prior to developing CDI, leading to the hypothesis that there is a yet unknown selection mechanism that favours the emergence of these strains (Freeman *et al.*, 2010; Hensgens *et al.*, 2012).

Diagnosis

The first indication of CDI is based on the presence of diarrhoea with a typical foul-smelling odour, conferred by the ability of *C. difficile* to produce and tolerate high concentrations of *p*-cresol, a bacteriostatic compound. However, detection of CDI based solely on the presence of an odour is obviously not an accurate and definitive diagnosis, and other methods are available (Carroll and Bartlett, 2011; Burnham and Carroll, 2013). For many years, the cell culture cytotoxicity neutralization assays (CCCNAs), which detect the cell toxicity of TcdA and TcdB present in a faecal eluate, were considered the gold standard. However, the method lacks sensitivity, and was recently replaced by new and improved diagnosis methods. Among these are the EIAs (enzyme immunoassays), in which polyclonal or monoclonal antibodies are used to target the TcdA and TcdB toxins, and the Glutamate dehydrogenase (GDH) test. GDH is a metabolic enzyme expressed at high levels in all *C. difficile* strains, toxinogenic and non-toxinogenic (Zheng *et al.*, 2004). Thus, the GDH test needs to be complemented with EIAs or other test that allows the detection of toxins (Snell *et al.*, 2004).

Lastly, nucleic acid amplification methods (NAATs), which are based on the isothermal amplification of conserved regions of the *tcdB* and *tcdA* genes, have recently become available (Goldenberg *et al.*, 2010). In the UK, new guidelines for CDI diagnosis recommend a combination of two tests, the

first of which should be a NAAT test or a GDH assay, followed by a toxin EIA test (DH/HCAI/Infectious disease, 2012).

Treatment

Conventional Treatment

Initial therapy for CDI includes stopping all antibiotic therapy, if possible, followed by the use of metronidazole or vancomycin, depending on illness severity and co-morbidities (Debast *et al.*, 2014). Metronidazole is the first line agent in treatment for mild to moderate CDI. However, it is becoming less effective as strains with decreased susceptibility to this antibiotic are appearing (Baines *et al.*, 2008; Moura *et al.*, 2013). This led to the adoption of vancomycin as the front-line antibiotic, together with fidaxomicin, a newly FDA approved antibiotic with cure rates equivalent to vancomycin, but that presents reduced recurrence rates (15% versus 25%) (Louie *et al.*, 2011). Unlike the other antibiotics, fidaxomicin specifically inhibits the formation of spores, whose accumulation in the host constitutes the most common cause of relapse (Babakhani *et al.*, 2012).

Reconstitution of the colonic microbiota

Recurrent CDI is an alarming and important problem. Patients with multiple episodes of CDI had markedly reduced microbial diversity when compared to patients that suffered only a single episode of CDI (Chang *et al.*, 2008). The main problem is that antibiotic treatment for a first infection can also eliminate many potential protective commensal bacteria, facilitating reinfection (Kelly and LaMont, 2008). Thus, the gut microbiota needs to be restored to protect the intestinal epithelium and to prevent residual spores from causing recurrent disease (Britton and Young, 2012). For this reason the development of non-antibiotic treatments for CDI has gained much attention in recent years. Among these, the faecal microbiota transplantation (FMT) is the most successful one, resolving recurrent CDI in more than 90%

of the cases (van Nood *et al.*, 2013). However, FMT is not an FDA-approved therapy, it requires time to identify a suitable donor and potentiates the transmission of undetected pathogens. These, together with the general patient aversion, represent major impediments to the broad application of FMT. Thus, a major goal in the field is to find specific bacterial species present in the gut that confer protection against CDI. A cocktail of only six cultivable bacterial strains that is able to cure chronic infection in mice was recently identified (Lawley *et al.*, 2012). This information is of paramount importance and can now be used to develop standardized treatment mixtures with the same efficacy as FMT, but increased safety and acceptance.

Pathogenesis

Toxin production

CDI symptoms are primarily caused by the glucosylating toxins TcdA and TcdB (Carter *et al.*, 2010; Carroll and Bartlett, 2011; Shen, 2012; Rupnik *et al.*, 2009). These are classified as large Clostridial toxins (LCTs), due to their large size (205 and 308 KDa, respectively), cytotoxicity and mechanism of action (Just and Gerhard, 2004). Both TcdA and TcdB catalyse the glycosylation, and hence inactivation, of Rho-GTPases (small regulatory proteins of the eukaryotic actin cytoskeleton), leading to disorganization of the cytoskeleton, cell rounding and, ultimately, cell death (Voth and Ballard, 2005; Jank and Aktories, 2008; Carter *et al.*, 2010; Shen, 2012). In addition to its cytotoxic effects, TcdA stimulates cytokine production as well as release of tumour necrosis factor from activated macrophages (Pothoulakis, 2000).

Genes encoding TcdA and TcdB are located in the pathogenicity locus (PaLoc) of *C. difficile* (Figure 1.3A). Besides the *tcdA* and *tcdB* genes, the PaLoc also comprises the *tcdE* gene, which encodes a holin whose pore forming activity allows the release of TcdA and TcdB, and two regulatory genes, *tcdC* and *tcdR.* TcdR is a sigma factor and activates *tcdA*, *tcdB* and *tcdE* transcription, while TcdC is an anti sigma factor that acts by destabilizing the

TcdR-polymerase complex, preventing gene transcription (Figure 1.3A) (Govind and Dupuy, 2012; Dupuy *et al.*, 2006; Dupuy and Matamouros, 2006) (Matamouros *et al.*, 2007; Carter *et al.*, 2011). Structural studies of TcdB have further elucidated the functional domains of these single-chain proteins, and suggest the presence of four structural domains: a C-terminal binding domain containing short combined repetitive oligopeptides (CROPs) for receptor binding, allowing the subsequent receptor-mediated endocytosis; a middle translocation domain, that upon acidification of the endosome suffers conformational changes leading to pore formation and translocation of the N-terminal region of the protein into the cytosol; a cysteine protease domain that is responsible for the auto cleavage of the protein; and finally a catalytic C-terminal domain with glucosyltransferase activity that upon cleavage is released into the host cell cytoplasm (Reinert *et al.*, 2005; Albesa-Jove *et al.*, 2010; Pruitt and Lacy, 2012). The glycoprotein gp96, located at the membrane surface of the colon epithelial cells, is a receptor for TcdA (Na *et al.*, 2008). However, the receptor for TcdB remains unknown.

Some *C. difficile* strains also produce a third, binary toxin known as CDT (*Clostridium difficile* Transferase), coded for by the CDT locus (Figure 1.3B). The CDTb component of CDT binds to the host cells and translocates CDTa, the catalytic component, that ADP-ribosylates actin molecules, leading to depolymerasion of the actin cytoskeleton, cell rounding, loss of fluid and cell death (Schwan *et al.*, 2014; Sundriyal *et al.*, 2010). The lipolysis-stimulated lipoprotein receptor (LSR) is the membrane receptor for CDT uptake by target cells (Papatheodorou *et al.*, 2011). *In vitro*, CDT induces the formation of microtubule protrusions that increase the adherence of *C. difficile* cells (Schwan *et al.*, 2014). Infection by strains producing CDT is associated with higher mortality rates than infection by strains that lack the binary toxin (Bacci *et al.*, 2011). However, strains that produce CDT but do not produce TcdA or TcdB colonize but do not kill hamsters (Geric *et al.*, 2006).

The fact that some pathogenic strains contain truncated versions of toxin A or B and lack CDT clearly suggests that other factors might account for the increased virulence observed for these strains.

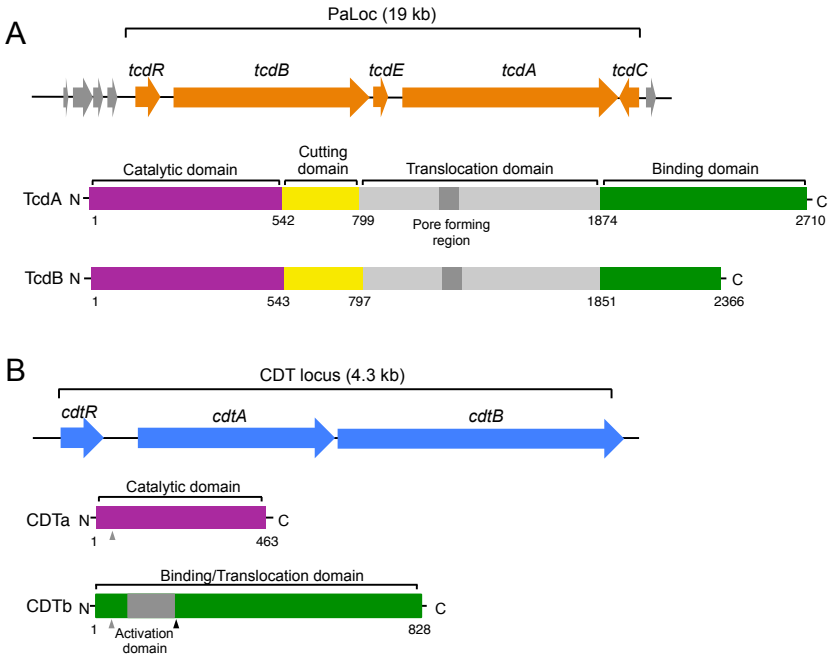


Figure 1.3. *C. difficile* toxins. (A) Schematic representation of the *C. difficile* pathogenicity locus (PaLoc). The structural organization of the two large toxins TcdA and TcdB, encoded by the *tcdA* and *tcdB* genes, respectively, is shown in detail. **(B)** Schematic representation of the *C. difficile* CDT locus, encoding the binary toxin. Structural domains of the CDTa and CDTb polypeptides are shown in detail. In Cdtb, grey arrowheads point to signal peptide cleavage sites, while the black arrowhead points to the proteolytic activation site. Adapted from Rupnik *et al.,* 2009 and Shen, 2012.

Cell surface and adhesion

Colonization is an important step in *C. difficile* pathogenesis. Adhesion factors, such as the S-layer proteins, cover the *C. difficile* cell surface and facilitate binding to the epithelial cells and components of the extracellular matrix fibres, allowing colonization (Spigaglia *et al.*, 2013). Several other proteins located at the cell surface play a role in the adherence to the intestinal epithelial cells. These include the cell wall proteins Cwp66 (Waligora *et al.*, 2001) and Cwp84 (Janoir *et al.*, 2007), the Fbp68

fibronectin-binding protein (Hennequin *et al.*, 2003), the lipoprotein CD0873 (Kovacs-Simon *et al.*, 2014) and the FliC-FliD components of the flagella (Tasteyre *et al.*, 2001). Beyond the role in adhesion, most of these proteins are immunodominant and are able to induce an immune response in the host, either inflammatory or regulatory. For these reasons, both the flagellar antigens FliC and FliD as well as the surface protease Cwp84 have been proposed as potential vaccine candidates (Pechine *et al.*, 2005).

SPORE FORMATION

Two classes of Firmicutes are able to produce bacterial endospores: the *Bacilli*, that includes the extensively studied model organism *Bacillus subtilis*, and the *Clostridia*, to which *C. difficile* belongs. Spores produced by *C. difficile* are crucial for the survival of this anaerobe outside the colonic environment. Spores constitute the most resilient cell form known, they are hard to eradicate and can accumulate and persist in the surfaces of health care institutions for long periods of time, without loosing viability (Nicholson *et al.*, 2000; Carroll and Bartlett, 2011; Maroo and Lamont, 2006; Lawley *et al.*, 2009). Animals infected with a *C. difficile* strain that is unable to sporulate are impaired in *C. difficile* mice-to-mice transmission and persistence within the infected animal (Deakin *et al.*, 2012). These findings highlight the importance of spores in *C. difficile* transmission and recurrence. In addition, higher *in vitro* sporulation rates have been reported for some 027 strains responsible for CDI outbreaks, but a straight correlation could not be established (Merrigan *et al.*, 2010; Akerlund *et al.*, 2008; Burns *et al.*, 2010a). Nevertheless, studies carried in animals show that at least one epidemic 027 strain is more efficiently transmitted to uninfected animals than virulent strains from other ribotypes (Lawley *et al.*, 2012).

Despite the central role played by the spores in *C. difficile* pathogenesis, spore formation has received little attention when in

comparison with other pathogenesis determinants, such as toxin production. In an effort to invert this bias, here we focused our attention on the regulation of sporulation and on the morphology of bacterial spores. The development of new tools to genetically manipulate *C. difficile* allowed us to gain insight on spore formation, structure and composition in this organism (Underwood *et al.*, 2009; Lawley *et al.*, 2009; Permpoonpattana *et al.*, 2011; Permpoonpattana *et al.*, 2013; Barra-Carrasco *et al.*, 2013; Pizarro-Guajardo *et al.*, 2014; Burns *et al.*, 2010b; Putnam *et al.*, 2013; Pettit *et al.*, 2014). Still, most of our knowledge about this process comes from the extensively studied model organism *B. subtilis*.

An overview of the sporulation process

In the model organism *B. subtilis* sporulation is mainly induced by nutrient exhaustion, and is controlled through the phosphorylation of the Spo0A response regulator via an expanded "two-component" signal transduction system called the phosphorelay (Burbulys *et al.*, 1991) (Sonenshein, 2000; Fujita and Losick, 2005). Once phosphorylated, Spo0A induces changes in the transcription of more than 500 genes (Molle *et al.*, 2003). Phosphorylated Spo0A activates the genes coding for the first sporulation-specific regulators, as well as genes required for the asymmetric division of the cell (Rosenbusch *et al.*, 2012; Errington, 2003; Hilbert and Piggot, 2004; Higgins and Dworkin, 2012). Asymmetric division originates two compartments of unequal sizes, the large mother cell compartment and the small forespore compartment (Figure 1.4A) (Piggot and Coote, 1976; Stragier and Losick, 1996; Piggot, 2002; Errington, 2003). Following asymmetric division, the mother cell membrane migrates around the forespore and engulfs it. Once this engulfment process is completed the forespore becomes a free protoplast inside the mother cell. A thick concentric layer of peptidoglycan, the cortex, is then synthesized between the forespore and the mother cell membranes, and a protective concentric

layer of dozens of proteins, named spore coat, is deposited around the cortex (Figure 1.4A). These layers will confer to the newly formed spore the ability to survive to harsh conditions once it is released from the mother cell (Driks, 1999; Henriques and Moran, 2007; McKenney *et al.*, 2012). Though sporulation takes about 8-10 hours, vegetative growth is resumed in only a few minutes, as soon as the spore senses the presence of nutrients in the surrounding medium, following a process called germination (Figure 1.4A).

Regulation of sporulation

Sporulation involves the expression of a large number of genes. In *B. subtilis*, the process is controlled by four RNA polymerase sigma factors that regulate gene expression during the course of spore morphogenesis (Higgins and Dworkin, 2012; Hilbert and Piggot, 2004; Piggot, 2002; Errington, 2003; Stragier and Losick, 1996). The sigma factors of sporulation belong to the σ^{70} family of bacterial sigma factors, which recognize promoters with conserved elements located near -35 and -10 with respect to the transcriptional start site (Doi and Wang, 1986).

The first sigma factors of sporulation, σ^F and σ^E, are activated soon after asymmetric division and control the early stages of spore assembly. At later times, these are replaced by σ^G and σ^K, respectively, which control the final stages of spore formation (Figure 1.4A and B). Therefore, inactivation of both σ^F and σ^E results in blockage of the process soon after asymmetric division, while inactivation of σ^G and σ^K results in a blockage of the process soon after engulfment completion (Piggot and Coote, 1976; Stragier and Losick, 1996; Piggot, 2002). The activity of these sigma factors is not only segregated in time, but also in space. While the activities of σ^F and σ^G are restricted to the forespore compartment, the activities of σ^E and σ^K are restricted to the mother cell compartment (Figure 1.4A and B). Activation of σ^G and σ^K requires the preceding σ^F and σ^E factors to be active in the respective cell compartment (Higgins and Dworkin, 2012; Hilbert and Piggot,

2004; Piggot, 2002; Errington, 2003; Stragier and Losick, 1996). Thus, asymmetric division sets in motion two parallel programs of gene expression, the forespore and the mother cell lines of gene expression.

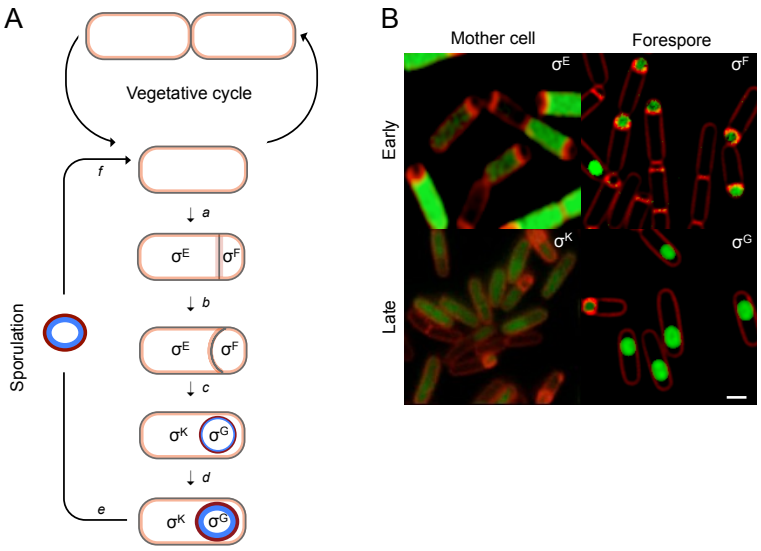


Figure 1.4. Morphological stages and compartmentalized gene expression of *B. subtilis* sporulation. (A) The life cycle of *B. subtilis*. In a nutrient rich medium, the cell grows and divides by symmetric division (vegetative cycle). However, upon starvation, the cell enters in sporulation. The process begins with an asymmetric cell division (*a*). The mother cell membrane then migrates around the forespore (*b*), engulfing it. At the end of this process, the forespore becomes a free protoplast in the mother cell cytoplasm (*c*). The cortex peptidoglycan (blue) and coat (red) layers are then synthesized and deposited around the developing spore (*d*). Upon mother cell lysis, a mature spore is released to the surrounding environment, where it remains in a dormant state (*e*). The spore can then germinate (*f*). The compartment and main periods of activity of the sporulation σ^F, σ^E, σ^G and σ^K sigma factors are indicated. **(B)** Activity of the sporulation σ^F, σ^E, σ^G and σ^K factors observed by fluorescence microscopy. Fluorescent microscopy images of sporulating cells carrying fusions of the promoters of the following genes to GFP: σ^F: *yuiC*; σ^E: *yhaX*; σ^G: *yhcV*; σ^K: *yxeE*. Sporulating cells were stained with the membrane dye FM4-64 (red). Images are the overlap of the green (GFP) and red (FM4-64) channels. Scale bar: 1 μm. The fluorescence images were obtained by Mónica Serrano.

The forespore line of gene expression

The first sporulation-specific sigma factor to be activated is σ^F, in the forespore compartment. σ^F is the product of the *spoIIAC* gene, the third gene of the *spoIIA* operon, whose transcription is controlled by phosphorylated Spo0A and by σ^H (a sigma factor that governs the transition from the late-exponential to the stationary phases) (Wu *et al.*, 1989; Wu *et al.*, 1991;

Piggot, 2002). Although produced in the predivisional cell, σ^F becomes active only after the formation of the asymmetric septum (Figures 1.4 and 1.5) (Gholamhoseinian and Piggot, 1989; Margolis et al., 1991). In the predivisional cell, σ^F is held inactive in a complex with the anti-sigma factor SpoIIAB, the second gene of the *spoIIA* operon (Schmidt *et al.*, 1990; Duncan and Losick, 1993; Min *et al.*, 1993). This inhibition is reversed by the anti-anti-sigma factor SpoIIAA, encoded by the first gene of the same operon (Alper *et al.*, 1994; Diederich *et al.*, 1994; Duncan *et al.*, 1996; Min *et al.*, 1993). SpoIIAA is regulated by its phosphorylation state: it is inactive when phosphorylated by SpoIIAB and active when dephosphorylated by the serine phosphatase SpoIIE (Alper *et al.*, 1994; Min *et al.*, 1993; Arigoni *et al.*, 1996; Duncan *et al.*, 1995; Arigoni *et al.*, 1996). SpoIIE is able to sense the presence of the asymmetric division complex, increasing its phosphatase activity (Barak *et al.*, 1996; Feucht *et al.*, 1996; Feucht *et al.*, 2002). Therefore, upon asymmetric division, SpoIIE dephosphorylates SpoIIAA in the forespore and σ^F is released from its inhibitory complex with SpoIIAB, becoming active (King *et al.*, 1999; Wu *et al.*, 1998). Two mechanisms contribute to the preferential dephosphorylation of SpoIIAA in the forespore compartment. The first is the higher concentration of the SpoIIE phosphatase in this compartment (Guberman *et al.*, 2008; Iber, 2006). The second mechanism is related with transient genetic asymmetry and the position of the *spoIIA* operon near the chromosome terminus. Upon asymmetric cell division only the origin-proximal one third of the chromosome is trapped into the forespore compartment (Wu *et al.*, 1998), leading to the exclusion of the *spoIIA* operon, and thus to depletion of the unstable SpoIIAB protein which is proteolytically unstable, from this compartment (Dworkin and Losick, 2001).

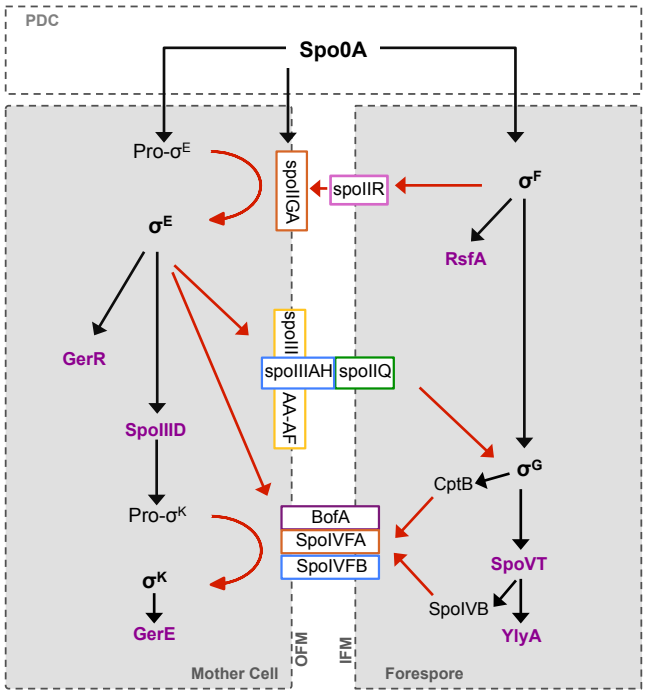


Figure 1.5. The sporulation regulatory network of *B. subtilis*. Temporal progression of sporulation is shown from top to bottom. Boxes represent the forespore and the mother cell, the two cell compartments formed after asymmetric division. The lighter box represents the predivisional cell (PDC), and the darker boxes represent the forespore and the mother cell. The two parallel vertical lines represent the membranes separating the forespore and the mother cell (OFM: outer forespore membrane; IFM: inner forespore membrane). Sigma factors are shown in black and bold, with the precursor proteins indicated as pro-$\sigma^{E/K}$, while transcriptional regulators are shown in purple. Proteins associated with the membranes or located into the intermembrane space (the space between the two parallel broken lines) are in square boxes and illustrated as contacting the parallel vertical line. Black solid arrows indicate activation at the transcriptional level. Red arrows indicate the intercellular signal pathways required for sigma factor activity.

Activation of σ^F leads to the expression of around 50 genes in the forespore (Wang *et al.*, 2006; Steil *et al.*, 2005). Among these are *katX*, a gene involved in spore maturation (Bagyan *et al.*, 1998); a protease encoded by *gpr* (Sussman and Setlow, 1991); *spoIIR*, which is required for the activation of the mother cell sigma factor σ^E (Hofmeister *et al.*, 1995; Karow *et al.*, 1995; Londono-Vallejo and Stragier, 1995); *spoIIQ*, required for engulfment under certain conditions and also for the activation of σ^G (Broder and Pogliano, 2006; Londono-Vallejo *et al.*, 1997); *spoIVB* and *bofC*, both required for the

activation of the late mother cell sigma factor σ^K (Cutting *et al.*, 1990; Gomez and Cutting, 1997) and *rsfA* (Figure 1.5). This last encodes a transcription factor that represses the expression of the σ^F-dependent *spoIIR* gene (Juan Wu and Errington, 2000; Steil *et al.*, 2005; Wang *et al.*, 2006).

Also among the genes transcribed by σ^F is *spoIIIG*, encoding σ^G, which replaces σ^F in the forespore line of gene expression (Figures 1.4 and 1.5) (Sun *et al.*, 1991b; Karmazyn-Campelli *et al.*, 1989). By a mechanism that is not yet fully understood, but that depends on a signal generated in the mother cell under the control of the σ^E factor, transcription of *spoIIIG* is delayed in comparison with the expression of other σ^F-controlled genes (Partridge and Errington, 1993). Therefore, *spoIIIG* expression is only detected towards the end of the engulfment sequence. However its activity is only detected upon engulfment completion and requires signalling from the mother cell compartment (see below; reviewed by Higgins and Dworkin, 2012). Pre-engulfment activity of σ^G in the forespore is prevented by the anti-sigma factor CsfB, which binds to and inactivates σ^G (Karmazyn-Campelli *et al.*, 2008; Serrano *et al.*, 2011). Since σ^G is able to drive its own expression, the inhibition by CsfB is crucial to prevent the generation of a positive feedback loop that causes its levels to increase rapidly in this compartment (Serrano *et al.*, 2011)

σ^G controls the expression of around 100 genes that contribute to proper spore morphogenesis, germination and protection of the spore from DNA damage (Wang *et al.*, 2006; Steil *et al.*, 2005). Examples are the *pdaA* gene, required for modification of the cortex peptidoglycan (Fukushima *et al.*, 2002); the *spoVA, gerA, gerB* and *gerK* operons important for spore germination (Paidhungat, 2002); and the genes coding for the core SASP proteins (Setlow, 1988). σ^G also drives the expression of several regulatory genes, such as *bofC, spoIVB* and *ctpB*, which are all required for the activation of the mother cell sigma factor σ^K (Cutting *et al.*, 1990; Gomez and Cutting, 1997; Campo and Rudner, 2007; Mastny *et al.*, 2013). Additionally, σ^G

transcribes the *spoVT* gene, coding for a transcriptional regulator that modulates the expression of about half of the σ^G-controlled genes, either activating or repressing it (Figure 1.5) (Bagyan *et al.*, 1996; Wang *et al.*, 2006). Inactivation of SpoVT causes an arrest in sporulation at the engulfment stage, most likely because this regulator is required for proper expression of *spoIVB* and thus for σ^K activation (Bagyan *et al.*, 1996). One other regulator that is under the control of σ^G is YlyA (Traag *et al.*, 2013). The gene coding for this protein makes part of the sporulation core (Traag *et al.*, 2013). YlyA binds to RNA polymerase allowing the expression of several σ^G-controlled genes required for efficient spore germination (Traag *et al.*, 2013).

The gene regulatory network governing sporulation in composed by a series of feed forward loops (FFLs). FFLs are circuits that involve two regulatory proteins in which a primary one controls the synthesis of a secondary protein, and together they control the expression of target genes. In type I FFLs, the primary regulator is an activator (Alon et al., 2007). In type I coherent FFLs, both proteins act positively on target genes, whereas in type I incoherent FFLs the first protein acts positively and the second acts negatively (Shen-Orr *et al.*, 2002). The forespore line of gene expression is composed of one coherent and two incoherent FFLs the first acting as a persistence detector and the later generating pulses of gene expression (Figure 1.6) (Traag *et al.*, 2013; Wang *et al.*, 2006). One of the incoherent FFLs is defined by and RsfA, which represses the σ^F-controlled transcription of *spoIIR*; the other is defined by SpoVT whose production is driven by σ^G and then represses a subset of the σ^G-controlled genes (Wang et al., 2006). The coherent FFL is defined by σ^G working with SpoVT in the context of an activator (Traag *et al.*, 2013; Wang *et al.*, 2006).

Lastly, some of the σ^G-dependent genes are also directly controlled by σ^F (Wang *et al.*, 2006). This is explained in part by the similarity in the consensus binding sequences of these two sigma factors, which differ only on the presence of a highly conserved G upstream of the -10 element in the σ^F

consensus motif that is replaced by an X in σ^G promoters (Wang *et al.*, 2006; Amaya *et al.*, 2001; Sun *et al.*, 1991a).

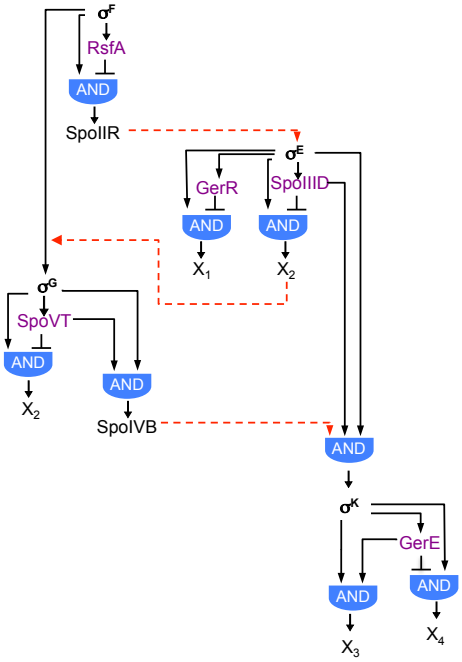


Figure 1.6. Regulatory circuits that govern gene expression during sporulation. The forespore regulatory circuit is composed of three coherent FFLs and one incoherent FFL, while the mother cell is composed of two coherent FFLs linked in series, and three incoherent FFLs. The AND symbols indicate that the FFLs output required the action of both transcription factors in the FFL. *X* denotes the subset of genes on which each FFL acts. Arrowheads represent positive inputs and short, horizontal lines negative inputs. The red broken arrows represent the cell-cell communication pathways that link the mother cell and the forespore. Adapted from Wang *et al.*, 2006.

The mother cell line of gene expression

Unlike the forespore sigma factors σ^F and σ^G, that are synthesized in an active state and held inactive by anti-sigma factors, the mother cell sigma factors σ^E and σ^K are synthesized as inactive pro-proteins and activated by proteolytic cleavage (Hilbert and Piggot, 2004; Higgins and Dworkin, 2012; Errington, 2003).

σ^E, the first sigma factor to be activated in the mother cell compartment, is synthesized in the form of an inactive precursor, pro-σ^E (Figures 1.4 and 1.5) (LaBell *et al.*, 1987). This precursor is encoded by *spoIIGB,* the second gene of the *spoIIG* operon, whose transcription depends on a σ^A-specific promoter that is activated by Spo0A (Kenney *et al.*, 1988). Spo0A activity substantially increases in the mother cell compartment soon after asymmetric division, leading to a corresponding increase in the transcription of the *spoIIG* locus in this compartment (Fujita and Losick, 2002; 2003). The increased transcription of *spoIIG* in the mother cell

following asymmetric division allows increased production of pro-σ^E relative to the forespore (pro-σ^E is also subject to proteolysis in the forespore). Processing of pro-σ^E into σ^E in the mother cell is mediated by the SpoIIGA protease, encoded by the first cistron of the *spoIIG* operon (Stragier *et al.*, 1988). Proteolysis occurs at the outer mother cell membrane and requires the product of *spoIIR* gene, which is transcribed in the forespore under σ^F control (see below) (Figure 1.5) (Fujita and Losick, 2002; Hofmeister *et al.*, 1995; Karow *et al.*, 1995; Londono-Vallejo and Stragier, 1995).

Activation of σ^E leads to expression of 262 genes (Eichenberger *et al.*, 2003; Eichenberger *et al.*, 2004; Steil *et al.*, 2005). A main function of σ^E is to direct the activation of the late forespore sigma factor σ^G as well as both the synthesis and activation of the late mother cell sigma factor σ^K (Figure 1.5). Aside a role in the activation of the σ factors that follow in the sporulation cascade, σ^E is also required for forespore engulfment and to inhibit the formation of a second asymmetric septum, through the activation of three genes: *spoIID*, *spoIIM* and *spoIIP* (Eichenberger *et al.*, 2001; Abanes-De Mello *et al.*, 2002). In addition, σ^E drives the transcription of the coat morphogenetic proteins *spoVM, spoIVA*, *spoVID, safA* and *cotE* (reviewed by McKenney and Eichenberger, 2012). σ^E is also responsible for the synthesis of two transcriptional regulators, SpoIIID and GerR (Figure 1.5)(Kunkel *et al.*, 1989; Juan Wu and Errington, 2000; Eichenberger *et al.*, 2003). GerR acts solely as a repressor of σ^E-activated genes and its inactivation has no major impact on spore resistance and viability, but results in impaired germination (Eichenberger *et al.*, 2004; Hornstra *et al.*, 2005; Cangiano *et al.*, 2010). Unlike GerR, inactivation of SpoIIID results in impaired spore formation and in an inability of sporulating cells to form both the cortex and coat layers (Yoshisue *et al.*, 1995). SpoIIID is a DNA-binding protein that acts together with the σ^E-containing RNA polymerase to repress the transcription of many genes and activate others (Eichenberger *et al.*, 2004, Halberg and Kroos, 1994). Among the genes whose transcription SpoIIID activates are the genes

involved in the appearance of the σ^K, the late mother cell regulator of gene expression (Figure 1.5) (Kroos *et al.*, 1989; Halberg and Kroos, 1994; Eichenberger *et al.*, 2004).

σ^K is the late sigma factor acting in the mother cell (Figures 1.4 and 1.5), and its activity is controlled both at the transcriptional and post-translational level. A particularity of *sigK* is that it is interrupted by a 48 kb prophage-like element called *skin* (for *sigK* intervening sequence) (Stragier *et al.*, 1989). Excision of the *skin* element is mediated by the site-specific recombinase SpoIVCA, resulting in the joining of the two halves of the *sigK* gene (encoded by *spoIIIC* and *spoIVCB*) (Kunkel *et al.*, 1990). Both *spoIVCA* and the intact *sigK* gene are transcribed in the mother cell by σ^E and SpoIIID (Halberg and Kroos, 1994; Kunkel *et al.*, 1990; Eichenberger *et al.*, 2004). SigK is then synthesized, but in the form of an inactive precursor, pro-σ^K (Kroos *et al.*, 1989). Removal of the pro-sequence requires a signal from the forespore compartment (see below) and results in the activation of σ^K (Figure 1.5)(Cutting *et al.*, 1990; Cutting *et al.*, 1991a; Lu *et al.*, 1990).

Activation of σ^K in the mother cell positively impacts on the expression of around 75 genes (Eichenberger *et al.*, 2004; Steil *et al.*, 2005). Most of these genes encode spore coat proteins, as the *cot* genes, or proteins required for spore maturation, as *spoVD* and *spoVK* (Daniel *et al.*, 1994; Fan *et al.*, 1992). Among the members of the σ^K regulon is the gene for the final regulatory protein in the sporulation cascade, GerE, which represses the transcription of over half of the σ^K targets, while switching on 36 others (Figure 1.5) (Cutting *et al.*, 1989; Zheng *et al.*, 1992; Eichenberger *et al.*, 2004).

Members of the σ^K regulon include genes that are under the dual control of σ^E and σ^K. In fact, as in the case of σ^F and σ^G, these two sigma factors are highly similar to each other, as are the promoter sequences that they recognize (Helmann, 2002; Eichenberger *et al.*, 2003; Eichenberger *et al.*, 2004). The motif recognized by σ^K is identical to σ^E in its -10 sequence.

The -35 elements only differ by a single base pair, a cytosine in the case of σ^K that replaces a thymine located in the fourth position of the σ^E element (Eichenberger *et al.*, 2004). Thus, the substitution of a single σ^E residue, the glutamine 217 which as been shown to contact the fourth position of the -35 element, confers to σ^E the ability to recognize σ^K-controlled promoters (Tatti *et al.*, 1995).

The mother-cell line of gene expression is formed by two coherent type-1 FFLs and three incoherent FFLs (Figure 1.6). σ^E activates transcription of SpoIIID, and both σ^E and SpoIIID activate transcription of target genes. Among those are the genes required for the production of σ^K. The second coherent FFL is defined by σ^K, which activates production of GerE. The later then activates transcription of target genes together with σ^K. The three incoherent type-1 FFLs are defined by SpoIIID and GerR, which repress subsets of σ^E-controlled genes, and by GerE, which acts as a repressor of a subset of genes that had been activated by σ^K (Figure 1.6) (Eichenberger *et al.*, 2004).

Cell-cell communication during sporulation

The programs of gene expression within each compartment described above can been seen as hierarchical regulatory cascades in the form of σ^F→RsfA/σ^G→SpoVT/YlyA in the forespore and σ^E→GerR/ SpoIIID→σ^K→GerE in the mother cell (Figure 1.5 and 1.6) (Wang *et al.*, 2006; Eichenberger *et al.*, 2004). Although the two lines of gene expression run in parallel, they are connected at the post-transcriptional level by intercellular signalling pathways (Figure 1.5). A first example is the activation of σ^E in the mother cell, which requires the product of the σ^F-controlled *spoIIR* gene (see above) (Hofmeister *et al.*, 1995; Karow *et al.*, 1995; Londono-Vallejo and Stragier, 1995). The SpoIIR protein contains a signal sequence and is secreted from the prespore to the intermembrane space, where it triggers the SpoIIGA-directed proteolysis of pro-σ^E into σ^E (Hofmeister *et al.*, 1995).

Active σ^E, in turn, is required for the activity of σ^G in the forespore. Activity of σ^G depends on the formation of a channel connecting the mother cell and the forespore compartment, which involves the eight SpoIIIAA-AH proteins, product of the σ^E-controled *spoIIIA* operon, and the forespore SpoIIQ protein, synthesized under σ^F control (Illing and Errington, 1990; Kellner *et al.*, 1996; Londono-Vallejo *et al.*, 1997; Meisner *et al.*, 2008; Camp and Losick, 2008; Doan *et al.*, 2009; Camp and Losick, 2009). The SpoIIIA-SpoIIQ complex allows the mother cell to nurture the forespore upon its isolation from the external medium, and to maintain a metabolic potential that permits macromolecular synthesis, including transcription (Camp and Losick, 2009). σ^G activity in the forespore is then required for the activation of σ^K in the mother cell. Activation of this last sigma factor requires the processing of its precursor protein pro-σ^K by proteolysis, a process that is mediated by the membrane-embedded metallo-protease SpoIVFB (Cutting *et al.*, 1990; Cutting *et al.*, 1991b; Lu *et al.*, 1990; Rudner *et al.*, 1999). SpoIVFB is held inactive in a complex with its inhibitor BofA and a third protein, SpoIVFA, which acts to bring BofA and SpoIVFB together (Ricca *et al.*, 1992; Zhou and Kroos, 2004; Rudner and Losick, 2001; Rudner and Losick, 2002). The concerted action of SpoIVFB and CptB, which are produced in the forespore under σ^G control and secreted to the intramembrane space, is essential to release of SpoIVFB from the inhibition of BofA, thus allowing pro-σ^K processing and σ^K activation (Rudner and Losick, 2002; Dong and Cutting, 2003; Pan *et al.*, 2003; Mastny *et al.*, 2013).

Regulation of sporulation in *Clostridia*

While sporulation has been extensively studied in *B. subtilis* and related species, comparatively little is known about this process in *Clostridia*. Still, many proteins that are known to be essential for sporulation in *Bacilli* are also present in Clostridial genomes. These include the master regulator of sporulation, Spo0A, which is present in the genome of all spore formers

(Galperin *et al.*, 2012; de Hoon *et al.*, 2010; Abecasis *et al.*, 2013; Paredes *et al.*, 2005). Despite the high level of conservation of Spo0A, no true homologues for the phosphorelay components have been found in *Clostridia*. Instead, evidences point to direct phosphorylation of Spo0A by different orphan kinases, such as CD1579 and CD2492 in *C. difficile* or CAC0323, CAC0903 and CAC3319 in *Clostridium acetobutylicum* (Underwood *et al.*, 2009; Steiner *et al.*, 2011; Durre, 2011; Paredes *et al.*, 2005). In the pathogen *C. difficile*, Spo0A controls 321 genes regulating not only sporulation, but also virulence, colonization and metabolism phenotypes, that most likely coordinate the adaptation of this pathogen to its ecological niche (Pettit *et al.*, 2014).

The sporulation sigma factors σ^F, σ^E, σ^G and σ^K, are also conserved in the vast majority of *Clostridia* spore formers (Galperin *et al.*, 2012; Traag *et al.*, 2013; de Hoon *et al.*, 2010; Abecasis *et al.*, 2013; Paredes *et al.*, 2005). However, their sequential activation is not always conserved. For instance, σ^K was shown to control early- instead of late- sporulation events in both *Clostridium botulinum* and *Clostridium perfringens.* In these organisms disruption of *sigK* prevents the formation of the asymmetric septum, and negatively impacts in the transcription of *sigE* in *C. perfringens*, or of *spo0A* and *sigF* in *C. botulinum* (Kirk *et al.*, 2012; Harry *et al.*, 2009). In *C. acetobutylicum* the sequential activation of σ^F, σ^E and σ^G conforms well to the *B. subtilis* paradigm, but σ^K plays a dual role in sporulation, one early, before asymmetric division, and one late, following σ^G activation (Jones *et al.*, 2008; Bi *et al.*, 2011; Jones *et al.*, 2011; Tracy *et al.*, 2011; Al-Hinai *et al.*, 2014). It is therefore conceivable to speculate that the σ^K of *C. perfringens* and *C. botulinum* also play two separate roles in sporulation, as in the case of *C. acetobutylicum*. Besides *B. subtilis*, the *sigK* gene is interrupted by a *skin*-like element in only two other spore formers, *Clostridium tetani* and *C. difficile* (Haraldsen and Sonenshein, 2003; Paredes *et al.*, 2005). In this last, early expression of an intact *sigK* gene negatively impacts on the capacity of at

least some strains of the organism to sporulate (Haraldsen and Sonenshein, 2003).

Most of the regulatory genes that play a role in fine-tuning the mother cell and forespore lines of gene expression as well as the genes involved in the cell-cell signalling pathways are present in *Clostridia*, but are less conserved than the sigma factors (de Hoon *et al.*, 2010; Galperin *et al.*, 2012; Traag *et al.*, 2013; Abecasis *et al.*, 2013; Paredes *et al.*, 2005). Finally, non-regulatory genes controlled by the sigma factor are the least conserved ones, suggesting that in these organisms the morphology and/or composition and functional properties of the spore may differ from *B. subtilis*. Differences in the constitution of the spore are expected, given the range of lifestyles and habitats of the different spore formers.

Spore morphology

Examination of spores by transmission electron microscopy (TEM) show that these consist of three main concentric compartments (the core, cortex and coat) (Figure 1.7) (Driks, 1999; Henriques and Moran, 2000 ; Piggot and Coote, 1976). The inner compartment harbours the DNA and constitutes the spore core. Immediately juxtaposed to the inner spore membrane that delimits the core is a thin layer of peptidoglycan called the germ cell wall, whose composition is similar to the vegetative cell wall (Figure 1.7C). This layer of peptidoglycan will serve as a primer for vegetative cell wall assembly (Cleveland, 1978). Assembled over the germ cell wall is a thick layer of a specialized peptidoglycan, the spore cortex. Surrounding the cortex is a proteinaceous coat layer. The spore coat is organized in two layers: a lamellar inner coat layer and an electrodense outer coat layer (Figure 1.7A and C). An additional external layer, termed exosporium, can also be found around the coat of spores of many Bacillus and Clostridium species (Figure 1.7B and C) (Henriques and Moran, 2007).

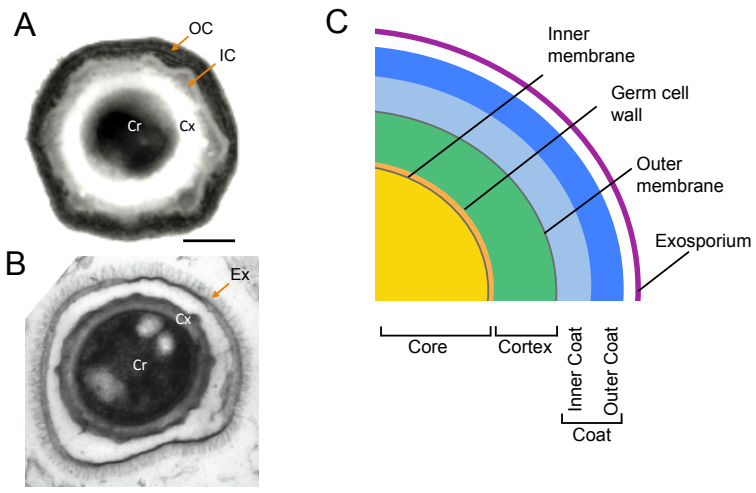


Figure 1.7. Spore ultrastructure. TEM images of *B. subtilis* **(A)** and *B. anthracis* **(B)** spores. The following spore compartments are recognized in (A) and/or in (B): Cr: core, Cx: cortex; IC: Inner Coat; OC: Outer coat; Ex: exosporium. The spore of *B. anthracis* is stained with osmium tetroxide to evidence the hair-nap projections of the exosporium layer that surrounds the spore. Scale bar: 200nm. **(C)** Schematic representation of the spore ultrastructure. The *B. anthracis* spore image is adapted from McKenney *et al.*, 2012.

As for the regulation of sporulation, most of our knowledge about the morphology of the spore is obtained from *B. subtilis* and related species. However, some studies involving the characterization of *C. difficile* spores have been recently conducted and we are starting to unravel the structure, composition and function of the *C. difficile* spore layers (Lawley *et al.*, 2009; Permpoonpattana *et al.*, 2011; Permpoonpattana *et al.*, 2013; Abhyankar *et al.*, 2013; Burns *et al.*, 2010b; Putnam *et al.*, 2013; Pizarro-Guajardo *et al.*, 2014; Barra-Carrasco *et al.*, 2013).

Spore core and cortex

The spore core contains a copy of the bacterial chromosome in a compacted state, determined by the action of SASPs. SASPs are abundant components of the spore proteome, which can bind to the DNA in a non-specific manner, altering its conformation (Setlow, 1988; 1995). By doing so, they provide resistance to a wide range of damaging agents such as dry heat,

UV mutagenesis, nucleases, chemicals and desiccation (Setlow, 2006). The genome of *B. subtilis* encodes 18 SASP, some of which are found almost exclusively in *Bacilli* (Driks, 2002; Galperin *et al.*, 2012). Indeed, one of the three most abundant SASP proteins of *B. subtilis*, SspE, is absent from the *C. difficile* and other Clostridial genomes (Sebaihia *et al.*, 2006; Galperin *et al.*, 2012). Besides the SASPs, the dehydrated state of the spore core and the presence of calcium and dipicolinic acid (DPA) in this compartment also play important roles in spore survival (Setlow, 2006).

The spore cortex is essential for the maintenance of the dehydrated state of the spore core, spore mineralization and dormancy (Gerhardt, 1978). The cortex is made from the same amino acids and sugar constituents that are found in vegetative cell wall peptidoglycan, but it has two unique spore-specific features. The first is a reduced level of cross-linking, which is important for the maintenance of heat resistance (Popham and Setlow, 1993). The second is the presence of δ-muramic acid residues (δ-lactam) (Warth and Strominger, 1969). This last modification plays a specific role during germination because the germination-specific lytic enzymes that degrade the cortex, such as the *B. subtilis* SleB and CwlJ and the *C. difficile* SleC, require the presence of this δ-lactam for substrate recognition (Popham *et al.*, 1996; Atrih and Foster, 1999).

Spore coat

The spore coat is composed of dozens of proteins that provide mechanical integrity and exclude toxic molecules. Moreover, it is this outermost layer that mediates the interaction of the spore with the surrounding environment (Driks, 1999; Henriques and Moran, 2000; Henriques and Moran, 2007; McKenney *et al.*, 2012). Therefore, and giving their distinct habitats and life styles, it is not surprising that only 25% of the *B. subtilis* spore coat proteins have homologues in *C. difficile*. These

conserved proteins cover morphogenetic effectors, several structural proteins and enzymes (Henriques and Moran, 2007).

Proper assembly of the spore coat relies on morphogenetic proteins that direct the assembly of the spore coat around the developing spore and mediate its attachment to the cortex layer. In *B. subtilis* five morphogenetic proteins are present: SpoVM, SpoIVA, SpoVID, SafA and CotE (reviewed by Henriques and Moran, 2007; McKenney *et al.*, 2012). Sporulating cells that lack either SpoVM, SpoIVA or SpoVID are able to synthesize the proteins of the spore coat, which are still able to self-assemble, but the resulting structure is detached form the surface of the developing spore, often forming swirls in the mother cell cytoplasm. The absence of SafA and CotE, in turn, results in the formation of only one of the coat layers: the outer coat in the case of SafA, or the inner coat in the case of CotE (reviewed by Henriques and Moran, 2007; McKenney *et al.*, 2012). In addition to defects on coat assembly, the absence of SpoVM and SpoIVA in *B. subtilis* also impairs cortex formation (Piggot and Coote, 1976; Levin *et al.*, 1993).

Among the conserved coat proteins between *B. subtilis* and *C. difficile* is the morphogenetic protein SpoIVA (Putnam *et al.*, 2013). A functional analogue of SpoVID (named SipL, for SpoIVA interacting protein involved in coat localization) was also identified in *C. difficile*, and despite the high degree of divergence, this protein interacts with SpoIVA, as in *B. subtilis*, but through a different region (Wang *et al.*, 2009; Putnam *et al.*, 2013). Importantly, inactivation of SipL results in a phenotype that is similar to its *B. subtilis* SpoVID counterpart (Putnam *et al.*, 2013; Beall *et al.*, 1993; Wang *et al.*, 2009). However, *C. difficile* cells that lack SpoIVA are able to form the cortex, unlike in *B. subtilis* (Roels *et al.*, 1992).

Other components of the *C. difficile* spore coat layers have been identified by proteomic studies (Permpoonpattana *et al.*, 2011; Abhyankar *et al.*, 2013). Enzymes constitute an important portion of the identified proteins and are responsible for the enzymatic activities present at the surface of the

C. difficile spores (Permpoonpattana *et al.*, 2011; Permpoonpattana *et al.*, 2013; Abhyankar *et al.*, 2013). Some of these enzymes, such as the manganese catalases CotCB, CotD and CotG, the rubrerythrins CD0825, CD1524 and CD2845 and the oxireductases CD0117, CD0176 and CD1623 may play a role in resistance against oxidative stress (Permpoonpattana *et al.*, 2011; Permpoonpattana *et al.*, 2013; Abhyankar *et al.*, 2013).

Still, the vast majority of the spore surface constituents identified have no assigned function nor have homologues in other spore formers and their function remains to be revealed (Permpoonpattana *et al.*, 2011; Abhyankar *et al.*, 2013).

Exosporium

Although dormant, bacterial spores of several species have mechanisms to interact with the environment in a manner that is favourable for their survival. In some species, mostly pathogenic, this interaction is mediated by a balloon-like structure that surrounds the spore coat, termed exosporium (Figure 1.7)(Henriques and Moran, 2007). The exosporium of *B. anthracis* and *B. cereus* is composed of a basal para-crystalline layer and a hair-like nap (Gerhardt and Ribi, 1964; Matz *et al.*, 1970; Kailas *et al.*, 2011). A collagen-like protein called BclA is a key component of this last layer and this protein is involved in the interaction of the spore with host surfaces (Sylvestre *et al.*, 2002; Xue *et al.*, 2011).

The presence and the nature of an exosporium encasing *C. difficile* spores are still under debate. When detected, the organization of such structure by TEM differs from strain to strain, but in general, it appears to be distinct from the *B. anthracis* and *B. cereus* exosporium (Lawley *et al.*, 2009; Joshi *et al.*, 2012; Paredes-Sabja and Sarker, 2012; Escobar-Cortes *et al.*, 2013; Permpoonpattana *et al.*, 2011). Nevertheless, the *C. difficile* genome encodes three homologues of the *B. anthracis bclA*: *bclA1*, *bclA2* and *bclA3*. The proteins encoded by these genes localize to the spore surface, and at

least one, BclA1, has a role in the assembly of the outer spore layers and is required for proper *C. difficile* colonization (Pizarro-Guajardo *et al.*, 2014; Lawley *et al.*, 2009; Abhyankar *et al.*, 2013; Phetcharaburanin *et al.*, 2014). Another spore component that plays a role in exosporium and also in coat assembly is the protein CdeC (for *C. difficile* exosporium cysteine-rich) (Barra-Carrasco *et al.*, 2013). No other components of the *C. difficile* exosporium have been identified yet.

Spore germination and outgrowth

Spores germinate in response to distinct environmental signals such as amino acids, sugars or peptidoglycan fragments (Paidhungat, 2002). The germinant has first to penetrate the coat and cortex layers of the spore before it reaches the germinant receptors located in the inner spore membrane (Paidhungat, 2002). In *B. subtilis* these receptors are encoded by the *gerA*, *gerB* and *gerK* operons, which are expressed in the forespore late in sporulation (Moir, 2006; Moir *et al.*, 2002; Hudson *et al.*, 2001; Paidhungat, 2002). No homologues of these germination receptors are present in the *C. difficile* genome. In fact, *C. difficile* responds to unique germinants, such as bile salts (Wilson *et al.*, 1982; Sorg and Sonenshein, 2008). While the bile salt cholate (CA) induces spore germination, another primary bile salt, chenodeoxycholate (CDCA), has been identified as a potent inhibitor of the process (Sorg and Sonenshein, 2009). Upon antibiotic administration, the metabolism of these two compounds is altered and the CA concentration in the gut becomes much higher than CDCA, triggering spore germination (Giel *et al.*, 2010). Germination induced by the cholate derivative taurocholate is also enhanced in the presence of amino acids, such as glycine and histidine that act as co-germinants (Sorg and Sonenshein, 2008; Wheeldon *et al.*, 2011). While the receptors for these aminoacids remain to be identified, a receptor for taurocholate, CspC, was recently identified. Spores that lack this

receptor show reduced virulence in a hamster model for CDI (Francis *et al.*, 2013).

Once the germinant reaches the receptor, a rapid efflux of cations is observed, together with a large release of calcium-DPA from the core, and its replacement by water (Paidhungat, 2002). The lytic enzymes that hydrolyse the specialized peptidoglycan of the cortex are then activated, leading to full rehydration. In the spore formers, alternative cortex hydrolysis pathways are present: the YpeB-SleB-CwlJ-GerQ (YSCQ) and the Csp-SleC (CS) pathway. While most *Bacilli* seem to rely only on the first, *Clostridia* possess either or both cortex lytic pathways (Paredes-Sabja *et al.*, 2009). In fact, in *C. difficile* members of the two pathways are present but only the Csp-SleC pathway is active (Burns *et al.*, 2010b). While in *C. perfringens* SleC is produced in an inactive pre-pro form that requires proteolytic activation by the Csp germination specific proteases, the *C. difficile* SleC full length protein is active *in vitro* (Gutelius *et al.*, 2014; Paredes-Sabja *et al.*, 2009). Degradation of the cortex peptidoglycan leads to full core hydration, enzyme activity, initiation of metabolism, macromolecular synthesis and spore outgrowth (Horsburgh *et al.*, 2001; Dembek *et al.*, 2013).

AIMS OF THIS WORK

Despite the crucial role played by the spores in the transmission of *C. difficile* and in CDI recurrence, our knowledge about the regulation of spore formation and spore composition in *C. difficile* is still very limited. Thus, the major goals of this work were to elucidate how is the sporulation regulatory network orchestrated in *C. difficile* and how is the *C. difficile* spore composed. We focused our attention on composition of the spore surface layers since they are key mediators of interactions with the host. To accomplish these goals, we have first developed a fluorescent reporter for single cell analysis

in *C. difficile*, since none was available for this anaerobic organism (Chapter 2). To understand how is spore formation regulated in *C. difficile*, we have generated and characterized the mutants for homologues of the four sigma factors that govern spore formation in *B. subtilis*. The employment of the fluorescent reporter was crucial to define the main periods of transcription and activity of each of these sigma factors, as well as the co-dependencies for their expression and activation (Chapter 3). To have a more complete view of the *C. difficile* sporulation regulatory network, we have also addressed the function of two additional regulators of sporulation, SpoIIID and SpoVT (Chapter 4). To achieve our second goal, i.e., to determine the composition of the *C. difficile* spores we have identified all the genes that are under the control of the sporulation sigma factors by whole-genome transcriptomics approaches (Chapter 4). Finally, to improve our knowledge about the composition of the spore surface layers and its interaction with the host, we have focused our attention on the characterization of a major component of the spore surface, the Sp17 protein (Chapter 5).

REFERENCES

Abanes-De Mello, A., Sun, Y. L., Aung, S. & Pogliano, K. (2002). A cytoskeleton-like role for the bacterial cell wall during engulfment of the Bacillus subtilis forespore. *Genes Dev* **16**, 3253-64.

Abecasis, A., Serrano, M., Alves, L., Quintais, L., Pereira-Leal, J. B. & Henriques, A. O. (2013). A genomic signature and the identification of new endosporulation genes. *J Bacteriol* **195**, 2101-2115.

Abhyankar, W., Hossain, A. H., Djajasaputra, A., Permpoonpattana, P., Ter Beek, A., Dekker, H. L., Cutting, S. M., Brul, S., de Koning, L. J. & de Koster, C. G. (2013). In pursuit of protein targets: proteomic characterization of bacterial spore outer layers. *J Proteome Res* **12**, 4507-21.

Akerlund, T., Persson, I., Unemo, M., Noren, T., Svenungsson, B., Wullt, M. & Burman, L. G. (2008). Increased sporulation rate of epidemic Clostridium difficile Type 027/NAP1. *J Clin Microbiol* **46**, 1530-3.

Al-Hinai, M. A., Jones, S. W. & Papoutsakis, E. T. (2014). sigmaK of Clostridium acetobutylicum is the first known sporulation-specific sigma factor with two developmentally separated roles, one early and one late in sporulation. *J Bacteriol* **196**, 287-99.

Albesa-Jove, D., Bertrand, T., Carpenter, E. P., Swain, G. V., Lim, J., Zhang, J., Haire, L. F., Vasisht, N., Braun, V., Lange, A., von Eichel-Streiber, C., Svergun, D. I., Fairweather, N. F. & Brown, K. A. (2010). Four distinct structural domains in Clostridium difficile toxin B visualized using SAXS. *J Mol Biol* **396**, 1260-70.

Alper, S., Duncan, L. & Losick, R. (1994). An adenosine nucleotide switch controlling the activity of a cell type-specific transcription factor in B. subtilis. *Cell* **77**, 195-205.

Amaya, E., Khvorova, A. & Piggot, P. J. (2001). Analysis of promoter recognition in vivo directed by sigma(F) of Bacillus subtilis by using random-sequence oligonucleotides. *J Bacteriol* **183**, 3623-30.

Antunes, W., Serrano, M., Santos, A., Rodrigues, J., Pereira, F. C., Oleastro, M. & Henriques, A. O. (2012). Characterization of Clostridium difficile 027 strains from an outbreak in a Portuguese hospital.: II International Conference on Antimicrobial Research, Lisbon, Portugal.

Arigoni, F., Duncan, L., Alper, S., Losick, R. & Stragier, P. (1996). SpoIIE governs the phosphorylation state of a protein regulating transcription factor sigma F during sporulation in Bacillus subtilis. *Proc Natl Acad Sci U S A* **93**, 3238-42.

Atrih, A. & Foster, S. J. (1999). The role of peptidoglycan structure and structural dynamics during endospore dormancy and germination. *Antonie Van Leeuwenhoek* **75**, 299-307.

Babakhani, F., Bouillaut, L., Gomez, A., Sears, P., Nguyen, L. & Sonenshein, A. L. (2012). Fidaxomicin inhibits spore production in Clostridium difficile. *Clin Infect Dis* **55 Suppl 2**, S162-9.

Bacci, S., Molbak, K., Kjeldsen, M. K. & Olsen, K. E. (2011). Binary toxin and death after Clostridium difficile infection. *Emerg Infect Dis* **17**, 976-82.

Backhed, F., Ding, H., Wang, T., Hooper, L. V., Koh, G. Y., Nagy, A., Semenkovich, C. F. & Gordon, J. I. (2004). The gut microbiota as an environmental factor that regulates fat storage. *Proc Natl Acad Sci U S A* **101**, 15718-23.

Bagyan, I., Casillas-Martinez, L. & Setlow, P. (1998). The katX gene, which codes for the catalase in spores of Bacillus subtilis, is a forespore-specific gene controlled by sigmaF, and KatX is essential for hydrogen peroxide resistance of the germinating spore. *J Bacteriol* **180**, 2057-62.

Bagyan, I., Hobot, J. & Cutting, S. (1996). A compartmentalized regulator of developmental gene expression in Bacillus subtilis. *J Bacteriol* **178**, 4500-7.

Baines, S. D., O'Connor, R., Freeman, J., Fawley, W. N., Harmanus, C., Mastrantonio, P., Kuijper, E. J. & Wilcox, M. H. (2008). Emergence of reduced susceptibility to metronidazole in Clostridium difficile. *J Antimicrob Chemother* **62**, 1046-52.

Barak, I., Behari, J., Olmedo, G., Guzman, P., Brown, D. P., Castro, E., Walker, D., Westpheling, J. & Youngman, P. (1996). Structure and function of the Bacillus SpoIIE protein and its localization to sites of sporulation septum assembly. *Mol Microbiol* **19**, 1047-60.

Barra-Carrasco, J., Olguin-Araneda, V., Plaza-Garrido, A., Miranda-Cardenas, C., Cofre-Araneda, G., Pizarro-Guajardo, M., Sarker, M. R. & Paredes-Sabja, D. (2013). The Clostridium difficile exosporium cysteine (CdeC)-rich protein is required for exosporium morphogenesis and coat assembly. *J Bacteriol* **195**, 3863-75.

Bartlett, J. G. (2010). Clostridium difficile: progress and challenges. *Ann N Y Acad Sci* **1213**, 62-9.

Bartlett, J. G., Moon, N., Chang, T. W., Taylor, N. & Onderdonk, A. B. (1978). Role of Clostridium difficile in antibiotic-associated pseudomembranous colitis. *Gastroenterology* **75**, 778-82.

Bartlett, J. G., Onderdonk, A. B., Cisneros, R. L. & Kasper, D. L. (1977). Clindamycin-associated colitis due to a toxin-producing species of Clostridium in hamsters. *J Infect Dis* **136**, 701-5.

Bauer, M. P., Notermans, D. W., van Benthem, B. H., Brazier, J. S., Wilcox, M. H., Rupnik, M., Monnet, D. L., van Dissel, J. T. & Kuijper, E. J. (2011). Clostridium difficile infection in Europe: a hospital-based survey. *Lancet* **377**, 63-73.

Beall, B., Driks, A., Losick, R. & Moran, C. P., Jr. (1993). Cloning and characterization of a gene required for assembly of the Bacillus subtilis spore coat. *J Bacteriol* **175**, 1705-16.

Berry, D., Stecher, B., Schintlmeister, A., Reichert, J., Brugiroux, S., Wild, B., Wanek, W., Richter, A., Rauch, I., Decker, T., Loy, A. & Wagner, M. (2013). Host-compound foraging by intestinal microbiota revealed by single-cell stable isotope probing. *Proc Natl Acad Sci U S A* **110**, 4720-5.

Bi, C., Jones, S. W., Hess, D. R., Tracy, B. P. & Papoutsakis, E. T. (2011). SpoIIE is necessary for asymmetric division, sporulation, and expression of sigmaF, sigmaE, and sigmaG but does not control solvent production in Clostridium acetobutylicum ATCC 824. *J Bacteriol* **193**, 5130-7.

Brazier, J. S. (2008). Clostridium difficile: from obscurity to superbug. *Br J Biomed Sci* **65**, 39-44.

Brazier, J. S. (2012). Clostridium difficile: the anaerobe that made the grade. *Anaerobe* **18**, 197-9.

Britton, R. A. & Young, V. B. (2012). Interaction between the intestinal microbiota and host in Clostridium difficile colonization resistance. *Trends Microbiol* **20**, 313-9.

Broder, D. H. & Pogliano, K. (2006). Forespore engulfment mediated by a ratchet-like mechanism. *Cell* **126**, 917-28.

Burbulys, D., Trach, K. A. & Hoch, J. A. (1991). Initiation of sporulation in B. subtilis is controlled by a multicomponent phosphorelay. *Cell* **64**, 545-52.

Burnham, C. A. & Carroll, K. C. (2013). Diagnosis of Clostridium difficile infection: an ongoing conundrum for clinicians and for clinical laboratories. *Clin Microbiol Rev* **26**, 604-30.

Burns, D. A., Heap, J. T. & Minton, N. P. (2010a). The diverse sporulation characteristics of Clostridium difficile clinical isolates are not associated with type. *Anaerobe* **16**, 618-22.

Burns, D. A., Heap, J. T. & Minton, N. P. (2010b). SleC is essential for germination of Clostridium difficile spores in nutrient-rich medium supplemented with the bile salt taurocholate. *J Bacteriol* **192**, 657-64.

Camp, A. H. & Losick, R. (2008). A novel pathway of intercellular signalling in Bacillus subtilis involves a protein with similarity to a component of type III secretion channels. *Mol Microbiol* **69**, 402-17.

Camp, A. H. & Losick, R. (2009). A feeding tube model for activation of a cell-specific transcription factor during sporulation in Bacillus subtilis. *Genes Dev* **23**, 1014-24.

Campo, N. & Rudner, D. Z. (2007). SpoIVB and CtpB are both forespore signals in the activation of the sporulation transcription factor sigmaK in Bacillus subtilis. *J Bacteriol* **189**, 6021-7.

Cangiano, G., Mazzone, A., Baccigalupi, L., Isticato, R., Eichenberger, P., De Felice, M. & Ricca, E. (2010). Direct and indirect control of late sporulation genes by GerR of Bacillus subtilis. *J Bacteriol* **192**, 3406-13.

Carroll, K. C. & Bartlett, J. G. (2011). Biology of Clostridium difficile: implications for epidemiology and diagnosis. *Annu Rev Microbiol* **65**, 501-21.

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., Dupuy, B., Rood, J. I. & Lyras, D. (2011). The anti-sigma factor TcdC modulates hypervirulence in an epidemic BI/NAP1/027 clinical isolate of Clostridium difficile. *PLoS Pathog* **7**, e1002317.

Carter, G. P., Rood, J. I. & Lyras, D. (2010). The role of toxin A and toxin B in Clostridium difficile-associated disease: Past and present perspectives. *Gut Microbes* **1**, 58-64.

Cartman, S. T., Heap, J. T., Kuehne, S. A., Cockayne, A. & Minton, N. P. (2010). The emergence of 'hypervirulence' in Clostridium difficile. *Int J Med Microbiol* **300**, 387-95.

Chang, J. Y., Antonopoulos, D. A., Kalra, A., Tonelli, A., Khalife, W. T., Schmidt, T. M. & Young, V. B. (2008). Decreased diversity of the fecal Microbiome in recurrent Clostridium difficile-associated diarrhea. *J Infect Dis* **197**, 435-8.

Clements, A. C., Magalhaes, R. J., Tatem, A. J., Paterson, D. L. & Riley, T. V. (2010). Clostridium difficile PCR ribotype 027: assessing the risks of further worldwide spread. *Lancet Infect Dis* **10**, 395-404.

Cleveland, E. F., Gilvarg, C. (1978). Selective degradation of peptidoglycan from Bacillus megaterium spores during germination. In *Spores VI* (ed. R. N. C. a. H. L. S. P. Gerhardt), pp. 458-464. ASM: Washington D.C.

Cutting, S., Driks, A., Schmidt, R., Kunkel, B. & Losick, R. (1991a). Forespore-specific transcription of a gene in the signal transduction pathway that governs Pro-sigma K processing in Bacillus subtilis. *Genes Dev* **5**, 456-66.

Cutting, S., Oke, V., Driks, A., Losick, R., Lu, S. & Kroos, L. (1990). A forespore checkpoint for mother cell gene expression during development in B. subtilis. *Cell* **62**, 239-50.

Cutting, S., Panzer, S. & Losick, R. (1989). Regulatory studies on the promoter for a gene governing synthesis and assembly of the spore coat in Bacillus subtilis. *J Mol Biol* **207**, 393-404.

Cutting, S., Roels, S. & Losick, R. (1991b). Sporulation operon spoIVF and the characterization of mutations that uncouple mother-cell from forespore gene expression in Bacillus subtilis. *J Mol Biol* **221**, 1237-56.

Daniel, R. A., Drake, S., Buchanan, C. E., Scholle, R. & Errington, J. (1994). The Bacillus subtilis spoVD gene encodes a mother-cell-specific penicillin-binding protein required for spore morphogenesis. *J Mol Biol* **235**, 209-20.

Dawson, L. F., Valiente, E. & Wren, B. W. (2009). Clostridium difficile--a continually evolving and problematic pathogen. *Infect Genet Evol* **9**, 1410-7.

de Hoon, M. J., Eichenberger, P. & Vitkup, D. (2010). Hierarchical evolution of the bacterial sporulation network. *Curr Biol* **20**, R735-45.

Deakin, L. J., Clare, S., Fagan, R. P., Dawson, L. F., Pickard, D. J., West, M. R., Wren, B. W., Fairweather, N. F., Dougan, G. & Lawley, T. D. (2012). The Clostridium difficile spo0A gene is a persistence and transmission factor. *Infect Immun* **80**, 2704-11.

Debast, S. B., Bauer, M. P. & Kuijper, E. J. (2014). European Society of Clinical Microbiology and Infectious Diseases: update of the treatment guidance document for Clostridium difficile infection. *Clin Microbiol Infect* **20 Suppl 2**, 1-26.

Dembek, M., Stabler, R. A., Witney, A. A., Wren, B. W. & Fairweather, N. F. (2013). Transcriptional analysis of temporal gene expression in germinating Clostridium difficile 630 endospores. *PLoS One* **8**, e64011.

Deneve, C., Janoir, C., Poilane, I., Fantinato, C. & Collignon, A. (2009). New trends in Clostridium difficile virulence and pathogenesis. *Int J Antimicrob Agents* **33 Suppl 1**, S24-8.

Diederich, B., Wilkinson, J. F., Magnin, T., Najafi, M., Erringston, J. & Yudkin, M. D. (1994). Role of interactions between SpoIIAA and SpoIIAB in regulating cell-specific transcription factor sigma F of Bacillus subtilis. *Genes Dev* **8**, 2653-63.

DH/HCAI/Infectious disease. Update guidance in the diagnosis and reporting of Clostridium difficile. UK Department of Health. 2012 6 March.

Doan, T., Morlot, C., Meisner, J., Serrano, M., Henriques, A. O., Moran, C. P., Jr. & Rudner, D. Z. (2009). Novel secretion apparatus maintains spore integrity and developmental gene expression in Bacillus subtilis. *PLoS Genet* **5**, e1000566.

Doi, R. H. & Wang, L. F. (1986). Multiple procaryotic ribonucleic acid polymerase sigma factors. *Microbiol Rev* **50**, 227-43.

Dong, T. C. & Cutting, S. M. (2003). SpoIVB-mediated cleavage of SpoIVFA could provide the intercellular signal to activate processing of Pro-sigmaK in Bacillus subtilis. *Mol Microbiol* **49**, 1425-34.

Driks, A. (1999). Bacillus subtilis spore coat. *Microbiol Mol Biol Rev* **63**, 1-20.

Driks, A. (2002). Proteins of the spore core and coat. In *Bacillus subtilis and its closest relatives: from genes to cells.* (ed. S. AL), pp. 526-535. ASM: Washington

Duncan, L., Alper, S., Arigoni, F., Losick, R. & Stragier, P. (1995). Activation of cell-specific transcription by a serine phosphatase at the site of asymmetric division. *Science* **270**, 641-4.

Duncan, L., Alper, S. & Losick, R. (1996). SpoIIAA governs the release of the cell-type specific transcription factor sigma F from its anti-sigma factor SpoIIAB. *J Mol Biol* **260**, 147-64.

Duncan, L. & Losick, R. (1993). SpoIIAB is an anti-sigma factor that binds to and inhibits transcription by regulatory protein sigma F from Bacillus subtilis. *Proc Natl Acad Sci U S A* **90**, 2325-9.

Dupuy, B. & Matamouros, S. (2006). Regulation of toxin and bacteriocin synthesis in Clostridium species by a new subgroup of RNA polymerase sigma-factors. *Res Microbiol* **157**, 201-5.

Dupuy, B., Raffestin, S., Matamouros, S., Mani, N., Popoff, M. R. & Sonenshein, A. L. (2006). Regulation of toxin and bacteriocin gene expression in Clostridium by interchangeable RNA polymerase sigma factors. *Mol Microbiol* **60**, 1044-57.

Durre, P. (2011). Ancestral sporulation initiation. *Mol Microbiol* **80**, 584-7.

Dworkin, J. & Losick, R. (2001). Differential gene expression governed by chromosomal spatial asymmetry. *Cell* **107**, 339-46.

Eichenberger, P., Fawcett, P. & Losick, R. (2001). A three-protein inhibitor of polar septation during sporulation in Bacillus subtilis. *Mol Microbiol* **42**, 1147-62.

Eichenberger, P., Fujita, M., Jensen, S. T., Conlon, E. M., Rudner, D. Z., Wang, S. T., Ferguson, C., Haga, K., Sato, T., Liu, J. S. & Losick, R. (2004). The program of gene transcription for a single differentiating cell type during sporulation in Bacillus subtilis. *PLoS Biol* **2**, e328.

Eichenberger, P., Jensen, S. T., Conlon, E. M., van Ooij, C., Silvaggi, J., Gonzalez-Pastor, J. E., Fujita, M., Ben-Yehuda, S., Stragier, P., Liu, J. S. & Losick, R. (2003). The sigmaE regulon and the identification of additional sporulation genes in Bacillus subtilis. *J Mol Biol* **327**, 945-72.

Errington, J. (2003). Regulation of endospore formation in Bacillus subtilis. *Nat Rev Microbiol* **1**, 117-26.

Escobar-Cortes, K., Barra-Carrasco, J. & Paredes-Sabja, D. (2013). Proteases and sonication specifically remove the exosporium layer of spores of Clostridium difficile strain 630. *J Microbiol Methods* **93**, 25-31.

Fan, N., Cutting, S. & Losick, R. (1992). Characterization of the Bacillus subtilis sporulation gene spoVK. *J Bacteriol* **174**, 1053-4.

Feucht, A., Abbotts, L. & Errington, J. (2002). The cell differentiation protein SpoIIE contains a regulatory site that controls its phosphatase activity in response to asymmetric septation. *Mol Microbiol* **45**, 1119-30.

Feucht, A., Magnin, T., Yudkin, M. D. & Errington, J. (1996). Bifunctional protein required for asymmetric cell division and cell-specific transcription in Bacillus subtilis. *Genes Dev* **10**, 794-803.

Francis, M. B., Allen, C. A., Shrestha, R. & Sorg, J. A. (2013). Bile acid recognition by the Clostridium difficile germinant receptor, CspC, is important for establishing infection. *PLoS Pathog* **9**, e1003356.

Freeman, J., Bauer, M. P., Baines, S. D., Corver, J., Fawley, W. N., Goorhuis, B., Kuijper, E. J. & Wilcox, M. H. (2010). The changing epidemiology of Clostridium difficile infections. *Clin Microbiol Rev* **23**, 529-49.

Fujita, M. & Losick, R. (2002). An investigation into the compartmentalization of the sporulation transcription factor sigmaE in Bacillus subtilis. *Mol Microbiol* **43**, 27-38.

Fujita, M. & Losick, R. (2003). The master regulator for entry into sporulation in Bacillus subtilis becomes a cell-specific transcription factor after asymmetric division. *Genes Dev* **17**, 1166-74.

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**, 2236-44.

Fukushima, T., Yamamoto, H., Atrih, A., Foster, S. J. & Sekiguchi, J. (2002). A polysaccharide deacetylase gene (pdaA) is required for germination and for production of muramic delta-lactam residues in the spore cortex of Bacillus subtilis. *J Bacteriol* **184**, 6007-15.

Galperin, M. Y., Mekhedov, S. L., Puigbo, P., Smirnov, S., Wolf, Y. I. & Rigden, D. J. (2012). Genomic determinants of sporulation in Bacilli and Clostridia: towards the minimal set of sporulation-specific genes. *Environ Microbiol* **14**, 2870-90.

Gerhardt, P., Murrel, W. G. (1978). Basis and mechanisms of spore resistance. In *Spores VII* (ed. G. H. C. a. J. C. Vary). ASM: Washington.

Gerhardt, P. & Ribi, E. (1964). ULTRASTRUCTURE OF THE EXOSPORIUM ENVELOPING SPORES OF BACILLUS CEREUS. *J Bacteriol* **88**, 1774-89.

Geric, B., Carman, R. J., Rupnik, M., Genheimer, C. W., Sambol, S. P., Lyerly, D. M., Gerding, D. N. & Johnson, S. (2006). Binary toxin-producing, large clostridial toxin-negative Clostridium difficile strains are enterotoxic but do not cause disease in hamsters. *J Infect Dis* **193**, 1143-50.

Gholamhoseinian, A. & Piggot, P. J. (1989). Timing of spoII gene expression relative to septum formation during sporulation of Bacillus subtilis. *J Bacteriol* **171**, 5747-9.

Giel, J. L., Sorg, J. A., Sonenshein, A. L. & Zhu, J. (2010). Metabolism of bile salts in mice influences spore germination in Clostridium difficile. *PLoS One* **5**, e8740.

Goldenberg, S. D., Dieringer, T. & French, G. L. (2010). Detection of toxigenic Clostridium difficile in diarrheal stools by rapid real-time polymerase chain reaction. *Diagn Microbiol Infect Dis* **67**, 304-7.

Gomez, M. & Cutting, S. M. (1997). BofC encodes a putative forespore regulator of the Bacillus subtilis sigma K checkpoint. *Microbiology* **143 (Pt 1)**, 157-70.

Govind, R. & Dupuy, B. (2012). Secretion of Clostridium difficile toxins A and B requires the holin-like protein TcdE. *PLoS Pathog* **8**, e1002727.

Guberman, J. M., Fay, A., Dworkin, J., Wingreen, N. S. & Gitai, Z. (2008). PSICIC: noise and asymmetry in bacterial division revealed by computational image analysis at sub-pixel resolution. *PLoS Comput Biol* **4**, e1000233.

Gutelius, D., Hokeness, K., Logan, S. M. & Reid, C. W. (2014). Functional analysis of SleC from Clostridium difficile: an essential lytic transglycosylase involved in spore germination. *Microbiology* **160**, 209-16.

Halberg, R. & Kroos, L. (1994). Sporulation regulatory protein SpoIIID from Bacillus subtilis activates and represses transcription by both mother-cell-specific forms of RNA polymerase. *J Mol Biol* **243**, 425-36.

Hall, I. C. and O'Toole, E. (1935). Intestinal flora in newborn infants: with a description of a new pathogenic anaerobe: Bacillus difficilis. *Am. J. Dis. Child.* **49**, 390.

Haraldsen, J. D. & Sonenshein, A. L. (2003). Efficient sporulation in Clostridium difficile requires disruption of the sigmaK gene. *Mol Microbiol* **48**, 811-21.

Harry, K. H., Zhou, R., Kroos, L. & Melville, S. B. (2009). Sporulation and enterotoxin (CPE) synthesis are controlled by the sporulation-specific sigma factors SigE and SigK in Clostridium perfringens. *J Bacteriol* **191**, 2728-42.

He, M., Miyajima, F., Roberts, P., Ellison, L., Pickard, D. J., Martin, M. J., Connor, T. R., Harris, S. R., Fairley, D., Bamford, K. B., D'Arc, S., Brazier, J., Brown, D., Coia, J. E., Douce, G., Gerding, D., Kim, H. J., Koh, T. H., Kato, H., Senoh, M., Louie, T., Michell, S., Butt, E., Peacock, S. J., Brown, N. M., Riley, T., Songer, G., Wilcox, M., Pirmohamed, M., Kuijper, E., Hawkey, P., Wren, B. W., Dougan, G., Parkhill, J. & Lawley, T. D. (2013). Emergence and global spread of epidemic healthcare-associated Clostridium difficile. *Nat Genet* **45**, 109-13.

Helmann, J. D., Moran Jr., C. P. (2002). RNA ploymerase and sigma factors. In *Bacillus subtilis and its closest relatives: from genes to cells.* (ed. S. AL), pp. 289-312. ASM: Washington DC.

Hennequin, C., Janoir, C., Barc, M. C., Collignon, A. & Karjalainen, T. (2003). Identification and characterization of a fibronectin-binding protein from Clostridium difficile. *Microbiology* **149**, 2779-87.

Henriques, A. O. & Moran, C. P., Jr. (2000). Structure and assembly of the bacterial endospore coat. *Methods* **20**, 95-110.

Henriques, A. O. & Moran, C. P., Jr. (2007). Structure, assembly, and function of the spore surface layers. *Annu Rev Microbiol* **61**, 555-88.

Hensgens, M. P., Keessen, E. C., Squire, M. M., Riley, T. V., Koene, M. G., de Boer, E., Lipman, L. J. & Kuijper, E. J. (2012). Clostridium difficile infection in the community: a zoonotic disease? *Clin Microbiol Infect* **18**, 635-45.

Higgins, D. & Dworkin, J. (2012). Recent progress in Bacillus subtilis sporulation. *FEMS Microbiol Rev* **36**, 131-48.

Hilbert, D. W. & Piggot, P. J. (2004). Compartmentalization of gene expression during Bacillus subtilis spore formation. *Microbiol Mol Biol Rev* **68**, 234-62.

Hofmeister, A. E., Londono-Vallejo, A., Harry, E., Stragier, P. & Losick, R. (1995). Extracellular signal protein triggering the proteolytic activation of a developmental transcription factor in B. subtilis. *Cell* **83**, 219-26.

Hornstra, L. M., de Vries, Y. P., de Vos, W. M., Abee, T. & Wells-Bennik, M. H. (2005). gerR, a novel ger operon involved in L-alanine- and inosine-initiated germination of Bacillus cereus ATCC 14579. *Appl Environ Microbiol* **71**, 774-81.

Horsburgh, M. J., Thackray, P. D. & Moir, A. (2001). Transcriptional responses during outgrowth of Bacillus subtilis endospores. *Microbiology* **147**, 2933-41.

Hudson, K. D., Corfe, B. M., Kemp, E. H., Feavers, I. M., Coote, P. J. & Moir, A. (2001). Localization of GerAA and GerAC germination proteins in the Bacillus subtilis spore. *J Bacteriol* **183**, 4317-22.

Hussain, H. A., Roberts, A. P. & Mullany, P. (2005). Generation of an erythromycin-sensitive derivative of Clostridium difficile strain 630 (630Deltaerm) and demonstration that the conjugative transposon Tn916DeltaE enters the genome of this strain at multiple sites. *J Med Microbiol* **54**, 137-41.

Iber, D. (2006). A computational analysis of the impact of the transient genetic imbalance on compartmentalized gene expression during sporulation in Bacillus subtilis. *J Mol Biol* **360**, 15-20.

Illing, N. & Errington, J. (1990). The spoIIIA locus is not a major determinant of prespore-specific gene expression during sporulation in Bacillus subtilis. *J Bacteriol* **172**, 6930-6.

Jank, T. & Aktories, K. (2008). Structure and mode of action of clostridial glucosylating toxins: the ABCD model. *Trends Microbiol* **16**, 222-9.

Janoir, C., Pechine, S., Grosdidier, C. & Collignon, A. (2007). Cwp84, a surface-associated protein of Clostridium difficile, is a cysteine protease with degrading activity on extracellular matrix proteins. *J Bacteriol* **189**, 7174-80.

Jones, A. M., Kuijper, E. J. & Wilcox, M. H. (2013). Clostridium difficile: a European perspective. *J Infect* **66**, 115-28.

Jones, S. W., Paredes, C. J., Tracy, B., Cheng, N., Sillers, R., Senger, R. S. & Papoutsakis, E. T. (2008). The transcriptional program underlying the physiology of clostridial sporulation. *Genome Biol* **9**, R114.

Jones, S. W., Tracy, B. P., Gaida, S. M. & Papoutsakis, E. T. (2011). Inactivation of sigmaF in Clostridium acetobutylicum ATCC 824 blocks sporulation prior to asymmetric division and abolishes sigmaE and sigmaG protein expression but does not block solvent formation. *J Bacteriol* **193**, 2429-40.

Joshi, L. T., Phillips, D. S., Williams, C. F., Alyousef, A. & Baillie, L. (2012). Contribution of spores to the ability of Clostridium difficile to adhere to surfaces. *Appl Environ Microbiol* **78**, 7671-9.

Juan Wu, L. & Errington, J. (2000). Identification and characterization of a new prespore-specific regulatory gene, rsfA, of Bacillus subtilis. *J Bacteriol* **182**, 418-24.

Just, I. & Gerhard, R. (2004). Large clostridial cytotoxins. *Rev Physiol Biochem Pharmacol* **152**, 23-47.

Kailas, L., Terry, C., Abbott, N., Taylor, R., Mullin, N., Tzokov, S. B., Todd, S. J., Wallace, B. A., Hobbs, J. K., Moir, A. & Bullough, P. A. (2011). Surface architecture of endospores of the Bacillus cereus/anthracis/thuringiensis family at the subnanometer scale. *Proc Natl Acad Sci U S A* **108**, 16014-9.

Karmazyn-Campelli, C., Bonamy, C., Savelli, B. & Stragier, P. (1989). Tandem genes encoding sigma-factors for consecutive steps of development in Bacillus subtilis. *Genes Dev* **3**, 150-7.

Karmazyn-Campelli, C., Rhayat, L., Carballido-Lopez, R., Duperrier, S., Frandsen, N. & Stragier, P. (2008). How the early sporulation sigma factor sigmaF delays the switch to late development in Bacillus subtilis. *Mol Microbiol* **67**, 1169-80.

Karow, M. L., Glaser, P. & Piggot, P. J. (1995). Identification of a gene, spoIIR, that links the activation of sigma E to the transcriptional activity of sigma F during sporulation in Bacillus subtilis. *Proc Natl Acad Sci U S A* **92**, 2012-6.

Kellner, E. M., Decatur, A. & Moran, C. P., Jr. (1996). Two-stage regulation of an anti-sigma factor determines developmental fate during bacterial endospore formation. *Mol Microbiol* **21**, 913-24.

Kelly, C. P. & LaMont, J. T. (2008). Clostridium difficile--more difficult than ever. *N Engl J Med* **359**, 1932-40.

Kenney, T. J., Kirchman, P. A. & Moran, C. P., Jr. (1988). Gene encoding sigma E is transcribed from a sigma A-like promoter in Bacillus subtilis. *J Bacteriol* **170**, 3058-64.

King, N., Dreesen, O., Stragier, P., Pogliano, K. & Losick, R. (1999). Septation, dephosphorylation, and the activation of sigmaF during sporulation in Bacillus subtilis. *Genes Dev* **13**, 1156-67.

Kirk, D. G., Dahlsten, E., Zhang, Z., Korkeala, H. & Lindstrom, M. (2012). Involvement of Clostridium botulinum ATCC 3502 sigma factor K in early-stage sporulation. *Appl Environ Microbiol* **78**, 4590-6.

Kovacs-Simon, A., Leuzzi, R., Kasendra, M., Minton, N., Titball, R. W. & Michell, S. L. (2014). Lipoprotein CD0873 Is a Novel Adhesin of Clostridium difficile. *J Infect Dis*.

Kroos, L., Kunkel, B. & Losick, R. (1989). Switch protein alters specificity of RNA polymerase containing a compartment-specific sigma factor. *Science* **243**, 526-9.

Kuijper, E. J., Barbut, F., Brazier, J. S., Kleinkauf, N., Eckmanns, T., Lambert, M. L., Drudy, D., Fitzpatrick, F., Wiuff, C., Brown, D. J., Coia, J. E., Pituch, H., Reichert, P., Even, J., Mossong, J., Widmer, A. F., Olsen, K. E., Allerberger, F., Notermans, D. W., Delmee, M., Coignard, B., Wilcox, M., Patel, B., Frei, R., Nagy, E., Bouza, E., Marin, M., Akerlund, T., Virolainen-Julkunen, A., Lyytikainen, O., Kotila, S.,

Ingebretsen, A., Smyth, B., Rooney, P., Poxton, I. R. & Monnet, D. L. (2008). Update of Clostridium difficile infection due to PCR ribotype 027 in Europe, 2008. *Euro Surveill* **13**.

Kuipers, E. J. & Surawicz, C. M. (2008). Clostridium difficile infection. *Lancet* **371**, 1486-8.

Kunkel, B., Kroos, L., Poth, H., Youngman, P. & Losick, R. (1989). Temporal and spatial control of the mother-cell regulatory gene spoIIID of Bacillus subtilis. *Genes Dev* **3**, 1735-44.

Kunkel, B., Losick, R. & Stragier, P. (1990). The Bacillus subtilis gene for the development transcription factor sigma K is generated by excision of a dispensable DNA element containing a sporulation recombinase gene. *Genes Dev* **4**, 525-35.

LaBell, T. L., Trempy, J. E. & Haldenwang, W. G. (1987). Sporulation-specific sigma factor sigma 29 of Bacillus subtilis is synthesized from a precursor protein, P31. *Proc Natl Acad Sci U S A* **84**, 1784-8.

Larson, H. E., Price, A. B., Honour, P. & Borriello, S. P. (1978). Clostridium difficile and the aetiology of pseudomembranous colitis. *Lancet* **1**, 1063-6.

Lawley, T. D., Clare, S., Walker, A. W., Stares, M. D., Connor, T. R., Raisen, C., Goulding, D., Rad, R., Schreiber, F., Brandt, C., Deakin, L. J., Pickard, D. J., Duncan, S. H., Flint, H. J., Clark, T. G., Parkhill, J. & Dougan, G. (2012). Targeted restoration of the intestinal microbiota with a simple, defined bacteriotherapy resolves relapsing Clostridium difficile disease in mice. *PLoS Pathog* **8**, e1002995.

Lawley, T. D., Croucher, N. J., Yu, L., Clare, S., Sebaihia, M., Goulding, D., Pickard, D. J., Parkhill, J., Choudhary, J. & Dougan, G. (2009). Proteomic and genomic characterization of highly infectious Clostridium difficile 630 spores. *J Bacteriol* **191**, 5377-86.

Levin, P. A., Fan, N., Ricca, E., Driks, A., Losick, R. & Cutting, S. (1993). An unusually small gene required for sporulation by Bacillus subtilis. *Mol Microbiol* **9**, 761-71.

Liao, C. H., Ko, W. C., Lu, J. J. & Hsueh, P. R. (2012). Characterizations of clinical isolates of clostridium difficile by toxin genotypes and by susceptibility to 12 antimicrobial agents, including fidaxomicin (OPT-80) and rifaximin: a multicenter study in Taiwan. *Antimicrob Agents Chemother* **56**, 3943-9.

Littman, D. R. & Pamer, E. G. (2011). Role of the commensal microbiota in normal and pathogenic host immune responses. *Cell Host Microbe* **10**, 311-23.

Londono-Vallejo, J. A., Frehel, C. & Stragier, P. (1997). SpoIIQ, a forespore-expressed gene required for engulfment in Bacillus subtilis. *Mol Microbiol* **24**, 29-39.

Londono-Vallejo, J. A. & Stragier, P. (1995). Cell-cell signaling pathway activating a developmental transcription factor in Bacillus subtilis. *Genes Dev* **9**, 503-8.

Louie, T. J., Miller, M. A., Mullane, K. M., Weiss, K., Lentnek, A., Golan, Y., Gorbach, S., Sears, P. & Shue, Y. K. (2011). Fidaxomicin versus vancomycin for Clostridium difficile infection. *N Engl J Med* **364**, 422-31.

Lozupone, C. A., Stombaugh, J. I., Gordon, J. I., Jansson, J. K. & Knight, R. (2012). Diversity, stability and resilience of the human gut microbiota. *Nature* **489**, 220-30.

Lu, S., Halberg, R. & Kroos, L. (1990). Processing of the mother-cell sigma factor, sigma K, may depend on events occurring in the forespore during Bacillus subtilis development. *Proc Natl Acad Sci U S A* **87**, 9722-6.

Lucado, J., Gould, C. & Elixhauser, A. (2009). Clostridium Difficile Infections (CDI) in Hospital Stays, 2009: Statistical Brief. In *Healthcare Cost and Utilization Project (HCUP) Statistical Briefs*. Agency for Health Care Policy and Research (US): Rockville (MD).

Lyras, D., Adams, V., Lucet, I. & Rood, J. I. (2004). The large resolvase TnpX is the only transposon-encoded protein required for transposition of the Tn4451/3 family of integrative mobilizable elements. *Mol Microbiol* **51**, 1787-800.

Margolis, P., Driks, A. & Losick, R. (1991). Establishment of cell type by compartmentalized activation of a transcription factor. *Science* **254**, 562-5.

Maroo, S. & Lamont, J. T. (2006). Recurrent clostridium difficile. *Gastroenterology* **130**, 1311-6.

Mastny, M., Heuck, A., Kurzbauer, R., Heiduk, A., Boisguerin, P., Volkmer, R., Ehrmann, M., Rodrigues, C. D., Rudner, D. Z. & Clausen, T. (2013). CtpB assembles a gated protease tunnel regulating cell-cell signaling during spore formation in Bacillus subtilis. *Cell* **155**, 647-58.

Matamouros, S., England, P. & Dupuy, B. (2007). Clostridium difficile toxin expression is inhibited by the novel regulator TcdC. *Mol Microbiol* **64**, 1274-88.

Matz, L. L., Beaman, T. C. & Gerhardt, P. (1970). Chemical composition of exosporium from spores of Bacillus cereus. *J Bacteriol* **101**, 196-201.

McKenney, P. T., Driks, A. & Eichenberger, P. (2012). The Bacillus subtilis endospore: assembly and functions of the multilayered coat. *Nat Rev Microbiol* **11**, 33-44.

McKenney, P. T. & Eichenberger, P. (2012). Dynamics of spore coat morphogenesis in Bacillus subtilis. *Mol Microbiol* **83**, 245-60.

Meisner, J., Wang, X., Serrano, M., Henriques, A. O. & Moran, C. P., Jr. (2008). A channel connecting the mother cell and forespore during bacterial endospore formation. *Proc Natl Acad Sci U S A* **105**, 15100-5.

Merrigan, M., Venugopal, A., Mallozzi, M., Roxas, B., Viswanathan, V. K., Johnson, S., Gerding, D. N. & Vedantam, G. (2010). Human hypervirulent Clostridium difficile

strains exhibit increased sporulation as well as robust toxin production. *J Bacteriol* **192**, 4904-11.

Min, K. T., Hilditch, C. M., Diederich, B., Errington, J. & Yudkin, M. D. (1993). Sigma F, the first compartment-specific transcription factor of B. subtilis, is regulated by an anti-sigma factor that is also a protein kinase. *Cell* **74**, 735-42.

Moir, A. (2006). How do spores germinate? *J Appl Microbiol* **101**, 526-30.

Moir, A., Corfe, B. M. & Behravan, J. (2002). Spore germination. *Cell Mol Life Sci* **59**, 403-9.

Molle, V., Fujita, M., Jensen, S. T., Eichenberger, P., Gonzalez-Pastor, J. E., Liu, J. S. & Losick, R. (2003). The Spo0A regulon of Bacillus subtilis. *Mol Microbiol* **50**, 1683-701.

Moura, I., Spigaglia, P., Barbanti, F. & Mastrantonio, P. (2013). Analysis of metronidazole susceptibility in different Clostridium difficile PCR ribotypes. *J Antimicrob Chemother* **68**, 362-5.

Na, X., Kim, H., Moyer, M. P., Pothoulakis, C. & LaMont, J. T. (2008). gp96 is a human colonocyte plasma membrane binding protein for Clostridium difficile toxin A. *Infect Immun* **76**, 2862-71.

Nagaro, K. J., Phillips, S. T., Cheknis, A. K., Sambol, S. P., Zukowski, W. E., Johnson, S. & Gerding, D. N. (2013). Nontoxigenic Clostridium difficile protects hamsters against challenge with historic and epidemic strains of toxigenic BI/NAP1/027 C. difficile. *Antimicrob Agents Chemother* **57**, 5266-70.

Ng, K. M., Ferreyra, J. A., Higginbottom, S. K., Lynch, J. B., Kashyap, P. C., Gopinath, S., Naidu, N., Choudhury, B., Weimer, B. C., Monack, D. M. & Sonnenburg, J. L. (2013). Microbiota-liberated host sugars facilitate post-antibiotic expansion of enteric pathogens. *Nature* **502**, 96-9.

Nicholson, W. L., Munakata, N., Horneck, G., Melosh, H. J. & Setlow, P. (2000). Resistance of Bacillus endospores to extreme terrestrial and extraterrestrial environments. *Microbiol Mol Biol Rev* **64**, 548-72.

O'Connor, J. R., Johnson, S. & Gerding, D. N. (2009). Clostridium difficile infection caused by the epidemic BI/NAP1/027 strain. *Gastroenterology* **136**, 1913-24.

Paidhungat, M., Setlow, P. (2002). Spore germination and outgrowth. In *Bacillus subtilis and its closest relatives: from genes to cells.* (ed. S. AL), pp. 537-548. ASM: Washington

Pan, Q., Losick, R. & Rudner, D. Z. (2003). A second PDZ-containing serine protease contributes to activation of the sporulation transcription factor sigmaK in Bacillus subtilis. *J Bacteriol* **185**, 6051-6.

Papatheodorou, P., Carette, J. E., Bell, G. W., Schwan, C., Guttenberg, G., Brummelkamp, T. R. & Aktories, K. (2011). Lipolysis-stimulated lipoprotein receptor (LSR) is the host receptor for the binary toxin Clostridium difficile transferase (CDT). *Proc Natl Acad Sci U S A* **108**, 16422-7.

Paredes, C. J., Alsaker, K. V. & Papoutsakis, E. T. (2005). A comparative genomic view of clostridial sporulation and physiology. *Nat Rev Microbiol* **3**, 969-78.

Paredes-Sabja, D. & Sarker, M. R. (2012). Adherence of Clostridium difficile spores to Caco-2 cells in culture. *J Med Microbiol* **61**, 1208-18.

Paredes-Sabja, D., Setlow, P. & Sarker, M. R. (2009). The protease CspB is essential for initiation of cortex hydrolysis and dipicolinic acid (DPA) release during germination of spores of Clostridium perfringens type A food poisoning isolates. *Microbiology* **155**, 3464-72.

Partridge, S. R. & Errington, J. (1993). The importance of morphological events and intercellular interactions in the regulation of prespore-specific gene expression during sporulation in Bacillus subtilis. *Mol Microbiol* **8**, 945-55.

Pechine, S., Gleizes, A., Janoir, C., Gorges-Kergot, R., Barc, M. C., Delmee, M. & Collignon, A. (2005). Immunological properties of surface proteins of Clostridium difficile. *J Med Microbiol* **54**, 193-6.

Permpoonpattana, P., Phetcharaburanin, J., Mikelsone, A., Dembek, M., Tan, S., Brisson, M. C., La Ragione, R., Brisson, A. R., Fairweather, N., Hong, H. A. & Cutting, S. M. (2013). Functional characterization of Clostridium difficile spore coat proteins. *J Bacteriol* **195**, 1492-503.

Permpoonpattana, P., Tolls, E. H., Nadem, R., Tan, S., Brisson, A. & Cutting, S. M. (2011). Surface layers of Clostridium difficile endospores. *J Bacteriol* **193**, 6461-70.

Pettit, L. J., Browne, H. P., Yu, L., Smits, W. K., Fagan, R. P., Barquist, L., Martin, M. J., Goulding, D., Duncan, S. H., Flint, H. J., Dougan, G., Choudhary, J. S. & Lawley, T. D. (2014). Functional genomics reveals that Clostridium difficile Spo0A coordinates sporulation, virulence and metabolism. *BMC Genomics* **15**, 160.

Piggot, P. J. & Coote, J. G. (1976). Genetic aspects of bacterial endospore formation. *Bacteriol Rev* **40**, 908-62.

Piggot, P. J., Losick, R. (2002). Sporulation genes and intercompartmental regulation. In *Bacillus subtilis and its closest relatives: from genes to cells.* (ed. A. L. Sonenshein), pp. 483-518. ASM: Washington.

Pizarro-Guajardo, M., Olguin-Araneda, V., Barra-Carrasco, J., Brito-Silva, C., Sarker, M. R. & Paredes-Sabja, D. (2014). Characterization of the collagen-like exosporium protein, BclA1, of Clostridium difficile spores. *Anaerobe* **25**, 18-30.

Popham, D. L., Helin, J., Costello, C. E. & Setlow, P. (1996). Muramic lactam in peptidoglycan of Bacillus subtilis spores is required for spore outgrowth but not for spore dehydration or heat resistance. *Proc Natl Acad Sci U S A* **93**, 15405-10.

Popham, D. L. & Setlow, P. (1993). The cortical peptidoglycan from spores of Bacillus megaterium and Bacillus subtilis is not highly cross-linked. *J Bacteriol* **175**, 2767-9.

Pothoulakis, C. (2000). Effects of Clostridium difficile toxins on epithelial cell barrier. *Ann N Y Acad Sci* **915**, 347-56.

Poutanen, S. M. & Simor, A. E. (2004). Clostridium difficile-associated diarrhea in adults. *Cmaj* **171**, 51-8.

Pruitt, R. N. & Lacy, D. B. (2012). Toward a structural understanding of Clostridium difficile toxins A and B. *Front Cell Infect Microbiol* **2**, 28.

Phetcharaburanin, J., Hong, H. A., Colenutt, C., Bianconi, I., Sempere, L., Permpoonpattana, P., Smith, K., Dembek, M., Tan, S., Brisson, M. C., Brisson, A. R., Fairweather, N. & Cutting, S. M. (2014). The Spore-Associated Protein BclA1 Affects the Susceptibility of Animals to Colonization and Infection by Clostridium difficile. *Mol Microbiol*.

Putnam, E. E., Nock, A. M., Lawley, T. D. & Shen, A. (2013). SpoIVA and SipL are Clostridium difficile spore morphogenetic proteins. *J Bacteriol*.

Rea, M. C., Sit, C. S., Clayton, E., O'Connor, P. M., Whittal, R. M., Zheng, J., Vederas, J. C., Ross, R. P. & Hill, C. (2010). Thuricin CD, a posttranslationally modified bacteriocin with a narrow spectrum of activity against Clostridium difficile. *Proc Natl Acad Sci U S A* **107**, 9352-7.

Reinert, D. J., Jank, T., Aktories, K. & Schulz, G. E. (2005). Structural basis for the function of Clostridium difficile toxin B. *J Mol Biol* **351**, 973-81.

Ricca, E., Cutting, S. & Losick, R. (1992). Characterization of bofA, a gene involved in intercompartmental regulation of pro-sigma K processing during sporulation in Bacillus subtilis. *J Bacteriol* **174**, 3177-84.

Roels, S., Driks, A. & Losick, R. (1992). Characterization of spoIVA, a sporulation gene involved in coat morphogenesis in Bacillus subtilis. *J Bacteriol* **174**, 575-85.

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

Rudner, D. Z., Fawcett, P. & Losick, R. (1999). A family of membrane-embedded metalloproteases involved in regulated proteolysis of membrane-associated transcription factors. *Proc Natl Acad Sci U S A* **96**, 14765-70.

Rudner, D. Z. & Losick, R. (2001). Morphological coupling in development: lessons from prokaryotes. *Dev Cell* **1**, 733-42.

Rudner, D. Z. & Losick, R. (2002). A sporulation membrane protein tethers the pro-sigmaK processing enzyme to its inhibitor and dictates its subcellular localization. *Genes Dev* **16**, 1007-18.

Rupnik, M., Wilcox, M. H. & Gerding, D. N. (2009). Clostridium difficile infection: new developments in epidemiology and pathogenesis. *Nat Rev Microbiol* **7**, 526-36.

Schmidt, R., Margolis, P., Duncan, L., Coppolecchia, R., Moran, C. P., Jr. & Losick, R. (1990). Control of developmental transcription factor sigma F by sporulation regulatory proteins SpoIIAA and SpoIIAB in Bacillus subtilis. *Proc Natl Acad Sci U S A* **87**, 9221-5.

Schwan, C., Kruppke, A. S., Nolke, T., Schumacher, L., Koch-Nolte, F., Kudryashev, M., Stahlberg, H. & Aktories, K. (2014). Clostridium difficile toxin CDT hijacks microtubule organization and reroutes vesicle traffic to increase pathogen adherence. *Proc Natl Acad Sci U S A* **111**, 2313-8.

Sebaihia, M., Wren, B. W., Mullany, P., Fairweather, N. F., Minton, N., Stabler, R., Thomson, N. R., Roberts, A. P., Cerdeno-Tarraga, A. M., Wang, H., Holden, M. T., Wright, A., Churcher, C., Quail, M. A., Baker, S., Bason, N., Brooks, K., Chillingworth, T., Cronin, A., Davis, P., Dowd, L., Fraser, A., Feltwell, T., Hance, Z., Holroyd, S., Jagels, K., Moule, S., Mungall, K., Price, C., Rabbinowitsch, E., Sharp, S., Simmonds, M., Stevens, K., Unwin, L., Whithead, S., Dupuy, B., Dougan, G., Barrell, B. & Parkhill, J. (2006). The multidrug-resistant human pathogen Clostridium difficile has a highly mobile, mosaic genome. *Nat Genet* **38**, 779-86.

Serrano, M., Real, G., Santos, J., Carneiro, J., Moran, C. P., Jr. & Henriques, A. O. (2011). A negative feedback loop that limits the ectopic activation of a cell type-specific sporulation sigma factor of Bacillus subtilis. *PLoS Genet* **7**, e1002220.

Setlow, P. (1988). Small, acid-soluble spore proteins of Bacillus species: structure, synthesis, genetics, function, and degradation. *Annu Rev Microbiol* **42**, 319-38.

Setlow, P. (1995). Mechanisms for the prevention of damage to DNA in spores of Bacillus species. *Annu Rev Microbiol* **49**, 29-54.

Setlow, P. (2006). Spores of Bacillus subtilis: their resistance to and killing by radiation, heat and chemicals. *J Appl Microbiol* **101**, 514-25.

Shen, A. (2012). Clostridium difficile toxins: mediators of inflammation. *J Innate Immun* **4**, 149-58.

Shen-Orr, S. S., Milo, R., Mangan, S. & Alon, U. (2002). Network motifs in the transcriptional regulation network of Escherichia coli. *Nat Genet* **31**, 64-8.

Snell, H., Ramos, M., Longo, S., John, M. & Hussain, Z. (2004). Performance of the TechLab C. DIFF CHEK-60 enzyme immunoassay (EIA) in combination with the C.

difficile Tox A/B II EIA kit, the Triage C. difficile panel immunoassay, and a cytotoxin assay for diagnosis of Clostridium difficile-associated diarrhea. *J Clin Microbiol* **42**, 4863-5.

Sonenshein, A. L. (2000). Control of sporulation initiation in Bacillus subtilis. *Curr Opin Microbiol* **3**, 561-6.

Sorg, J. A. & Sonenshein, A. L. (2008). Bile salts and glycine as cogerminants for Clostridium difficile spores. *J Bacteriol* **190**, 2505-12.

Sorg, J. A. & Sonenshein, A. L. (2009). Chenodeoxycholate is an inhibitor of Clostridium difficile spore germination. *J Bacteriol* **191**, 1115-7.

Sorg, J. A. & Sonenshein, A. L. (2010). Inhibiting the initiation of Clostridium difficile spore germination using analogs of chenodeoxycholic acid, a bile acid. *J Bacteriol* **192**, 4983-90.

Spigaglia, P., Barketi-Klai, A., Collignon, A., Mastrantonio, P., Barbanti, F., Rupnik, M., Janezic, S. & Kansau, I. (2013). Surface-layer (S-layer) of human and animal Clostridium difficile strains and their behaviour in adherence to epithelial cells and intestinal colonization. *J Med Microbiol* **62**, 1386-93.

Stecher, B. & Hardt, W. D. (2011). Mechanisms controlling pathogen colonization of the gut. *Curr Opin Microbiol* **14**, 82-91.

Steil, L., Serrano, M., Henriques, A. O. & Volker, U. (2005). Genome-wide analysis of temporally regulated and compartment-specific gene expression in sporulating cells of Bacillus subtilis. *Microbiology* **151**, 399-420.

Steiner, E., Dago, A. E., Young, D. I., Heap, J. T., Minton, N. P., Hoch, J. A. & Young, M. (2011). Multiple orphan histidine kinases interact directly with Spo0A to control the initiation of endospore formation in Clostridium acetobutylicum. *Mol Microbiol* **80**, 641-54.

Stragier, P., Bonamy, C. & Karmazyn-Campelli, C. (1988). Processing of a sporulation sigma factor in Bacillus subtilis: how morphological structure could control gene expression. *Cell* **52**, 697-704.

Stragier, P., Kunkel, B., Kroos, L. & Losick, R. (1989). Chromosomal rearrangement generating a composite gene for a developmental transcription factor. *Science* **243**, 507-12.

Stragier, P. & Losick, R. (1996). Molecular genetics of sporulation in Bacillus subtilis. *Annu Rev Genet* **30**, 297-41.

Sun, D., Fajardo-Cavazos, P., Sussman, M. D., Tovar-Rojo, F., Cabrera-Martinez, R. M. & Setlow, P. (1991a). Effect of chromosome location of Bacillus subtilis forespore genes on their spo gene dependence and transcription by E sigma F: identification of features of good E sigma F-dependent promoters. *J Bacteriol* **173**, 7867-74.

Sun, D. X., Cabrera-Martinez, R. M. & Setlow, P. (1991b). Control of transcription of the Bacillus subtilis spoIIIG gene, which codes for the forespore-specific transcription factor sigma G. *J Bacteriol* **173**, 2977-84.

Sundriyal, A., Roberts, A. K., Ling, R., McGlashan, J., Shone, C. C. & Acharya, K. R. (2010). Expression, purification and cell cytotoxicity of actin-modifying binary toxin from Clostridium difficile. *Protein Expr Purif* **74**, 42-8.

Sussman, M. D. & Setlow, P. (1991). Cloning, nucleotide sequence, and regulation of the Bacillus subtilis gpr gene, which codes for the protease that initiates degradation of small, acid-soluble proteins during spore germination. *J Bacteriol* **173**, 291-300.

Sylvestre, P., Couture-Tosi, E. & Mock, M. (2002). A collagen-like surface glycoprotein is a structural component of the Bacillus anthracis exosporium. *Mol Microbiol* **45**, 169-78.

Tasteyre, A., Barc, M. C., Collignon, A., Boureau, H. & Karjalainen, T. (2001). Role of FliC and FliD flagellar proteins of Clostridium difficile in adherence and gut colonization. *Infect Immun* **69**, 7937-40.

Tatti, K. M., Shuler, M. F. & Moran, C. P., Jr. (1995). Sequence-specific interactions between promoter DNA and the RNA polymerase sigma factor E. *J Mol Biol* **253**, 8-16.

Theriot, C. M., Koenigsknecht, M. J., Carlson, P. E., Jr., Hatton, G. E., Nelson, A. M., Li, B., Huffnagle, G. B., J, Z. L. & Young, V. B. (2014). Antibiotic-induced shifts in the mouse gut microbiome and metabolome increase susceptibility to Clostridium difficile infection. *Nat Commun* **5**, 3114.

Traag, B. A., Pugliese, A., Eisen, J. A. & Losick, R. (2013). Gene conservation among endospore-forming bacteria reveals additional sporulation genes in Bacillus subtilis. *J Bacteriol* **195**, 253-260.

Traag, B. A., Ramirez-Peralta, A., Wang Erickson, A. F., Setlow, P. & Losick, R. (2013). A novel RNA polymerase-binding protein controlling genes involved in spore germination in Bacillus subtilis. *Mol Microbiol* **89**, 113-22.

Tracy, B. P., Jones, S. W. & Papoutsakis, E. T. (2011). Inactivation of sigmaE and sigmaG in Clostridium acetobutylicum illuminates their roles in clostridial-cell-form biogenesis, granulose synthesis, solventogenesis, and spore morphogenesis. *J Bacteriol* **193**, 1414-26.

Underwood, S., Guan, S., Vijayasubhash, V., Baines, S. D., Graham, L., Lewis, R. J., Wilcox, M. H. & Stephenson, K. (2009). Characterization of the sporulation initiation pathway of Clostridium difficile and its role in toxin production. *J Bacteriol* **191**, 7296-305.

van Nood, E., Vrieze, A., Nieuwdorp, M., Fuentes, S., Zoetendal, E. G., de Vos, W. M., Visser, C. E., Kuijper, E. J., Bartelsman, J. F., Tijssen, J. G., Speelman, P., Dijkgraaf, M. G. & Keller, J. J. (2013). Duodenal infusion of donor feces for recurrent Clostridium difficile. *N Engl J Med* **368**, 407-15.

Voth, D. E. & Ballard, J. D. (2005). Clostridium difficile toxins: mechanism of action and role in disease. *Clin Microbiol Rev* **18**, 247-63.

Waligora, A. J., Hennequin, C., Mullany, P., Bourlioux, P., Collignon, A. & Karjalainen, T. (2001). Characterization of a cell surface protein of Clostridium difficile with adhesive properties. *Infect Immun* **69**, 2144-53.

Wang, K. H., Isidro, A. L., Domingues, L., Eskandarian, H. A., McKenney, P. T., Drew, K., Grabowski, P., Chua, M. H., Barry, S. N., Guan, M., Bonneau, R., Henriques, A. O. & Eichenberger, P. (2009). The coat morphogenetic protein SpoVID is necessary for spore encasement in Bacillus subtilis. *Mol Microbiol* **74**, 634-49.

Wang, S. T., Setlow, B., Conlon, E. M., Lyon, J. L., Imamura, D., Sato, T., Setlow, P., Losick, R. & Eichenberger, P. (2006). The forespore line of gene expression in Bacillus subtilis. *J Mol Biol* **358**, 16-37.

Warth, A. D. & Strominger, J. L. (1969). Structure of the peptidoglycan of bacterial spores: occurrence of the lactam of muramic acid. *Proc Natl Acad Sci U S A* **64**, 528-35.

Wheeldon, L. J., Worthington, T. & Lambert, P. A. (2011). Histidine acts as a co-germinant with glycine and taurocholate for Clostridium difficile spores. *J Appl Microbiol*.

Wilson, K. H., Kennedy, M. J. & Fekety, F. R. (1982). Use of sodium taurocholate to enhance spore recovery on a medium selective for Clostridium difficile. *J Clin Microbiol* **15**, 443-6.

Wu, J. J., Howard, M. G. & Piggot, P. J. (1989). Regulation of transcription of the Bacillus subtilis spoIIA locus. *J Bacteriol* **171**, 692-8.

Wu, J. J., Piggot, P. J., Tatti, K. M. & Moran, C. P., Jr. (1991). Transcription of the Bacillus subtilis spoIIA locus. *Gene* **101**, 113-6.

Wu, L. J., Feucht, A. & Errington, J. (1998). Prespore-specific gene expression in Bacillus subtilis is driven by sequestration of SpoIIE phosphatase to the prespore side of the asymmetric septum. *Genes Dev* **12**, 1371-80.

Xue, Q., Gu, C., Rivera, J., Hook, M., Chen, X., Pozzi, A. & Xu, Y. (2011). Entry of Bacillus anthracis spores into epithelial cells is mediated by the spore surface protein BclA, integrin alpha2beta1 and complement component C1q. *Cell Microbiol* **13**, 620-34.

Yoshisue, H., Ihara, K., Nishimoto, T., Sakai, H. & Komano, T. (1995). Cloning and characterization of a Bacillus thuringiensis homolog of the spoIIID gene from Bacillus subtilis. *Gene* **154**, 23-9.

Zheng, L., Halberg, R., Roels, S., Ichikawa, H., Kroos, L. & Losick, R. (1992). Sporulation regulatory protein GerE from Bacillus subtilis binds to and can activate

or repress transcription from promoters for mother-cell-specific genes. *J Mol Biol* **226**, 1037-50.

Zheng, L., Keller, S. F., Lyerly, D. M., Carman, R. J., Genheimer, C. W., Gleaves, C. A., Kohlhepp, S. J., Young, S., Perez, S. & Ye, K. (2004). Multicenter evaluation of a new screening test that detects Clostridium difficile in fecal specimens. *J Clin Microbiol* **42**, 3837-40.

Zhou, R. & Kroos, L. (2004). BofA protein inhibits intramembrane proteolysis of pro-sigmaK in an intercompartmental signaling pathway during Bacillus subtilis sporulation. *Proc Natl Acad Sci U S A* **101**, 6385-90.

Chapter 2

Development of a fluorescent reporter for single cell analysis of gene expression and protein localization in *C. difficile*

This chapter contains data published in: Pereira, F. C., Saujet, L., Tomé, A. R., Serrano, M., Monot, M., Couture-Tosi, E., Martin-Verstraete, I., Dupuy, B., and Henriques, A. O. (2013). The spore differentiation pathway in the enteric pathogen *Clostridium difficile*. *PLoS Genet*, 9: e1003782

SUMMARY

Spores are central for the transmission of the human pathogen *C. difficile*. Despite the important role played by spores, the spore differentiation pathway is poorly characterized in *C. difficile,* in comparison with the model organism *B. subtilis*. Full dissection of the sporulation pathway in *B. subtilis* relied on the employment of transcriptional or translational fusions to GFP and its derivatives as reporters for single cell analysis. The use of GFP as a fluorescent reporter requires, however, the presence of oxygen, hindering its application to strict anaerobic organisms like *C. difficile*. Here we describe a fluorescent reporter system for *C. difficile* based on the SNAP-tag. The SNAP-tag is derived from the human O^6-alkylguanine-DNA alkyltransferase (hAGT) that can be covalently labeled in living cells using O^6-benzylguanine (BG) derivatives. We have optimized the SNAP coding sequence for expression in *C. difficile*, and we describe the construction of vectors to generate transcriptional and translational fusions in this organism. We show that the SNAP-tag can be specifically and quantitatively labeled and thus that it can be used as a reporter for single cell analysis of gene expression in *C. difficile*. We further demonstrate specific labeling of the orthogonal CLIP-tag in *C. difficile*. In addition, the SNAP-tag was also successfully employed for studies of protein localization in *C. difficile*. The ability to monitor gene expression at the single cell and population levels, as well as to localize specific proteins within a cell, will allow detailed analysis of processes such as sporulation and toxin production that are central to pathogenesis.

INTRODUCTION

C. difficile is a Gram-positive, spore-forming, obligate anaerobe that is able to cause infection and disease, with symptoms ranging from mild diarrhea to pseudomembranous colitis, toxic megacolon, bowel perforation and death (Deneve *et al.*, 2009; Rupnik *et al.*, 2009). *C. difficile* infections (CDI) are steadily increasing worldwide and it is nowadays the most common cause of healthcare-acquired infections in the western hemisphere (Ghantoji *et al.*, 2010; Redelings *et al.*, 2007; Freeman *et al.*, 2010). A key element to the success of *C. difficile* as a pathogen is its ability to produce spores. Spores are resistant to most disinfectants and antibiotics, making them difficult to eradicate both from infected humans and the environment (Nicholson *et al.*, 2000). As a result, *C. difficile* spores disseminate from person to person and cause high rates of recurrent infections, which can lead to serious illness or even death (Deneve *et al.*, 2009; Maroo and Lamont, 2006; Deakin *et al.*, 2012). Although spores are critical to the pathogenesis of *C. difficile*, their composition and formation remain poorly characterized.

The use of transcriptional and translational fusions to GFP and its variants has proved invaluable as a tool to study the process of spore formation in *B. subtilis* (eg, Wang *et al.*, 2009). Fusions of promoters of genes under the control of each of the sporulation sigma factors have enabled the establishment of the cell compartment as well as the main periods of activity of these factors (Hilbert and Piggot, 2004). However this methodology cannot be applied in strict anaerobes like *C. difficile*, since emission of fluorescence by the GFP chromophore requires cyclization and oxidation of an internal tripeptide motif, and the last step of this reaction, which is autocatalytic, requires oxygen (Reid and Flynn, 1997). To overcome the limitations associated with the use of autofluorescent proteins in strict anaerobes, new proteins and protein tags to be used as fluorescent reporters

in the absence of oxygen have been developed. Flavin mononucleotide-based (FMN) fluorescent proteins, engineered from the light oxygen voltage (LOV) domains of the *Bacillus subtilis* or *Pseudomonas putida* blue-light photoreceptors can be mentioned as a good example (Drepper *et al.*, 2007). These proteins were shown to fluoresce in the absence of oxygen, when produced in the facultative anaerobic bacterium *Rhodobacter capsulatus* (Drepper *et al.*, 2007), and also during hypoxia in the pathogen *Candida albicans* (Ernst and Tielker, 2009; Tielker *et al.*, 2009). However, this approach comes with handicaps. All of the FMN proteins generated so far emit fluorescence at the same wavelength, which restricts multicolour imaging. Moreover, the wavelength of emission may coincide with autofluorescence of the cells, as is the case for *C. difficile* (George *et al.*, 1979; see below).

Other tools developed to allow fluorescence-based studies in the absence of oxygen consist on protein tags that can be site-specific labeled with chemical probes (Keppler *et al.*, 2003; Johnsson and Johnsson, 2007; Hinner and Johnsson, 2010) or fluorescent non-natural amino acids (Liu and Schultz, 2010). Protein tags are fused to a protein or placed under the command of a promoter of interest, and can be covalently labeled with a small molecule, thereby combining the simplicity of fusion protein expression with the diversity of the molecular probes provided by chemistry. The tetracysteine-tag (Hoffmann *et al.*, 2010), the Halo-tag (Los *et al.*, 2008), and the SNAP-tag (Keppler *et al.*, 2003), represent some of these tags. The SNAP-tag is a 20 kDa engineered variant of the human repair protein O^6-alkylguanine-DNA alkyltransferase (hAGT) that covalently reacts with O^6-benzylguanine derivatives, in an irreversible manner (Keppler *et al.*, 2003; Keppler *et al.*, 2004a; Gronemeyer *et al.*, 2006) (Figure 2.1). The SNAP crystal structure was resolved and is shown on Figure 2.1A (pdb code: 3KZY; see also 3KZZ for the structure of the SNAP-tag bound to its substrate benzylguanine). During the reaction with a substrate, a stable thioether bond

is formed between the reactive cysteine of the tag and the chemical probe (Figure 2.1B). The fact that it also reacts with nucleotides containing derivatives of O^6-benzylguanoside (BG) with substituted benzyl rings allowed the development of several chemical substrates that specifically react with the SNAP-tag (Keppler *et al.*, 2003). BG substrates are inert, therefore avoiding unspecific labeling, and allow labeling of the SNAP-tag in any cellular compartment (Keppler *et al.*, 2004b), characteristics that increase the attractiveness of the SNAP-tag over other tags. In addition to labeling with fluorescent probes, SNAP-tag proteins can be modified with affinity ligands or other binding moieties (Kindermann *et al.*, 2004), and used for crosslinking experiments to monitor protein-protein interactions inside living cells (Lemercier *et al.*, 2007; Gautier *et al.*, 2009), or immobilized on solid surfaces for purification, pull-downs or protein microarray experiments (Kindermann *et al.*, 2004). A recently developed split-SNAP system also allows for the detection of protein-protein interactions in living cells (Mie et al., 2012).

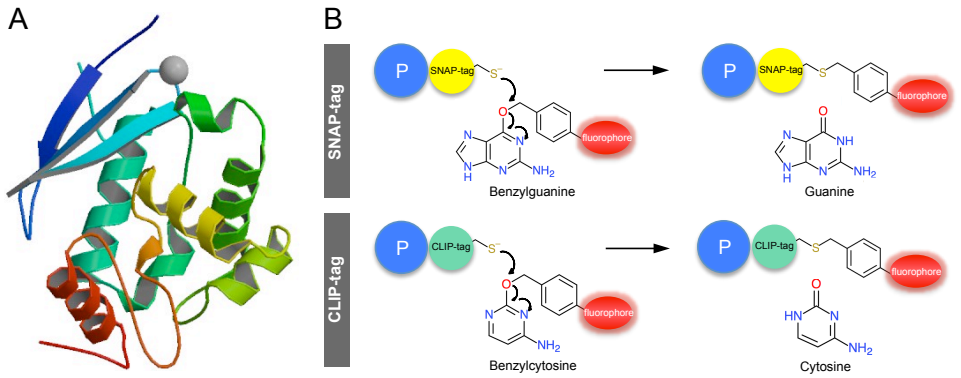


Figure 2.1. The SNAP-tag technology. (A) Crystal structure of SNAP-tag (PDB: 3KZY). Proteins chains are colored from the N-terminal to the C-terminal using a rainbow (spectral) color gradient. **(B)** Schematic representation of the mechanism of reaction of SNAP and CLIP with their substrates. SNAP-tag or CLIP-tag fused to the protein of interest reacts with with a fluorophore-modified benzylguanine or benzylcytosine substrate, releasing guanine or cytosine.

All the applications described above highlight the flexibility of the SNAP-tag. However, a handicap of this approach was the simultaneous labeling of tagged proteins. A step further in this direction was given by mutagenizing eight amino acids in the SNAP-tag, generating the CLIP-tag, which irreversibly reacts with O^2-benzylcitosine (BC) derivatives (Gautier *et al.*, 2008) (Figure 2.1B). Because SNAP-tag shows high selectivity for BG over BC derivatives, the SNAP-tag and CLIP-tag can be simultaneously used due to orthogonal labeling with different chemical probes (Figure 2.1B). Importantly, the similar properties of these two protein tags represents an advantage when comparisons of the properties of one fusion protein to another need to be made.

The SNAP-tag has been mainly used to visualize proteins inside mammalian cells. Nevertheless, this technique has also been successfully applied to studies of gene expression and protein localization in anaerobic organisms (Regoes and Hehl, 2005; Martincova *et al.*, 2012; Nicolle *et al.*, 2010). Here we have extended the application of the SNAP-tag technology to the strict anaerobe *C. difficile*. Conditions for optimal SNAP accumulation and complete SNAP labeling with its substrate when SNAP is produced under the control of the inducible P_{tet} promoter were established. The P_{tet} is the only inducible system fully functional in *C. difficile* (Fagan and Fairweather, 2011) and was employed here with the aim of optimizing the use of the SNAP-tag as a fluorescent reporter system for single cell and population level studies of gene expression in this organism. The experiments performed and the toolbox herein developed represent the basis for the use if the SNAP technology to dissect of the *C. difficile* spore formation pathway.

MATERIALS AND METHODS

General methods. The *Escherichia coli* strain DH5α (Bethesda Research laboratories) was used for molecular cloning. Luria-Bertani medium was routinely used for growth and maintenance of *E. coli*. The *C. difficile* strains used in this study are congenic derivatives of the wild-type strain 630Δ*erm* (Hussain *et al.*, 2005) and were routinely grown anaerobically (5% H2, 15% CO2, 80% N2) at 37ºC in brain heart infusion (BHI) medium (Difco). To induce sporulation, the wild-type strain 630Δ*erm* was grown in Sporulation Medium (SM) for 24 hours. When necessary thiamphenicol (15 μg /ml) was added to *C. difficile* cultures.

Transcriptional *SNAP^{Cd}* fusions. We first obtained a synthetic version of the *SNAP26b* gene (encoding a mutant form of the human gene for *O^6*-alkylguanine-DNA-alkyltransferase (New England Biolabs) (Keppler *et al.*, 2003) codon usage-optimized for expression in *C. difficile* (DNA 2.0, Menlo Park, CA). The synthetic gene cassette, hereinafter termed *SNAP^{Cd}*, includes a ribosome-binding site (RBS) and flanking XhoI and HindIII sites. To construct a P$_{tet}$-*SNAP^{Cd}* fusion (*SNAP^{Cd}* under the control of P$_{tet}$ inducible promotor), the synthetic *SNAP^{Cd}* cassette was PCR-amplified using primer pair SNAPtag_SacI Fw/SNAPtag_BamHI Rev and inserted between the SacI and BamHI sites of pRPF185 (Fagan and Fairweather, 2010), replacing the *gusA* gene, and yielding pFT46 (Figure 2.2A). Plasmids used or constructed in this study are listed in Table A1 of Appendices and the oligonucleotide primers used for PCR, mutagenesis or sequencing are listed in Table A2. pFT46 was subsequently introduced into *E. coli* HB101 (RP4), originating strain AHCD095, and transferred to *C. difficile* 630Δ*erm* strain by conjugation (Heap *et al.*, 2007), yielding strain AHCD586. Anhydrotetracycline was used for induction of the P$_{tet}$ promoter present in *C. difficile* strains bearing pFT46.

The synthetic *SNAP*^{Cd} sequence was also cloned between the XhoI/HindIII sites of pMTL84121, to produce pFT47 (Table A2 and Figure 2.2C), allowing transcriptional fusion of genes of interest to *SNAP*^{Cd} in the future.

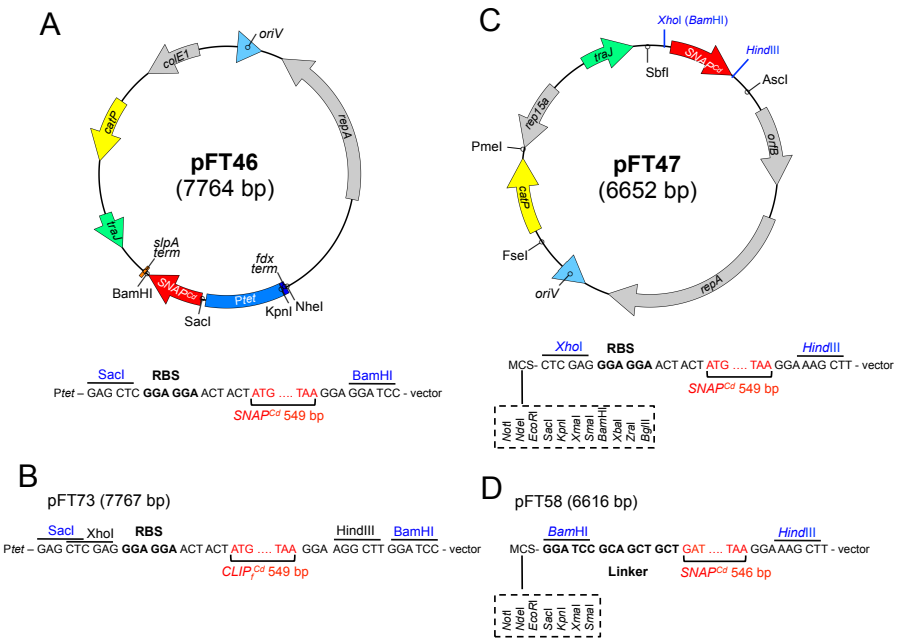


Figure 2.2. SNAP-tag vectors used in the study. (A) Schematic representation of pFT46. The P$_{tet}$ -*SNAP*^{Cd} fusion is flanked by the transcriptional terminator sites of the *slpA* and *fdx* genes. Unique restricition sites are indicated. A detailed representation of the region between SacI/and BamHI sites (in blue) is shown on the bottom (for pFT47) and in (B) (for pFT73). **(B)** For the P$_{tet}$ -*CLIP*$_f^{Cd}$ fusion, the region between SacI and BamHI sites of pFT46 (in blue) was replaced by the one depicted, yielding pFT73. **(C)** Schematic representation and features of pFT47 vector. Unique restriction sites (SbfI, AscI, FseI, PmeI) that allow replacement of functional blocks among plasmids of the pMTL80000 series are also indicated along the circular plasmid map. A detailed representation of the region between XhoI/BamHI and HindIII sites (in blue) is shown on the bottom (for pFT47) and in (D) (for pFT58). The *SNAP*^{Cd} gene, preceded by an RBS, was inserted between the XhoI and HindIII sites within the multiple cloning site (MCS) of pMTL84121 (unique sites upstream of XhoI, that are part of the pMTL84121 MCS, are indicated), yielding pFT47. **(D)** For translational fusions, the region between BamHI and HindIII sites of pFT47 (in blue) was replaced by the one depicted, yielding pFT58.

Transcriptional *CLIP*$_f^{Cd}$ fusion. A synthetic version of the *Clip*$_f$ gene was obtained as described for the *SNAP*^{Cd}. Briefly, a synthetic gene cassette, codon usage-optimization for expression in *C. difficile* was synthesized (DNA

2.0, Menlo Park, CA). The cassete, hereinafter termed *CLIP*$_f^{Cd}$, includes a ribosome-binding site (RBS) and flanking SacI/XhoI and HindIII/BamHI sites. To construct a P$_{tet}$-*CLIP*$_f^{Cd}$ fusion (*CLIP*$_f^{Cd}$ under the control of P$_{tet}$ inducible promotor), the synthetic *CLIP*$_f^{Cd}$ cassette was cut from DNA2.0 cloning vector using SacI and BamHI restriction enzymes and inserted between the SacI and BamHI sites of pRPF185 (Fagan and Fairweather, 2010), replacing the *gusA* gene, and yielding pFT73 (Figure 2.2A and B and Table A1). pFT73 was subsequently introduced into *E. coli* HB101 (RP4), originating strain AHCD165, and transferred to *C. difficile* 630Δ*erm* strain by conjugation, yielding strain AHCD714. Anhydrotetracycline (ATc) was used for induction of the P$_{tet}$ promoter present in *C. difficile* strains bearing pFT73.

Translational SNAP^{Cd} fusions. To allow the construction of C-terminal SNAP-tag protein fusions, the *SNAP*^{Cd} sequence was amplified from pFT47 using primers SNAP-linker BamHI Fw and SNAP-tag Rev and cloned between BamHI and HindIII sites of pMTL84121. The resulting plasmid, pFT58, contains the *SNAP*^{Cd} sequence without the start codon (Figure 2.2C and D). To construct a SpoIIQ-SNAP fusion, a fragment encompassing the coding sequence of *spoIIQ* (666 bp length) and 322 bp of its regulatory region was PCR amplified using primer pair CDspoIIQ-40D (EcoRI) and CDspoIIQ-R (BamHI) (Table A2; it should be referred that this primer removes the STOP codon from *spoIIQ*, and contains part of the linker sequence that connects the proteins of interest to SNAP). The resulting 988 bp fragment was cloned between the NotI and EcoRI sites of pFT58 to yield pMS480 (Table A1). To construct a SpoIIIAH-SNAP fusion, the entire coding sequence of the *spoIIIAH* gene (654 bp length) and its expected promoter region comprising the 533 bp upstream of *spoIIIAA* were independently PCR amplified using primer pairs PspoIIIAA-SNAPEcoRI_Fw/ CDspoIIIAA-spoIIIAH_Rev and CDspoIIIAH_D/CDspoIIIAH_R(BamHI) (Table A2; it should be referred again that this last Rev primer removes the STOP codon from *spoIIIAH*, and

contains part of the linker sequence that connects the proteins of interest to SNAP), joined by SOE PCR and the resulting 1187 bp fragment cloned between the EcoRI and BamHI sites of pFT58 to create pMS481 (Table A1). Plasmids pMS480 and pMS481 were subsequently transformed into *E. coli* HB101 (RP4), originating strains AHCD166 and AHCD167, and then transferred to *C. difficile* 630Δ*erm* strain by conjugation (Heap *et al.*, 2007), yielding strains AHCD727 and AHCD728, respectively.

SNAP labeling and analysis. Whole cell extracts were obtained by withdrawing 10 ml samples from *C. difficile* cultures in brain heart infusion (BHI) of the P_{tet}-SNAP-bearing strain at the desired times. The extracts were prepared immediately following labeling with a range of concentrations of the TMR-Star substrate (New England Biolabs), for various labeling times, in the dark. Following labeling, the cells were collected by centrifugation (4000xg, for 5 min at 4ºC), the cell sediment was washed with phosphate-buffered saline (PBS) and resuspended in 1 ml French press buffer (10 mM Tris pH 8.0, 10 mM MgCl2, 0.5 mM EDTA, 0.2 mM NaCl, 10% Glycerol, 1 mM PMSF). The cells were lysed using a French pressure cell (18000 lb/in^2). Proteins in the extracts were resolved on 15% SDS-PAGE gels. The gels were first scanned in a Fuji TLA-5100 fluorimager, and then subject to immunoblot analysis as described before (Serrano *et al*., 2011). The anti-SNAP antibody (New England Biolabs) was used at a 1:1000 dilution, and a rabbit secondary antibody conjugated to horseradish peroxidase (Sigma) was used at dilution 1:10000. The immunoblots were developed with enhanced chemiluminescence reagents (Amersham Pharmacia Biotech).

Microscopy and image analysis. Cells present in 1ml culture samples were labeled with TMR-Star (as above), collected by centrifugation (4000xg, 3 min, at room temperature), washed four times with 1 ml of PBS, and finally ressupended in 1ml of PBS containing the DNA stain DAPI (4',6-

diamidino-2-phenylindole; 50 µ/ml) or the membrane dye Mitotracker Green (0.5 µg/ml) (both from Molecular Probes/Invitrogen), when indicated. For phase contrast and fluorescence microscopy, cells were mounted on 1.7% agarose coated glass slides and observed on a Leica DM6000B microscope equipped with a phase contrast Uplan F1 100x objective and a CCD Ixon^{EM} camera (Andor Technologies) (Serrano *et al.*, 2011). Images were acquired and analyzed using the Metamorph software suite version 5.8 (Universal Imaging), and adjusted and cropped using ImageJ (http://rsbweb.nih.gov/ij/). For quantification of the *SNAP^{Cd}* signal resulting from transcriptional fusions, 6x6 pixel regions were defined in the desired cell and the average pixel intensity was calculated, and corrected by subtracting the average pixel intensity of the background. Small fluctuations of fluorescence among different fields were corrected by normalizing to the average pixel intensity obtained for the intrinsic autofluorescence of *C. difficile* cells (George *et al.*, 1979).

RESULTS

Specific labeling of *C. difficile* cells with the TMR-Star SNAP substrate

With the aim of developing a fluorescent-based reporter for single cell analysis in *C. difficile* we have synthesized a variant of the *SNAP26b* gene, which we termed *SNAP^{Cd}*, preceded by a ribosome-binding site, that was codon-usage optimized for expression in *C. difficile*. This cassette was placed under the control of the anhydrotetracycline (ATc) responsive P_{tet} promoter in a replicative plasmid (Fagan and Fairweather, 2011), to yield pFT46 (Figure 2.2A). Plasmid pFT46 was introduced into the wild type strain 630Δ*erm*, and the resulting strain AHCD586 was grown in BHI and induced by addition of 1000 ng/ml of ATc for 2 hours (Fagan and Fairweather, 2011). Cells were labeled with 50 nM of a BG derivative coupled to the

tetramethylrodamine fluorophore (TMR-Star; Figure 2.3A), for 30 min. Both the concentration and incubation time were selected based on the literature (Donovan and Bramkamp, 2009; Nicolle *et al.*, 2010) and on preliminary assays where we noticed that the use of 10 times more TMR-Star results in high levels of background fluorescence (not shown). To confirm that labeling of SNAP with this substrate is specific of cells expressing *SNAP*^{Cd}, we have labeled both induced and non-induced cells of strain AHCD586, and these last were further labeled with the DNA dye DAPI to allow differentiation from the induced ones under the same microscopy field. Cells were then mixed and examined by phase contrast and fluorescence microscopy (Figure 2.3B). Results clearly suggest that labeling of *C. difficile* cells is specific of cells expressing *SNAP*^{Cd} as there is no overlap between the red (SNAP-TMR-Star) and blue (DAPI) channels (Figure 2.3B). Other SNAP-tag cell-permeable substrates commercially available were tested, but the TMR-Star fluorophore was the one that produced the highest signal-to-noise ratio, consistent with the fact that its emission peak is the furthermost from the wavelength at which *C. difficile* cells show autofluoresce (George *et al.*, 1979).

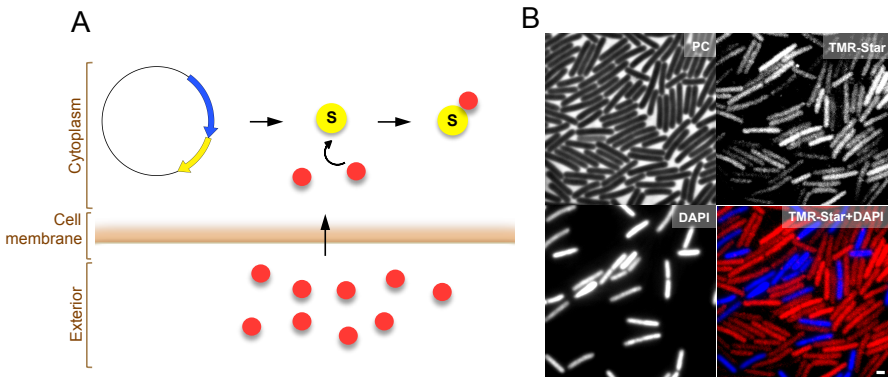


Figure 2.3. **The SNAP technology extended to *C. difficile*. (A)** Schematic representation of SNAP labeling. *C. difficile* cells carrying a replicative plasmid where the *SNAP*^{Cd} sequence (yellow) is present under the control or fused to a gene of interest (blue), are cultivated under appropriate conditions that allow production of the SNAP protein, time at which a cell permeable label (red) should be added. The permeable label enters into the cell, where a covalent modification occurs, allowing labeling and visualization of the SNAP protein by fluorescence. **(B)** Microscopy analysis of *C. difficile* cells producing SNAP under the control of P_{tet} (in pFT46). Cells were labeled with 50 nM of

the TMR-Star cell-permeable SNAP substrate (red) and mixed non-induced cells, labeled in the same conditions but posteriorly stained also with DAPI (blue), to allow distinction of the two cell types. Scale bar: 1 μm.

Optimization of SNAP labeling times

Labeling of *C. difficile* cells expressing *SNAP*^{Cd} with the TMR-Star substrate seems specific in the conditions shown above. However, the SNAP was produced under the control of an inducible promoter whose induction might lead to higher levels of SNAP than the ones that can be achieved with promoters of sporulation genes we plan to analyse in the future. Being so, we went to further optimize the different labeling parameters in order to get as high fluorescent signal as possible, keeping it specific. First, we labeled cells from strain AHCD586 with SNAP TMR-Star substrate for various times (Figure 2.4), and analysed cells by fluorescence microscopy. Cells to which TMR-Star was not added, ("0" in Figure 2.4A), did not display fluorescence, as judged by microscopy and quantification of the signal across the population (Figure 2.4A and B), confirming that no background fluorescence was registered. The average fluorescence signal for the cell population increased linearly with the labeling time up to 30 min (Figure 2.4A and B). Importantly, at any labeling time tested, about 20% of the cell population did not display fluorescence above the background ("0" time labeling) (Figure 2.4C). This number of unlabeled cells might represent a sub-population of cells refractory to either ATc or TMR-Star internalization, or might represent a normal response to this promoter, where P_{tet} activation is possibly a stochastic event that is not achieved in a fraction of cells. However we did not investigated this issue further. In any event, this experiment establishes 30 min as an optimal labeling time of SNAP with the TMR-Star substrate in *C. difficile* when the tag is produced from the P_{tet} promoter in the presence of 1000 ng/ml ATc.

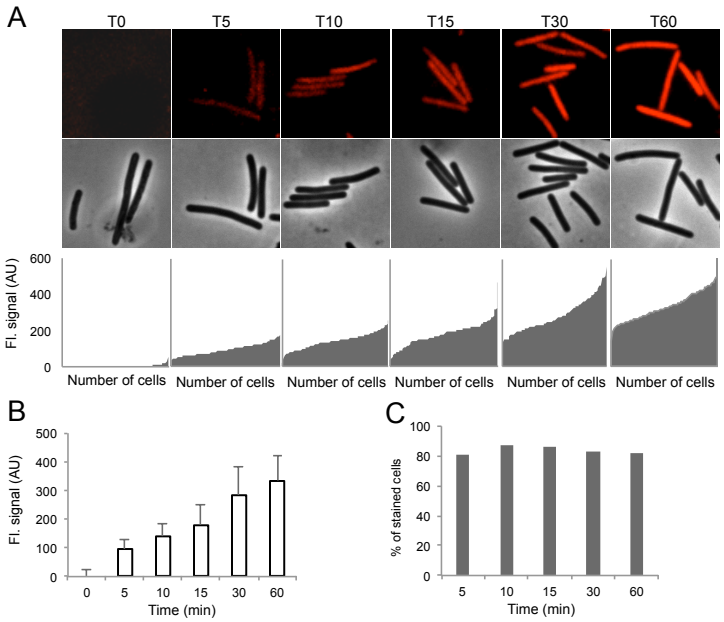


Figure 2.4. **Optimal labeling time of *C. difficile* cells producing SNAP under the control of P_{tet} (in pFT46) with the TMR-Star substrate. (A)** Phase contrast (PC) and fluorescence microscopy analysis of cells labeled with 50 NM of the TMR-Star substrate during the indicated times (in minutes). The panels on the bottom show the relative levels of SNAP-TMR-Star fluorescence (Fl.) within individual cells (n=150), for the indicated labeling times, in arbitrary units (AU). Scale bar, 1μm. **(B)** Values represented by the bars are the average ± SD of the fluorescence (Fl.) intensity shown in A. **(C)** Percentage of cells showing a fluorescence signal above the background level for each labeling time (in minutes).

Maximal accumulation of SNAP under the P_{tet} control

Induction of the SNAP promotor with ATc is dose-dependent and linear within a range of 0-500 ng/ml of ATc (Fagan and Fairweather, 2010). This led us to test whether in our conditions (1000 ng/ml of ATc for 2 hours) the accumulation of SNAP is maximal. To answer this question, we took advantage of the covalent label of SNAP by the fluorescent probe. Proteins in total cell extracts were resolved by SDS-PAGE and the gels scanned on a fluorimager to detect the signal from SNAP-TMR-Star. Gels were further processed for immunoblot analysis with an anti-SNAP antibody. No SNAP-TMR-Star signal is detected in wild type 630Δ*erm* cells that do not carry pFT46, even for the maximum time of induction or concentration of inducer tested, neither in cell extracts immediately after induction ("0" in Figure

2.5A). Importantly, accumulation of SNAP increases, reaching its maximum, 60 min after induction (Figure 2.5A). By monitoring the accumulation of SNAP in these conditions, but varying the concentrations of the inducer (0-1000 ng/ml ATc) we could also observe that there is no increase in the SNAP or SNAP-TMR-Star levels for ATc concentrations higher than 250 ng/ml (Figure 2.5B). Therefore, we conclude that future analysis using the P_{tet} system should performed with induction with 250 ng/ml ATc for 1 hour. Interestingly, immunoblot analysis of cell extracts of strain 630Δ*erm* bearing the P_{tet}-*SNAP*^{Cd} fusion revealed the presence of at least two clearly distinct bands corresponding to the accumulation of the SNAP protein (Figure 2.5A and B). The expected size for the SNAP is of 20 kDa (Keppler *et al.*, 2003). We noticed that when cells were not labeled with the TMR-Star substrate, used as a control for specificity of the fluorimager signal, the band corresponding to the higher molecular form of the SNAP protein is absent (Figure 2.5B). Labeling with 50 nM of TMR-Star resulted in the appearance of a band with a slightly higher size, consistent with formation the covalent modification of the SNAP by the TMR-Star substrate (this reaction involves the formation of a stable thioether bond between the reactive cysteine of the SNAP protein and the substrate) (Keppler *et al.*, 2003). Importantly, these results suggest the presence of unlabeled SNAP protein inside the cells in our conditions. Indeed, increasing the concentration of SNAP to 250 nM of TMR-Star resulted in a higher SNAP-TMR-Star signal detected under the fluorimager. This band was fluorescent, confirming reaction with the SNAP (Figure 2.5C, red arrows). Concomitantly with this increase, the fast migrating band forms of SNAP disappear from the WB, rendering a single SNAP band (Figure 2.5C). We therefore conclude that labeling with concentrations higher than 50 nM of TMR-substrate are needed in order to fully label the SNAP produced under the P_{tet} control in our conditions.

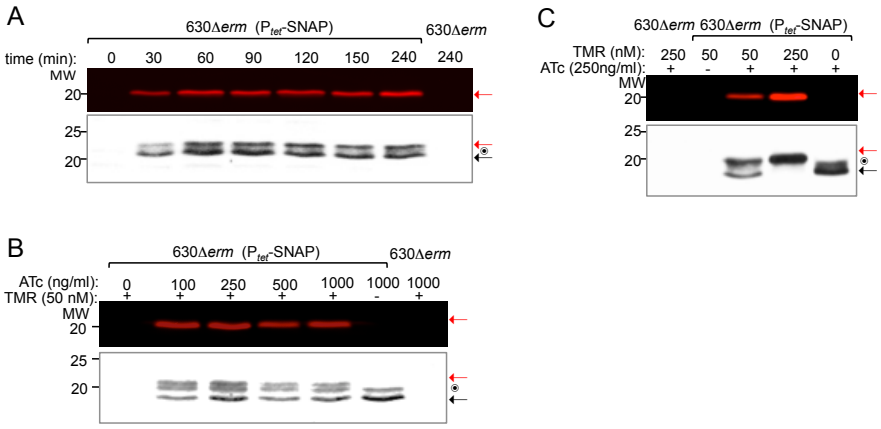


Figure 2.5. SNAP accumulation under the control of P_{tet} (in pFT46), for different induction conditions. Fluorescence scanning using a fluorimager (top) and immunoblot analysis using anti-SNAP antibody (bottom) of whole cell extracts prepared from: **(A)** TMR-Star (50 nM) labeled cells after induction with 1000 ng/ml of ATc, for the different indicated times; **(B)** TMR-Star (50 nM) labeled cells induced for 1 hour with the indicated ATc concentrations; **(C)** cells induced with 250 ng/ml of ATc for 1 hour and labeled with different concentrations (50 nM or 250 nM) of TMR-Star, as indicated. In (A), (B) and (C) black or red arrows point to unlabeled or labeled SNAP, respectively. The dot indicates a second, slow migrating form of the unlabeled SNAP protein. The position of molecular weight markers (in kDa) is shown on the left side of the panels.

Complete labeling of SNAP with the TMR-Star substrate

Having established both the concentration of inducer and induction times that lead to optimal accumulation of SNAP under the the control of P_{tet}, we went to determine which is the minimal concentration of TMR-Star that leads to full labeling of the SNAP protein. To do so, *C. difficile* cells bearing pFT46 and induced with 250 ng/ml of ATc for 1 hour were collected and labeled with increasing concentrations of the TMR-Star substrate for 30 min. Cells were first examined by phase contrast and fluorescence microscopy. Cells to which TMR-Star was not added, ("0" in Figure 2.6A), or cells that were not induced with ATc but labeled with the highest concentration of the substrate tested (250 nM), did not display fluorescence, as judged by microscopy and quantification of the signal across the population (Figure 2.6A). Therefore, even in the presence of the TMR-Star substrate, no background fluorescence was registered. Moreover, staining of non-induced

cells with 250 nM of TMR-Star did not result in any detectable fluorescence signal, regardless of the incubation time (up to 1 hour; not shown). The average fluorescence signal for the cell population increased linearly with the concentration of TMR-Star in the labeling reaction (Figure 2.6A and B). For any concentration of TMR-Star tested, about 20% of the cell population did not display fluorescence above the background value ("0" time labeling) (Figure 2.6C), which it is also in agreement with results obtained before (Figure 2.4C). We presume that the number of unlabeled cells reflects a normal response of the P_{tet} promoter across the cell population (with transcription not activated in the 20% sub-population), rather than for instance a sub-population of cells refractory to TMR internalization. We base this assumption on the observation that no unlabeled SNAP could be detected by immunoblotting, for the highest concentration of substrate used (see below).

To monitor the extent of labeling of the SNAP protein, we prepared whole cell extracts from cells of strain 630Δ*erm* bearing the P_{tet}-*SNAP*^{Cd} fusion, following incubation with increasing (50-250 nM) concentrations of TMR-Star. No SNAP production was detected, by immunobloting, for the uninduced strain (Figure 2.6D). For the sample corresponding to the induced culture, but unlabeled cells, the immunoblot analysis showed accumulation of the SNAP protein, which migrates just below the 20kDa size marker (Figure 2.6D). This species was non-fluorescent, as expected and revealed by the fluorimager scan (Figure 2.6D, black arrow). Labeling with 50 nM of TMR-Star resulted in the appearance of a band with a slightly higher size, consistent with the covalent modification of the SNAP by the TMR-Star substrate (see also Figure 2.5D). This band was fluorescent, confirming reaction with the SNAP (Figure 2.6D, red arrows). However, the unlabeled SNAP was still detected (Figure 2.6D, black arrow). Complete labeling of the SNAP-tag produced from the P_{tet}-*SNAP*^{Cd} fusion was only achieved for a concentration of TMR-Star of 250 nM (Figure 2.6D). In all, these experiments

show the usefulness of the SNAP tag as a reporter for gene expression, and the need to optimize the concentration of the TMR-Star substrate to a particular reporter construct, so that complete labeling can be achieved.

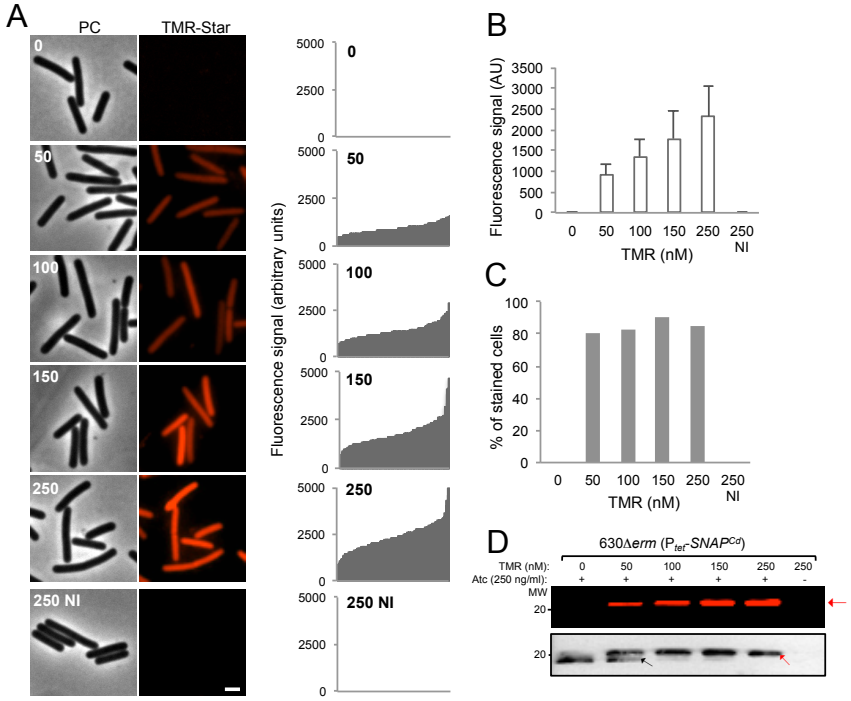


Figure 2.6. **SNAP labeling of *C. difficile* cells producing SNAP under the control of P_{tet} (in pFT46) with different concentrations of TMR-Star substrate**. **(A)** Phase contrast (PC) and fluorescence microscopy analysis of cells labeled with the indicated concentrations (nM) of the SNAP substrate TMR-Star. The panels on the right show the relative levels of SNAP-TMR-Star fluorescence within individual cells (n=150), for the indicated labeling times. Scale bar, 1μm. **(B)** Values represented by the bars are the average ± SD of the fluorescence intensity shown in A. **(C)** Percentage of cells showing a fluorescent signal above the background level for each labeling concentration. **(D)** Fuorimager scanning (top) and immunoblot analysis (bottom) of extracts from cells were prepared immediately following labeling with increasing concentrations of TMR-Star (nM), as described in (A). Black or red arrows point to unlabeled or labeled SNAP, respectively. The position of molecular weight markers (in kDa) is shown on the left side of the panels.

Finally, to allow the use of the SNAP technology as a reporter system for gene expression in *C. difficile*, we have further constructed a promoter-less plasmid, pFT47, for the construction of transcriptional fusions to *SNAP^{Cd}* (Figure 2.2B). The construction of this plasmid is detailed in Materials and

Methods. Work presented in Chapters 3 to 5 illustrate its use for studies of the time and cell type-specific expression of genes involved in the control of spore morphogenesis.

Protein localization in *C.difficile* using the SNAP-tag

Once established the conditions that allow the use of the SNAP-tag as a reporter for single cell analysis of gene expression in *C. difficile*, we went to determine whether the SNAP-tag is also suitable for protein localization in this organism. To address this issue, we have constructed a plasmid, named pFT58 (see Materials and Methods and Figure 2.2C and D) that allows the fusion of the SNAP-tag to the C-terminal of a protein of interest. Bridging the protein of interest to the SNAP protein is a small linker of nine aminoacids length (Figure 2.2C and D). We have then constructed fusions of the *C. difficile* SpoIIQ and SpoIIIAH, two proteins involved in spore formation in the model organism *B. subtilis*, to SNAP (see Materials and Methods). In this organism, SpoIIQ and SpoIIIAH are known to localize at the forespore inner and outer membranes, respectively, and to interact with each other forming a channel that connects the mother cell and forespore compartments during sporulation (Camp and Losick, 2008; Higgins and Dworkin, 2012; Meisner *et al.*, 2008). Plasmids carrying either SpoIIQ or SpoIIIAH-SNAP fusions were introduced into *C. difficile* wild type cells. Cells were labeled with 250 nM of the TMR-Star substrate and analysed by fluorescence microscopy. Results show that both proteins localize around the engulfed forespore, in a pattern that mimics the one observed in the model organism *B. subtilis* (Figure 2.7A and B). Moreover, fusion of these proteins to SNAP does not interfere with the process of spore formation in an otherwise wild type strain. This suggests that these fusions are functional although more studies are need to fully clarify this issue. Overall, we can conclude that the SNAP-tag is a suitable tool for studies of protein localization in *C. difficile*.

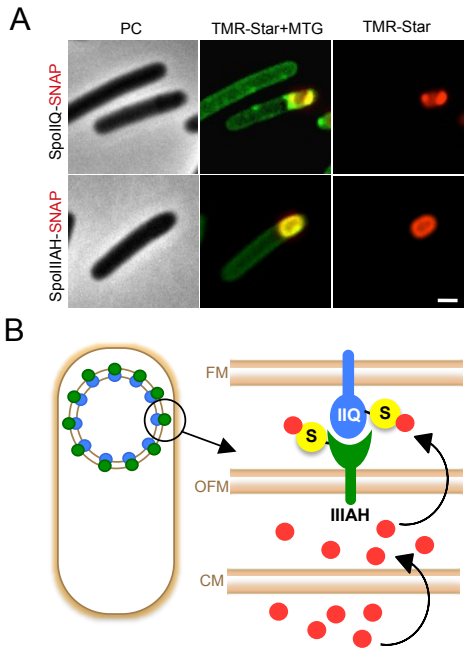


Figure 2.7. Protein localization in *C. difficile* using the SNAP-tag. (A) Fluorescence microscopy of sporulating cells carrying SpoIIQ- and SpoIIIAH-SNAP fusions. Cells were collected after 24h of growth in SM medium and labeled with 250nM of the TMR-Star substrate (red) and the membrane dye MTG (green). Scale bar: 1 µm. **(B)** Schematic representation of the channel formed by the interacting proteins SpoIIQ and SpoIIIAH during endospore formation. SpoIIQ and SpoIIIAH are colored in blue and green, respectively. At the point of the sporulation process depicted, the forespore is completely enclosed in the mother cell cytoplasm. SpoIIQ and SpoIIIAH are associated with the inner and outer forespore membranes (IFM and OFM, respectively), as depicted. The C-terminal fused SNAP protein (yellow), physically localizes at the inter-membrane space, where it can be covalently labeled by the cell permeable TMR-Star substrate (red), which is taken up across the cell membrane (CM) and the OFM.

Expanding the SNAP technology in *C. difficile* to the CLIP-tag

In *B. subtilis,* fluorescent proteins are often combined allowing colocalization and coexpression studies in a single cell. To allow similar studies, a variant of the SNAP-tag, named the CLIP-tag was created, engineering the substrate specificity of the SNAP-tag (Gautier *et al.*, 2008). The fact that the SNAP-tag specifically reacts with BG derivatives while the CLIP-tag specifically reacts with BC derivatives (Figure 2.1B), allows the simultaneous labeling of SNAP and CLIP proteins in the same cell. Recentlly, a faster labeling variant of SNAP-tag, termed $SNAP_f$, was generated, and extended to the CLIP-tag ($CLIP_f$) (Sun *et al.*, 2011). We have synthesized a variant of the *$CLIP_f$* gene, which we termed *$CLIP_f^{Cd}$*, preceded by the same ribosome-binding site that was used previously in the *$SNAP^{Cd}$* sequence. This cassette was placed under the control of the P_{tet} promoter, yielding pFT73 (Figure 2.2A and B, see also Materials and Methods). Plasmid pFT73 was introduced into the wild type strain 630Δ*erm*. The resulting strain AHCD714

was grown in BHI until mid-exponential phase, induced by addition of 250 ng/ml of ATc for 1 hour, and labeled with 250 nM of the CLIP-tag cell-permeable substrate TMR-Star for 30 minutes. To confirm that labeling of $CLIP_f$ with this substrate in *C. difficile* is specific of cells expressing $CLIP_f^{Cd}$, we have labeled both induced and non-induced cells of strain AHCD714, and the non-induced cells were further labeled with the DNA dye DAPI to allow differentiation from the induced ones. Cells were then mixed and examined by phase contrast and fluorescence microscopy (Figure 2.8). The results clearly suggest that labeling is specific to cells expressing $CLIP_f^{Cd}$, as there is no overlap between the red (CLIP-TMR-Star) and blue (DAPI) channels (Figure 2.8). Other CLIP-tag cell-permeable substrates commercially available were tested. However, the TMR-Star fluorophore was the one that produced the best results, consistent with the fact that its emission wavelength is the furthermost from the wavelength at which *C. difficile* cells usually autofluoresce (George *et al.*, 1979). Additional studies are needed to determine whether the conditions that lead to maximal accumulation of $CLIP_f$ are the same as the ones established for SNAP, and which concentrations of substrate are needed to quantitatively label the $CLIP_f$-tag. Nevertheless, our results open the way for uses of the SNAP- and CLIP-tags for orthogonal dual labeling studies.

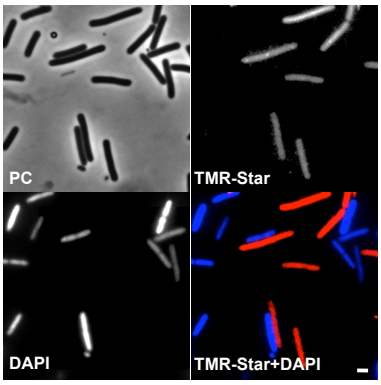


Figure 2.8. Microscopy analysis of *C. difficile* cells producing $CLIP_f$ under the control of P_{tet} (in pFT73). Cells were labeled with 250 nM of the TMR-Star cell-permeable CLIP substrate (red) and mixed with non-induced cells, labeled in the same conditions but subsequently stained with DAPI (blue), to allow distinction between the CLIP-labeled and the unlabeled cells. Scale bar: 1 µm.

DISCUSSION

Autofluorescent fusion proteins have revolutionized the way protein function is studied in the living cells (Shaner *et al.*, 2005; Giepmans *et al.*, 2006). However, oxygen requirement has limited its application to anaerobic organisms. One alternative approach for the fluorescent proteins uses covalent labeling of tagged proteins to fluorescent chemical probes (Keppler *et al.*, 2003; Johnsson and Johnsson, 2007; Hinner and Johnsson, 2010). Such tags, like the Halo-tag and SNAP-tag, which can be specifically labeled independently of the presence of oxygen, have been used to study biofilm formation in *Porphyromonas gingivalis* (Nicolle *et al.*, 2010), or protein traficking and organelle dynamics in the model organism *Giardia Intestinalis* (Regoes and Hehl, 2005; Martincova *et al.*, 2012). Here we have shown that the SNAP-tag can be used as fluorescent reporter for both gene expression and protein localization analysis in the anaerobe *C. difficile*. Fluorimager analysis combined with immunoblot experiments have shown that complete labeling of the all the SNAP produced under the control of P_{tet} can be achieved under optimal labeling conditions. This conclusion was only possible due to the slower migration of the SNAP protein when covalently attached to the TMR-Star substrate on SDS-PAGE gels, as revealed by the fluorescent scanning and immunoblot analysis. These results demonstrate that the SNAP-tag can be used as a quantitative reporter for analysis of gene expression at the single cell level in *C. difficile.* They also underscore the need to adjust labeling conditions to the specific promoter under study.

The P_{tet} system was the system used here to determine whether the SNAP-tag would work in *C. difficile*. So far studies involving this promoter took only into account the response of the whole cell population (Fagan and Fairweather, 2011; Govind and Dupuy, 2012). The use of a fluorescent reporter allowed us to analyse the response to this promoter at the single

cell level. By quantifying the fluorescence microscopy signals of the SNAP-TMR-Star in single cells, we have demonstrated that not all the cells respond in the same way to the presence of ATc, and that different amounts of the SNAP protein are produced from cell to cell. Moreover, our results also indicate that about 20% of the cells do not respond to the inducer. Our analysis therefore reveals new features of this inducible system that might need to be taken into account in the analysis of future experiments.

The SNAP-tag is highly versatile. For instance, labeling a SNAP-fusion protein at different time points with different fluorophores allows distinguishing between young and old copies of that same protein (Vivero-Pol *et al.*, 2005; Jansen *et al.*, 2007). This approach is an elegant alternative to the use of photo-activatable or -switchable autofluorescent proteins to track protein over time. An obvious application is the measurement of protein half-life in living cells (Bojkowska *et al.*, 2011). Additionally, quenched derivative probes that become fluorescent only upon reaction with the SNAP-tag have been recently developed, eliminating the need of a washing step to remove unreacted fluorescent probes during sequential labeling (Sun *et al.*, 2011). The use of such substrates combined with the faster labeling variant of SNAP, the $SNAP_f$ (Sun *et al.*, 2011), will broaden the applications of the SNAP technology in *C. difficile*. A recent, highly promising development is the use of a Split-SNAP system to monitor in time and space protein-protein interactions *in vivo* (Mie *et al.*, 2012). Lastly, substrates coupled to molecules other than fluorophores can be used and be an advantage to study different facets of protein function.

ACKNOWLEDGMENTS

The author of this dissertation participated in all the experiments described in this chapter. We thank Rita Tomé and Mónica Serrano for the

valuable help in some of the experiments. We also thank Nigel Minton and Neil Fairweather for the gift of plasmids, strains and helpful discussions and Lars Jansen for advice on the use of the SNAP-tag. This work was supported by grant ERA-PTG/SAU/0002/2008 (ERA-NET PathoGenoMics) and PEst-OE/EQB/LA0004/2011 from the "Fundação para a Ciência e a Tecnologia" (FCT). F.P. (SFRH/BD/45459/08) was the recipient of doctoral fellowship from the FCT.

REFERENCES

Bojkowska, K., Santoni de Sio, F., Barde, I., Offner, S., Verp, S., Heinis, C., Johnsson, K. & Trono, D. (2011). Measuring in vivo protein half-life. *Chem Biol* **18**, 805-15.

Camp, A. H. & Losick, R. (2008). A novel pathway of intercellular signalling in Bacillus subtilis involves a protein with similarity to a component of type III secretion channels. *Mol Microbiol* **69**, 402-17.

Deakin, L. J., Clare, S., Fagan, R. P., Dawson, L. F., Pickard, D. J., West, M. R., Wren, B. W., Fairweather, N. F., Dougan, G. & Lawley, T. D. (2012). The Clostridium difficile spo0A gene is a persistence and transmission factor. *Infect Immun* **80**, 2704-11.

Deneve, C., Janoir, C., Poilane, I., Fantinato, C. & Collignon, A. (2009). New trends in Clostridium difficile virulence and pathogenesis. *Int J Antimicrob Agents* **33 Suppl 1**, S24-8.

Donovan, C. & Bramkamp, M. (2009). Characterization and subcellular localization of a bacterial flotillin homologue. *Microbiology* **155**, 1786-99.

Drepper, T., Eggert, T., Circolone, F., Heck, A., Krauss, U., Guterl, J. K., Wendorff, M., Losi, A., Gartner, W. & Jaeger, K. E. (2007). Reporter proteins for in vivo fluorescence without oxygen. *Nat Biotechnol* **25**, 443-5.

Ernst, J. F. & Tielker, D. (2009). Responses to hypoxia in fungal pathogens. *Cell Microbiol* **11**, 183-90.

Fagan, R. & Fairweather, N. (2010). Dissecting the cell surface. *Methods Mol Biol* **646**, 117-34.

Fagan, R. P. & Fairweather, N. F. (2011). Clostridium difficile has two parallel and essential Sec secretion systems. *J Biol Chem* **286**, 27483-93.

Freeman, J., Bauer, M. P., Baines, S. D., Corver, J., Fawley, W. N., Goorhuis, B., Kuijper, E. J. & Wilcox, M. H. (2010). The changing epidemiology of Clostridium difficile infections. *Clin Microbiol Rev* **23**, 529-49.

Gautier, A., Juillerat, A., Heinis, C., Correa, I. R., Jr., Kindermann, M., Beaufils, F. & Johnsson, K. (2008). An engineered protein tag for multiprotein labeling in living cells. *Chem Biol* **15**, 128-36.

Gautier, A., Nakata, E., Lukinavicius, G., Tan, K. T. & Johnsson, K. (2009). Selective cross-linking of interacting proteins using self-labeling tags. *J Am Chem Soc* **131**, 17954-62.

George, W. L., Sutter, V. L., Citron, D. & Finegold, S. M. (1979). Selective and differential medium for isolation of Clostridium difficile. *J Clin Microbiol* **9**, 214-9.

Ghantoji, S. S., Sail, K., Lairson, D. R., DuPont, H. L. & Garey, K. W. (2010). Economic healthcare costs of Clostridium difficile infection: a systematic review. *J Hosp Infect* **74**, 309-18.

Giepmans, B. N., Adams, S. R., Ellisman, M. H. & Tsien, R. Y. (2006). The fluorescent toolbox for assessing protein location and function. *Science* **312**, 217-24.

Govind, R. & Dupuy, B. (2012). Secretion of Clostridium difficile toxins A and B requires the holin-like protein TcdE. *PLoS Pathog* **8**, e1002727.

Gronemeyer, T., Chidley, C., Juillerat, A., Heinis, C. & Johnsson, K. (2006). Directed evolution of O6-alkylguanine-DNA alkyltransferase for applications in protein labeling. *Protein Eng Des Sel* **19**, 309-16.

Heap, J. T., Pennington, O. J., Cartman, S. T., Carter, G. P. & Minton, N. P. (2007). The ClosTron: a universal gene knock-out system for the genus Clostridium. *J Microbiol Methods* **70**, 452-64.

Henriques, A. O. & Moran, C. P., Jr. (2007). Structure, assembly, and function of the spore surface layers. *Annu Rev Microbiol* **61**, 555-88.

Higgins, D. & Dworkin, J. (2012). Recent progress in Bacillus subtilis sporulation. *FEMS Microbiol Rev* **36**, 131-48.

Hilbert, D. W. & Piggot, P. J. (2004). Compartmentalization of gene expression during Bacillus subtilis spore formation. *Microbiol Mol Biol Rev* **68**, 234-62.

Hinner, M. J. & Johnsson, K. (2010). How to obtain labeled proteins and what to do with them. *Curr Opin Biotechnol* **21**, 766-76.

Hoffmann, C., Gaietta, G., Zurn, A., Adams, S. R., Terrillon, S., Ellisman, M. H., Tsien, R. Y. & Lohse, M. J. (2010). Fluorescent labeling of tetracysteine-tagged proteins in intact cells. *Nat Protoc* **5**, 1666-77.

Hussain, H. A., Roberts, A. P. & Mullany, P. (2005). Generation of an erythromycin-sensitive derivative of Clostridium difficile strain 630 (630Deltaerm) and demonstration that the conjugative transposon Tn916DeltaE enters the genome of this strain at multiple sites. *J Med Microbiol* **54**, 137-41.

Jansen, L. E., Black, B. E., Foltz, D. R. & Cleveland, D. W. (2007). Propagation of centromeric chromatin requires exit from mitosis. *J Cell Biol* **176**, 795-805.

Johnsson, N. & Johnsson, K. (2007). Chemical tools for biomolecular imaging. *ACS Chem Biol* **2**, 31-8.

Keppler, A., Gendreizig, S., Gronemeyer, T., Pick, H., Vogel, H. & Johnsson, K. (2003). A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nat Biotechnol* **21**, 86-9.

Keppler, A., Kindermann, M., Gendreizig, S., Pick, H., Vogel, H. & Johnsson, K. (2004a). Labeling of fusion proteins of O6-alkylguanine-DNA alkyltransferase with small molecules in vivo and in vitro. *Methods* **32**, 437-44.

Keppler, A., Pick, H., Arrivoli, C., Vogel, H. & Johnsson, K. (2004b). Labeling of fusion proteins with synthetic fluorophores in live cells. *Proc Natl Acad Sci U S A* **101**, 9955-9.

Kindermann, M., Sielaff, I. & Johnsson, K. (2004). Synthesis and characterization of bifunctional probes for the specific labeling of fusion proteins. *Bioorg Med Chem Lett* **14**, 2725-8.

Lemercier, G., Gendreizig, S., Kindermann, M. & Johnsson, K. (2007). Inducing and sensing protein--protein interactions in living cells by selective cross-linking. *Angew Chem Int Ed Engl* **46**, 4281-4.

Los, G. V., Encell, L. P., McDougall, M. G., Hartzell, D. D., Karassina, N., Zimprich, C., Wood, M. G., Learish, R., Ohana, R. F., Urh, M., Simpson, D., Mendez, J., Zimmerman, K., Otto, P., Vidugiris, G., Zhu, J., Darzins, A., Klaubert, D. H., Bulleit, R. F. & Wood, K. V. (2008). HaloTag: a novel protein labeling technology for cell imaging and protein analysis. *ACS Chem Biol* **3**, 373-82.

Maroo, S. & Lamont, J. T. (2006). Recurrent clostridium difficile. *Gastroenterology* **130**, 1311-6.

Martincova, E., Voleman, L., Najdrova, V., De Napoli, M., Eshar, S., Gualdron, M., Hopp, C. S., Sanin, D. E., Tembo, D. L., Van Tyne, D., Walker, D., Marcincikova, M., Tachezy, J. & Dolezal, P. (2012). Live imaging of mitosomes and hydrogenosomes by HaloTag technology. *PLoS One* **7**, e36314.

Meisner, J., Wang, X., Serrano, M., Henriques, A. O. & Moran, C. P., Jr. (2008). A channel connecting the mother cell and forespore during bacterial endospore formation. *Proc Natl Acad Sci U S A* **105**, 15100-5.

Mie, M., Naoki, T., Uchida, K. & Kobatake, E. (2012). Development of a split SNAP-tag protein complementation assay for visualization of protein-protein interactions in living cells. *Analyst* **137**, 4760-5.

Nicholson, W. L., Munakata, N., Horneck, G., Melosh, H. J. & Setlow, P. (2000). Resistance of Bacillus endospores to extreme terrestrial and extraterrestrial environments. *Microbiol Mol Biol Rev* **64**, 548-72.

Nicolle, O., Rouillon, A., Guyodo, H., Tamanai-Shacoori, Z., Chandad, F., Meuric, V. & Bonnaure-Mallet, M. (2010). Development of SNAP-tag-mediated live cell labeling as an alternative to GFP in Porphyromonas gingivalis. *FEMS Immunol Med Microbiol* **59**, 357-63.

Redelings, M. D., Sorvillo, F. & Mascola, L. (2007). Increase in Clostridium difficile-related mortality rates, United States, 1999-2004. *Emerg Infect Dis* **13**, 1417-9.

Regoes, A. & Hehl, A. B. (2005). SNAP-tag mediated live cell labeling as an alternative to GFP in anaerobic organisms. *Biotechniques* **39**, 809-10, 812.

Reid, B. G. & Flynn, G. C. (1997). Chromophore formation in green fluorescent protein. *Biochemistry* **36**, 6786-91.

Rupnik, M., Wilcox, M. H. & Gerding, D. N. (2009). Clostridium difficile infection: new developments in epidemiology and pathogenesis. *Nat Rev Microbiol* **7**, 526-36.

Serrano, M., Real, G., Santos, J., Carneiro, J., Moran, C. P., Jr. & Henriques, A. O. (2011). A negative feedback loop that limits the ectopic activation of a cell type-specific sporulation sigma factor of Bacillus subtilis. *PLoS Genet* **7**, e1002220.

Shaner, N. C., Steinbach, P. A. & Tsien, R. Y. (2005). A guide to choosing fluorescent proteins. *Nat Methods* **2**, 905-9.

Sun, X., Zhang, A., Baker, B., Sun, L., Howard, A., Buswell, J., Maurel, D., Masharina, A., Johnsson, K., Noren, C. J., Xu, M. Q. & Correa, I. R., Jr. (2011). Development of SNAP-tag fluorogenic probes for wash-free fluorescence imaging. *Chembiochem* **12**, 2217-26.

Tielker, D., Eichhof, I., Jaeger, K. E. & Ernst, J. F. (2009). Flavin mononucleotide-based fluorescent protein as an oxygen-independent reporter in Candida albicans and Saccharomyces cerevisiae. *Eukaryot Cell* **8**, 913-5.

Vivero-Pol, L., George, N., Krumm, H., Johnsson, K. & Johnsson, N. (2005). Multicolor imaging of cell surface proteins. *J Am Chem Soc* **127**, 12770-1.

Wang, K. H., Isidro, A. L., Domingues, L., Eskandarian, H. A., McKenney, P. T., Drew, K., Grabowski, P., Chua, M. H., Barry, S. N., Guan, M., Bonneau, R., Henriques, A. O. & Eichenberger, P. (2009). The coat morphogenetic protein SpoVID is necessary for spore encasement in Bacillus subtilis. *Mol Microbiol* **74**, 634-49.

Chapter 3

The pathway of spore differentiation in *C. difficile*

This chapter contains data published in: Pereira, F. C., Saujet, L., Tomé, A. R., Serrano, M., Monot, M., Couture-Tosi, E., Martin-Verstraete, I., Dupuy, B., and Henriques, A. O. (2013). The spore differentiation pathway in the enteric pathogen *Clostridium difficile*. *PLoS Genet* 9: e1003782.

SUMMARY

Endosporulation is an ancient bacterial developmental program that culminates with the differentiation of a highly resistant endospore. In the model organism *B. subtilis*, gene expression in the forespore and in the mother cell, the two cells that participate in endospore development, is governed by cell type-specific RNA polymerase sigma subunits. σ^F in the forespore, and σ^E in the mother cell control early stages of development and are replaced, at later stages, by σ^G and σ^K, respectively. Starting with σ^F, the activation of the sigma factors is sequential, requires the preceding factor, and involves cell-cell signaling pathways that operate at key morphological stages. Here, we have studied the function and regulation of the sporulation sigma factors in the intestinal pathogen *C. difficile*, an obligate anaerobe in which the endospores are central to the infectious cycle. The morphological characterization of mutants for the sporulation sigma factors, in parallel with use of a fluorescence reporter for single cell analysis of gene expression, unraveled important deviations from the *B. subtilis* paradigm. While the main periods of activity of the sigma factors are conserved, we show that the activity of σ^E is partially independent of σ^F, that σ^G activity is not dependent on σ^E, and that the activity of σ^K does not require σ^G. We also show that σ^K is not strictly required for heat resistant spore formation. In all, our results indicate reduced temporal segregation between the activities of the early and late sigma factors, and reduced requirement for the σ^F-to-σ^E, σ^E-to-σ^G, and σ^G-to-σ^K cell-cell signaling pathways. Nevertheless, our results support the view that the top level of the endosporulation network is conserved in evolution, with the sigma factors acting as the key regulators of the pathway, established some 2.5 billion years ago upon its emergence at the base of the Firmicutes Phylum.

INTRODUCTION

Bacterial spores, as those formed by species of the well-known *Bacillus* and *Clostridium* genera, but also by many other groups within the Firmicutes phylum, resist to extremes of physical and chemical parameters that would rapidly destroy the vegetative cells, and are the most resistant cellular structure known (Nicholson *et al.*, 2000). Infection by the intestinal human and animal pathogen *C. difficile*, an obligate anaerobe, and the subject of the present investigation, often also starts with the ingestion of spores (Deneve *et al.*, 2009; Rupnik *et al.*, 2009) Ingested spores of this organism germinate in the colon, to establish a population of vegetative cells that will produce two potent cytotoxins and more spores (Rupnik *et al.*, 2009; Carter *et al.*, 2012; Burns *et al.*, 2010; Sarker and Paredes-Sabja, 2012). Therefore C. difficile spores are central for infection and disease transmission (Rupnik *et al.*, 2009; Carter *et al.*, 2012; Burns *et al.*, 2010; Sarker and Paredes-Sabja, 2012; Deakin *et al.*, 2012). A fraction of these spores can also remain latent in the gut, being responsible for the persistence of the organism inside the host and recurrence.

The basic spore plan is conserved (Henriques and Moran, 2007; McKenney *et al.*, 2012; de Hoon *et al.*, 2010). The genome is deposited in a central compartment delimited by a lipid bilayer with a layer of peptidoglycan (PG) apposed to its external leaflet. This layer of PG, known as the germ cell wall, will serve as the wall of the outgrowing cell that forms when the spore completes germination. The germ cell wall is encased in a thick layer of a modified form of PG, the cortex, essential for the acquisition and maintenance of heat resistance (Henriques and Moran, 2007; McKenney *et al.*, 2012). The cortex is wrapped by a multiprotein coat, which protects it from the action of PG-breaking enzymes produced by host organisms or predators (Henriques and Moran, 2007; McKenney *et al.*, 2012). In some

species, including the pathogens *B. anthracis*, *B. cereus* and *C. difficile*, the coat is further enclosed within a structure known as the exosporium. The coat and the exosporium, when present, mediate the immediate interactions of the spore with the environment, including the interaction with small molecules that trigger germination (Oliva *et al.*, 2009; Henriques and Moran, 2007; Panessa-Warren *et al.*, 2007; Paredes-Sabja and Sarker, 2012; Paredes-Sabja *et al.*, 2012).

The process of spore differentiation has been extensively studied in the model organism *B. subtilis* (Hilbert and Piggot, 2004; Higgins and Dworkin, 2012). Rod-shaped vegetative cells, growing by binary fission, will switch to an asymmetric (polar) division when facing severe nutritional stress. Polar division yields a larger mother cell and a smaller forespore, the future spore. The mother cell then engulfs the forespore. This process, akin to phagocytosis and a hallmark of endosporulation, isolates the forespore from the surrounding medium, and releases it as a cell, surrounded by a double membrane, within the mother cell cytoplasm (Hilbert and Piggot, 2004; Higgins and Dworkin, 2012). With the exception of the germ cell wall, which is formed from the forespore, the assembly of the main spore protective structures is mostly a function of the mother cell (Henriques and Moran, 2007; McKenney *et al.*, 2012). At the end of the process, and following a period of spore maturation, the mother cell undergoes autolysis, to release the finished spore. For the organisms that have been studied to date, mostly by transmission electron microscopy (TEM), this basic sequence of morphological events appears conserved (Henriques and Moran, 2007; Hilbert and Piggot, 2004).

The developmental regulatory network of sporulation shows a hierarchical organization and functional logic (de Hoon *et al.*, 2010). A master regulatory protein, Spo0A, activated by phosphorylation, governs entry into sporulation, including the switch to asymmetric division (Hilbert and Piggot, 2004; Galperin *et al.*, 2012). Gene expression in the forespore

and mother cell is controlled by 4 cell type-specific sigma factors, which are sequentially activated, alternating between the two cells. σ^F and σ^E control the early stages of development in the forespore and the mother cell, respectively, and are replaced by σ^G and σ^K when engulfment of the forespore is completed (Hilbert and Piggot, 2004; Higgins and Dworkin, 2012; Piggot and Hilbert, 2004). Activation of the sporulation sigma factors coincides with the completion of key morphological intermediates in the process, at which stages cell-cell signaling events further allow the alignment of the forespore and mother cell programs of gene expression. The result is the coordinated deployment of the forespore and mother cell lines of gene expression, in close register with the course of cellular morphogenesis (Hilbert and Piggot, 2004; Higgins and Dworkin, 2012; Piggot and Hilbert, 2004). A large number of genes of *B. subtilis*, distributed in the four cell type-specific regulons, participate in spore morphogenesis (Galperin *et al.*, 2012; Traag *et al.*, 2012; Miller *et al.*, 2012; Abecasis *et al.*, 2013). The key regulatory factors, Spo0A and the sporulation sigma factors, which define the highest level in the functional and evolutionary hierarchy of the sporulation network, are conserved in sporeformers. The conservation of the sporulation sigma factors has suggested that their role and sequential activation is also maintained across species (de Hoon *et al.*, 2010). However, recent studies have revealed differences in the roles and time of activity of the sigma factors during spore morphogenesis in several Clostridial species (Paredes *et al.*, 2005; Harry *et al.*, 2009; Li and McClane, 2010; Tracy *et al.*, 2011; Jones *et al.*, 2008; Jones *et al.*, 2011; Bi *et al.*, 2011). For instance, σ^F and σ^E are active prior to asymmetric division in *C. acetobutylicum* and *C. perfringens* (Li and McClane, 2010; Tracy *et al.*, 2011; Jones *et al.*, 2011; Bi *et al.*, 2011). Also, σ^K, which in *B. subtilis* controls late stages of morphogenesis in the mother cell, is active in pre-divisional cells of C. *perfringens* and *C. botulinum* (Harry *et al.*, 2009; Kirk *et al.*, 2012). Collectively, and relative to the aerobic *Bacilli*, the *Clostridia* represent an older group within the

Firmicutes phylum, at the base of which endosporulation has emerged some 2.5 billion years ago, before the initial rise in oxygen levels (Galperin *et al.*, 2012; Traag *et al.*, 2012; Miller *et al.*, 2012; Abecasis *et al.*, 2013; Stragier, 2002).

Despite the importance of *C. difficile* for human health and activities, and the central role of sporulation in the infection cycle, a cytological and molecular description of sporulation has been lacking. Here, we have combined cytological and genetics methodologies to define the sequence of sporulation events in *C. difficile* and the function of the cell type-specific sigma factors. In addition, by using a fluorescent reporter for studies of gene expression at the single cell level (described in Chapter 2), we were able to correlate the expression and activity of the sporulation-specific sigma factors with the course of morphogenesis. A key observation is that during *C. difficile* sporulation the forespore and mother cell programs of gene expression are less tightly coupled. Our study also provides a platform for additional studies of the regulatory network and for integrating the expression and function of the effector genes, many of which will be species-specific, and possibly related to host colonization and transmission.

MATERIALS AND METHODS

Strains and general techniques. Bacterial strains and their relevant properties are listed in Table 3.1. The *Escherichia coli* strain DH5α (Bethesda Research laboratories) was used for molecular cloning. Luria-Bertani medium was routinely used for growth and maintenance of *E. coli* and *B. subtilis.* The *B. subtilis* strains are congenic derivatives of the Spo$^+$ strain MB24 (*trpC2 metC3*). Sporulation of *B. subtilis* was induced by growth and exhaustion in Difco sporulation medium (DSM) (Henriques *et al.*, 1995). When indicated, ampicillin (100 μg/ml) or chloramphenicol (15 μg/ml) was

added to the culture medium. The *Bacillus cereus* strain used in this study is listed in Table 3.1. Sporulation in *B. cereus* was induced by growth in Leighton-Doi medium at 30ºC, for 48h (Leighton and Doi, 1971). The *C. difficile* strains used in this study are congenic derivatives of the wild-type strain 630Δ*erm* (Hussain *et al.*, 2005) and were routinely grown anaerobically (5% H2, 15% CO2, 80% N2) at 37ºC in brain heart infusion (BHI) medium (Difco), BHIS [BHI medium supplemented with yeast extract (5 mg/ml) and L-cysteine (0.1%)], or SM medium [for 1l: 90 g Bacto-tryptone, 5g Bacto-peptone, 1 g $(NH_4)_2SO_4$ and 1.5 g Tris base)] (Wilson *et al.*, 1982). Sporulation assays were performed in SM medium (Wilson *et al.*, 1982). When necessary, cefoxitin (25 μg/ml), thiamphenicol (15 μg/ml), or erythromycin (5 μg/ml) was added to *C. difficile* cultures.

Construction of gene knockout mutants in *C. difficile*. The ClosTron gene knockout system (Heap *et al.*, 2007) was used to inactivate the *sigF*, *sigE*, *sigG* and *sigK* genes. Primers to retarget the group II intron of pMTL007 to these genes (all oligonucleotide primers used in this work are listed in Table A2) were designed with the Targetron design software (Sigma-Aldrich). The PCR primer sets were used with the EBS universal primer and intron template DNA to generate by overlap extension PCR a 353-bp product for each gene that allows intron retargeting. These PCR products were cloned between the HindIII and BsrGI restriction sites of pMTL007 giving plasmids pMTL007::Cdi-s*igF*-459s, pMTL007::Cdi-*sigE*-453s, pMTL007::Cdi-*sigG*-546s and pMTL007::Cdi-*sigK*-102s (all plasmids are listed in Table A1). DNA sequencing was performed to verify plasmid constructs, using the pMTL007-specific primers pMTL007-F and pMTL007-R. The derivative pMTL007 plasmids were transformed into *E. coli* HB101 (RP4) and subsequently mated with *C. difficile* 630Δ*erm* (Heap *et al.*, 2007). *C. difficile* transconjugants were selected by subculturing on BHI agar containing thiamphenicol (15 μg/ml) and cefoxitin (25 μg/ml) and then

plated on BHI agar containing erythromycin (5 μg/ml). This produced the 630Δ*erm* derivatives AHCD533 (*sigF*::*erm*), AHCD532 (*sigE*::*erm*), AHCD534 (*sigG*::*erm*) and AHCD535 (*sigK*::*erm*), respectively (Table 3.1). Chromosomal DNA was isolated from the transconjugants, and the structure of the various mutants in the vicinity of the disrupted gene was verified by PCR.

Table 3.1. Bacterial Strains used in this study

Strain	Relevant Properties	Origin
E.coli		
Dh5α		Invitrogen
HB101 (RP4)		Laboratory stock
AHCD018	HB101 (RP4)/ pMTL007::Cdi-*sigE*-453s	This work
AHCD019	HB101 (RP4)/ pMTL007::Cdi-*sigF*-459s	"
AHCD020	HB101 (RP4)/ pMTL007::Cdi-*sigG*-546s	"
AHCD021	HB101 (RP4)/ pMTL007::Cdi-*sigK*-102s	"
AHCD072	HB101 (RP4)/ pFT32	"
AHCD073	HB101 (RP4)/ pFT38	"
AHCD074	HB101 (RP4)/ pFT39	"
AHCD075	HB101 (RP4)/ pFT40	"
AHCD076	HB101 (RP4)/ pMTL84121	"
AHCD077	HB101 (RP4)/ pFT42	"
AHCD095	HB101 (RP4)/ pFT46	"
AHCD108	HB101 (RP4)/ pFT49	"
AHCD109	HB101 (RP4)/ pFT50	"
AHCD110	HB101 (RP4)/ pFT51	"
AHCD112	HB101 (RP4)/ pFT48	"
AHCD114	HB101 (RP4)/ pFT53	"
AHCD115	HB101 (RP4)/ pFT54	"
AHCD116	HB101 (RP4)/ pFT55	"
AHCD131	HB101 (RP4)/ pFT47	"
AHCD142	HB101 (RP4)/ pFT64	"
AHCD149	HB101 (RP4)/ pFT63	"
AHCD150	HB101 (RP4)/ pFT69	"
B. subtilis		
MB24	*trpC2 metC3* /Spo$^+$	Laboratory stock
AH10170	MB24 *tgl-cfp*	"
B. cereus		
ATCC4342	*B. cereus* wild type strain	(Schuch *et al*., 2002)
C. difficile		
630Δ*erm*	*C. difficile* 630Δ*erm*	(Hussain *et al*., 2005)
630Δ*erm spo0A*	630Δ*erm spo0A*::intron *ermB*	(Saujet *et al*., 2011)
AHCD532	630Δ*erm sigE*::intron *ermB*	This work
AHCD533	630Δ*erm sigF*::intron *ermB*	"
AHCD534	630Δ*erm sigG*::intron *ermB*	"
AHCD535	630Δ*erm sigK*::intron *ermB*	"
AHCD543	630Δ*erm* containing pMTL84121	"
AHCD544	AHCD533 containing pMTL84121	"

Table 3.1. Bacterial Strains used in this study

Strain	Relevant Properties	Origin
AHCD545	AHCD532 containing pMTL84121	This work
AHCD546	AHCD534 containing pMTL84121	“
AHCD547	AHCD535 containing pMTL84121	“
AHCD548	AHCD533 containing pFT32	“
AHCD549	AHCD532 containing pFT39	“
AHCD550	AHCD534 containing pFT40	“
AHCD551	AHCD535 containing pFT38	“
AHCD577	AHCD535 containing pFT42	“
AHCD586	630Δ*erm* containing pFT46	“
AHCD597	630Δ*erm* containing pFT49	“
AHCD598	630Δ*erm* containing pFT48	“
AHCD600	630Δ*erm* containing pFT50	“
AHCD601	630Δ*erm* containing pFT51	“
AHCD602	630Δ*erm* containing pFT53	“
AHCD603	630Δ*erm* containing pFT54	“
AHCD604	630Δ*erm* containing pFT55	“
AHCD610	AHCD533 containing pFT53	“
AHCD611	AHCD533 containing pFT54	“
AHCD612	AHCD533 containing pFT55	“
AHCD613	AHCD532 containing pFT53	“
AHCD614	AHCD532 containing pFT54	“
AHCD615	AHCD532 containing pFT55	“
AHCD619	AHCD534 containing pFT55	“
AHCD621	AHCD535 containing pFT55	“
AHCD630	AHCD533 containing pFT48	“
AHCD631	AHCD533 containing pFT49	“
AHCD632	AHCD533 containing pFT50	“
AHCD633	AHCD533 containing pFT51	“
AHCD634	AHCD532 containing pFT48	“
AHCD635	AHCD532 containing pFT49	“
AHCD636	AHCD532 containing pFT50	“
AHCD637	AHCD532 containing pFT51	“
AHCD640	AHCD534 containing pFT50	“
AHCD641	AHCD534 containing pFT51	“
AHCD644	AHCD535 containing pFT50	“
AHCD645	AHCD535 containing pFT51	“
AHCD646	630Δ*erm* containing pFT47	“
AHCD656	630Δ*erm* containing pFT64	“
AHCD674	AHCD534 containing pFT64	“
AHCD675	AHCD535 containing pFT64	“
AHCD678	630Δ*erm* containing pFT63	“
AHCD679	AHCD534 containing pFT63	“
AHCD680	AHCD535 containing pFT63	“
AHCD683	AHCD534 containing pFT53	“
AHCD684	AHCD535 containing pFT53	“
AHCD685	AHCD534 containing pFT54	“
AHCD686	AHCD535 containing pFT54	“
AHCD692	630Δ*erm spo0A* containing pFT48	“
AHCD693	630Δ*erm spo0A* containing pFT49	“
AHCD695	630Δ*erm* containing pFT69	“
AHCD696	AHCD533 containing pFT69	“
AHCD697	AHCD532 containing pFT69	“
AHCD698	AHCD534 containing pFT69	“
AHCD695	AHCD535 containing pFT69	“

Complementation in *C. difficile*. To complement the *sigF* mutation, the coding sequence of the *sigF* gene (765 bp length) and its expected promoter region comprising 514 bp upstream of *spoIIAA* were independently amplified by PCR using primer pairs Prom_CDsigF Fw/Prom_CD sigF Rev and CDsigF Fw/CDsigF-EcoRI Rev (Table A2). The two pieces were joined into a 1279 bp fragment by splicing by overlaping extension (SOE) PCR and cloned between the NotI and EcoRI sites of pMTL84121 (Heap *et al.*, 2009) to produce pFT32 (Table A1). To complement the *sigE* mutation, the entire coding sequence of *sigE* gene (742 bp) and its promoter region (459 bp upstream of *spoIIGA*) were independently PCR amplified using primer pairs Prom_CDsigE Fw/Prom_CD sigE Rev and CDsigE Fw/CDsigE-EcoRI Rev, joined by SOE PCR and the resulting 1143 bp fragment cloned between the NotI and EcoRI sites of pMTL84121 to create pFT39. To complement the *sigG* mutation, a fragment encompassing the coding sequence of *sigG* (774 bp) and 459 bp of its regulatory region was PCR amplified using primer pair Prom_CDsigG Fw/CDsigG-EcoRI Rev. The resulting 1233 bp fragment was cloned between the NotI and EcoRI sites of pMTL84121 to yield pFT40. Finally, to complement the *sigK* mutation two constructs were generated. First, the interrupted 5´ coding sequence of *sigK* gene (404 bp downstream of the translational start site) and its expected promoter region comprising the 415 bp upstream of *sigK* gene was PCR amplified using primers Prom_CDsigK Fw and CDsigK5' Rev yielding a 819 bp fragment. In a second PCR, the region containing the interrupted 3´ coding sequence of *sigK*, the gene (CD1231) coding for the recombinase, and a 466 bp fragment upstream of this gene was PCR amplified using primer pair CDsigK3' Fw/CDsigK3'_EcoRI Rev. This fragment, of 2285 bp was joined to the 819 bp fragment by SOE PCR and cloned between the NotI and EcoRI sites of pMTL84121 to produce pFT38. A second construct was generated in which an intact *sigK* gene with 419 bp of its regulatory region was amplified using primer pair Prom_CDsigK Fw/ CDsigK3'_EcoRI Rev from DNA prepared from 48h cultures in BHI, a time at

which most of the cells already bear a re-arranged *sigK* gene. The resulting 1061 bp fragment was cloned between the NotI/EcoRI sites of pMTL84121 yielding pFT42. Using the *E. coli* HB101 (RP4) strain as donor, pFT32, pFT38, pFT39, pFT40 and pFT42 were transferred by conjugation into the various *sig*::*erm* mutants of *C. difficile* 630Δ*erm*, giving strains AHCD548, AHCD551, AHCD549, AHCD550 and AHCD577 (Table 3.1).

Southern Blot. For Southern blot analysis, 6 μg of genomic DNA from *C. difficile* strain 630Δ*erm* and the congenic *sigF*, *sigE*, *sigG* or *sigK* mutant strains were digested to completion with HindIII. The Southern blot probe was generated by PCR using pMTL007 plasmid as a template and primer pair OBD522 and OBD523 (Table A2), yielding a 374 bp PCR product that hybridizes within the group II intron. Southern blot analyses were performed using Amersham ECL Direct Nucleic Acid labeling and detection reagents, according to the manufacturer's guidelines. The hybridization signal was detected using Super Signal West Femto Maximum Sensitivity Substrate (Thermo Scientific).

Sporulation assays. Overnight cultures grown at 37°C in BHI were used to inoculate SM medium (at a dilution of 1:200). At specific time points, 1 ml of culture was withdrawn, serially diluted in phosphate-buffered saline (PBS; 137 mM NaCl, 10 mM Phosphate, 2.7 mM KCl, pH 7.4), and plated before and after heat treatment (30min at 60ºC), to determine the total and heat-resistant colony forming units (CFU). The samples were plated onto BHI plates supplemented with 0.1% taurocholate (Sigma-Aldrich), to promote efficient spore germination (Wilson *et al.*, 1982). The percentage of sporulation was determined as the ratio between the number of spores/ml and the total number of viable bacteria/ml times 100.

Spore production, purification and decoating. For spore production, 5 ml of BHI media was inoculated with an isolated colony of *C. difficile* 630Δ*erm* and cultured overnight at 37ºC in anaerobic conditions. 100 ml of fresh BHI media was then inoculated with 1 ml of the overnight culture and incubated at 37ºC under anaerobic conditions for 10 days. Cells were collected by centrifugation at 4800xg, ressuspended in cold water and stored over night at 4ºC. Spores were then purified with a 40-50% Gastrografin (Schering) step gradient, as previously described (Henriques *et al.*, 1995). The pellet was ressuspended in cold water, washed 10 times with cold water, and stored at 4ºC until further use. To remove the spore coat, *C. difficile* purified spores were resuspended in 50 ml of extraction buffer (0.1 M NaOH, 0.1 M NaCl, 1% SDS, 0.1 M DTT) to a final OD_{600nm} of 4.0. Spores were then incubated at 70ºC for 30 minutes, extensively washed with water, to remove all traces of NaOH and SDS, and then stained with the FM4-64 dye as described below.

RNA isolation and real-time quantitative RT-PCR. In a preliminary set of experiments, we defined the time for which the difference in expression of a selected target gene between the wild type and the corresponding mutant strain was highest. To study σ^E- or σ^F-dependent control, we harvested cells from 630Δ*erm*, *sigF* and *sigE* mutants after 14 h of growth in SM medium. Strain 630Δ*erm* and the *sigG* or the *sigK* mutants were harvested after 19 h (630Δ*erm*, *sigG* mutant) and 24 h (630Δ*erm*, *sigK* mutant) of growth in SM medium. Total RNA was extracted from at least two independent cultures. After centrifugation, the culture pellets were resuspended in RNApro™ solution (MP Biomedicals) and RNA extracted using the FastRNA Pro Blue Kit, according to the manufacturer's instructions. The RNA quality was determined using RNA 6000 Nano Reagents (Agilent). For quantitative RT-PCR experiments, 1 μg of total RNA was heated at 70°C for 10 min along with 1 μg of hexamer oligonucleotide primers $p(dN)_6$

(Roche). After slow cooling, cDNAs were synthesized as previously described (Saujet *et al.*, 2011). The reverse transcriptase was inactivated by incubation at 85°C for 5 min. Real-time quantitative RT-PCR was performed twice in a 20 µl reaction volume containing 20 ng of cDNAs, 10 µl of FastStart SYBR Green Master mix (ROX, Roche) and 200 nM gene-specific primers in a AB7300 real-time PCR instrument (Applied Biosystems). The primers used for each marker are listed in Table A2. Amplification and detection were performed as previously described In each sample, the quantity of cDNAs of a gene was normalized to the quantity of cDNAs of the DNApolIII gene (Saujet *et al.*, 2011). The relative transcript changes were calculated using the $2^{-\Delta\Delta Ct}$ method as described (Saujet *et al.*, 2011).

Transcriptional *SNAP*^{Cd} fusions. To construct *sigF*-, *sigE*-, *sigG*- and *sigK*- transcriptional *SNAP*^{Cd} fusions, the promoter regions of these genes were PCR-amplified using genomic DNA from strain 630Δ*erm* and primer pairs PCDsigF-Fw and PCDsigF-XhoI Rev, PCDsigE-Fw and PCDsigE-XhoI Rev, PCDsigG-Fw and PCDsigG-XhoI Rev and PCDsigK5´-Fw and PCDsigK-XhoI Rev, to produce 539, 482, 494 and 437 bp products, respectively (Table A2). These were inserted between the EcoRI and XhoI sites of pFT47 to create pFT48, pFT49, pFT50 and pFT51 (Table A1). To monitor the activities of σ^F, σ^E, σ^G and σ^K, the promoter regions for the *gpr*, *spoIIIAA*, *sspA* and *cotE* genes were PCR-amplified from 630Δ*erm* genomic DNA using primer pairs Pgpr-EcoRI Fw and Pgpr-XhoI Rev, PspoIIIAA-EcoRI Fw and PspoIIIAA-XhoI Rev, PsspA-EcoRI Fw and PsspA-XhoI Rev, and PcotE-EcoRI Fw and PcotE-XhoI Rev. The resulting 433, 533, 489, and 303 bp fragments, respectively, were inserted between the EcoRI and XhoI sites of pFT47 to create pFT53, pF54, pFT55, and pFT69. The absence of unwanted mutations was verified by sequencing the insert in all the plasmids. All plasmids bearing promoter-*SNAP*^{Cd} fusions were introduced into *E. coli* HB101 (RP4) and then transferred to *C. difficile* 630Δ*erm* and *sigF* (AHCD533), *sigE* (AHCD532), *sigG*

(AHCD534) and *sigK* (AHCD535) mutant strains by conjugation (Heap *et al.,* 2007) (Table 3.1).

Translational *SNAP^{Cd}* fusions. The promoter and coding regions of *cotB* (*CD1511*) and *cotE* (*CD1433*) were amplified from 630Δ*erm* genomic DNA using primer pairs PcotB-EcoRI Fw and cotB-Linker-Rev, and PcotE-EcoRI Fw and cotE-Linker-Rev (Table A2). The resulting PCR products were inserted between the EcoRI and BamHI sites of pFT58, yielding pFT63 and pFT64, respectively (Table A1). In the final constructs, the *cotB* and *cotE* sequences are separated from the SNAP-coding sequence by a sequence coding for a 9 amino acid linker (LGGGGSAAA; see also Chapter 2). Plasmids pFT63 and pFT64 were then transferred to *C. difficile* 630Δ*erm*, *sigG* (AHCD534) and *sigK* (AHCD535) mutant strains by conjugation (Heap *et al.*, 2007) (Table 3.1).

SNAP labeling and analysis. Whole cell extracts were obtained by withdrawing 10 ml samples from *C. difficile* cultures in brain heart infusion (BHI) for the P_{tet}-SNAP-bearing strains, or in SM medium for the sporulation experiments, at the desired times. The extracts were prepared immediately following labeling with 250 nM of the TMR-Star substrate (New England Biolabs), for 30 min in the dark. Following labeling, the cells were collected by centrifugation (4000xg, for 5 min at 4ºC), the cell sediment was washed with phosphate-buffered saline (PBS) and resuspended in 1 ml French press buffer (10 mM Tris pH 8.0, 10 mM MgCl2, 0.5 mM EDTA, 0.2 mM NaCl, 10% Glycerol, 1 mM PMSF). The cells were lysed using a French pressure cell (18000 lb/in^2). Proteins in the extracts were resolved on 15% SDS-PAGE gels. The gels were first scanned in a Fuji TLA-5100 fluorimager, and then subject to immunoblot analysis as described before (Serrano *et al.*, 2011). The anti-SNAP antibody (New England Biolabs) was used at a 1:1000 dilution, and a rabbit secondary antibody conjugated to horseradish

peroxidase (Sigma) was used at dilution 1:10000. The immunoblots were developed with enhanced chemiluminescence reagents (Amersham Pharmacia Biotech).

Microscopy and image analysis. Samples of 1 ml were withdrawn from BHI or SM cultures at the desired times following inoculation, and the cells collected by centrifugation (4000xg for 5 min). The cells were washed with 1ml of PBS, and ressuspended in 0.1 ml of PBS supplemented with the lipophilic styryl membrane dye *N*-(3-triethylammoniumprpyl)-4-(*p*-diethylaminophenyl-hexatrienyl) pyridinium dibromide (FM4-64; 10 μg/ml) (Pogliano *et al.*, 1999) (Vida and Emr, 1995) and the DNA stain DAPI (4',6-diamidino-2-phenylindole; 50 μg/ml) (both from Molecular Probes, Invitrogen). For the live/dead assay, samples were collected as described above, ressuspended in 0.05 ml of PBS and mixed with an equal volume of 2x LIVE/DEAD BacLight 2x staining reagent mixture (Molecular Probes, Invitrogen) containing Propidium iodide (30 μM final concentration) and syto9 (6 μM final concentration).

For SNAP labeling experiments, cells in culture samples were labeled with TMR-Star (as above), collected by centrifugation (4000xg, 3 min, at room temperature), washed four times with 1 ml of PBS, and finally ressupended in 1ml of PBS containing the membrane dye Mitotracker Green (0.5 μg/ml) (Molecular Probes, Invitrogen).

For phase contrast and fluorescence microscopy, cells were mounted on 1.7% agarose coated glass slides and observed on a Leica DM6000B microscope equipped with a phase contrast Uplan F1 100x objective and a CCD Ixon^{EM} camera (Andor Technologies) (Serrano *et al.*, 2011). Images were acquired and analyzed using the Metamorph software suite version 5.8 (Universal Imaging), and adjusted and cropped using ImageJ (http://rsbweb.nih.gov/ij/). Exposure times were adjusted and defined for each SNAP transcriptional or translational fusion analyzed. For

quantification of the SNAP-TMR Star signal resulting from transcriptional fusions, 6x6 pixel regions were defined in the desired cell and the average pixel intensity was calculated, and corrected by subtracting the average pixel intensity of the background. Small fluctuations of fluorescence among different fields were corrected by normalizing to the average pixel intensity obtained for the intrinsic autofluorescence of *C. difficile* cells (George *et al.*, 1979).

Statistical analysis. Statistical analysis was carried out using GraphPad Prism (Version 6.0; GraphPad Software Inc.). The non-parametric Kolmogorov-Smirnov test (KS-test) was applied to compare distributions obtained from quantifications of the SNAP-TMR signal. The P-value is indicated for all comparisons whose differences were found to be statistically significant. Although the results presented are from a single experiment, all experiments involving quantification of a fluorescence signal were performed independently three times and only results that were considered statistically significant by a KS-test in all three experiments were considered to be statistically relevant.

Electron microscopy. For transmission electron microscopy electron (TEM) analysis, cells of the wild type 630Δ*erm* strains and of the various *sig* mutants were collected at various times following inoculation onto Columbia Horse Blood Agar plates (BioMérieux). The high fraction of sporulating cells under these growth conditions facilitates the TEM analysis (Joshi *et al.*, 2012). Samples were processed for TEM as described previously (Balomenou *et al.*, 2013).

RESULTS

Sporulation in sporulation medium (SM)

Earlier studies using TEM have suggested that the main stages of sporulation are conserved amongst *Bacillus* and Clostridial species (Piggot and Coote, 1976). Here, we examined sporulation of *C. difficile* using phase contrast and fluorescence microscopy with the goal of establishing a platform for both the phenotypic analysis of mutants blocked in the process and for the analysis of developmental gene expression in relation to the course of morphogenesis. This approach requires the individual scoring of a relative large number of cells. However, under culturing conditions widely used for *C. difficile* sporulation, as in BHI medium, supplemented or not with cysteine and yeast extract (BHIS), the process is highly heterogeneous, or asynchronous (Burns *et al.*, 2010; Cartman *et al.*, 2010; Haraldsen and Sonenshein, 2003; Burns *et al.*, 2010; Burns *et al.*, 2011; Burns and Minton, 2011; Cartman *et al.*, 2010; Haraldsen and Sonenshein, 2003) High titers of spores have been reported following 48h incubation of liquid cultures in the Sporulation Medium (or SM) described by Wilson and co-authors (Wilson *et al.*, 1982) but how the spore titer developed over time was not reported. More recently, SM was used, with some modifications to the original formulation, for high yield spore production on agar plates (Putnam *et al.*, 2013). We determined the spore titer during growth of the wild type strain 630Δ*erm* in liquid SM cultures. As shown in Figure 3.1A, no heat resistant spores could be detected at the time of inoculation, or during the first 10 hours of growth. Heat resistant spores, $3.7x10^2$ spores/ml, were first detected at hour 12, a titer that increased to $2.4x10^5$ at hour 24 (about 14 hours after entry into the stationary phase of growth), the later number corresponding to a percentage of sporulation of 0.3% (Figure 3.1A). From hour 24 onwards, the spore titer increased slowly, to reach $4.7x10^6$

spores/ml 72 hours following inoculation, corresponding to 43.8% sporulation (Figure 3.1A). Importantly, the percentage of sporulation in SM medium was higher than in BHI or BHIS for all the time points tested (Figure 3.1B). In particular, the titer of spores in SM was two orders of magnitude higher than in BHIS, when measured 24 hours following inoculation (Figure 3.1B). For our studies of spore morphogenesis and cell type-specific gene expression, SM was adopted.

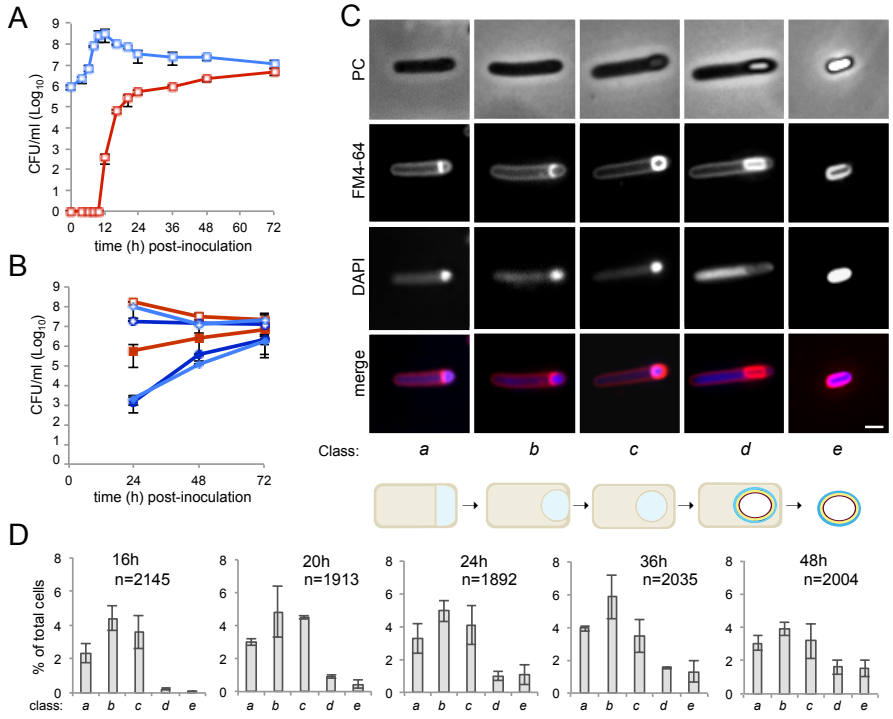


Figure 3.1. Sporulation in *C. difficile* 630Δ*erm*. (A) The spore (red symbols) and total cell titer (blue symbols) was measured for a culture of strain 630Δ*erm* at the indicated times post-inoculation in SM. Asterisk means that no heat resistant CFUs were detected for an undiluted 100 μl culture sample (CFU/ml: ≤10^1). **(B)** Sporulation of strain 630Δ*erm* in SM (red), BHI (dark blue) and BHIS (light blue) liquid media. The heat resistant spore titer (filled symbols) and the total cell titer (open symbols) was measured for culture samples collected at the indicated time points. **(C)** Samples of an SM broth culture of strain 630Δ*erm* were collected 24h after inoculation, stained with DAPI and FM4-64 and examined by phase contrast (PC) and fluorescence microscopy. The panel illustrates the stages in the sporulation pathway, according to the classes defined in the text and represented schematically at the bottom of the panel. Scale bar, 1μm. **(D)** Quantification of the percentage of cells in the morphological classes represented in (B) (as defined in the text), relative to the total viable cell population, for strain 630Δ*erm* at the indicated times following inoculation in SM broth. The total number of cells scored (n) is indicated in each panel. In (A), (B) and (D) the data represent the average ± standard deviation of three independent experiments.

Table 3.2 – Total and heat resistant (H) cell counts (CFU/ml) for the wild type strain (630Δ*erm*) and congenic *sigF*, *sigE*, *sigG* and *sigK* mutants.

t	630Δ*erm*		*sigF*		*sigE*		*sigG*		*sigK*	
(h)	Total	H	Total	H	Total	H	Total	H	Total	H
24	$7.3x10^7$ ± $1.6x10^7$	$3.1x10^5$ ± $5.6x10^4$	$2.3x10^8$ ± $6.4x10^7$	0	$9.6x10^7$ ± $7.2x10^7$	0	$9.0x10^7$ ± $4.5x10^7$	0	$4.4x10^7$ ± $2.7x10^7$	0
48	$4.2x10^7$ ± $8.5x10^6$	$2.4x10^6$ ± $3.1x10^5$	$7.9x10^7$ ± $6.3x10^7$	0	$1.3x10^7$ ± $7.8x10^7$	0	$9.6x10^6$ ± $4.5x10^6$	0	$1.2x10^7$ ± $2.9x10^6$	$2.1x10^2$ ± $3.6x10^1$
72	$2.2.x10^7$ ± $2.9x10^6$	$1.5x10^7$ ± $6.8x10^6$	$4.3x10^7$ ± $2.1x10^7$	0	$7.1x10^6$ ± $2.8x10^6$	0	$3.0x10^6$ ± $3.9x10^6$	0	$6.1x10^6$ ± $4.3x10^7$	$1.7x10^3$ ± $6.0x10^2$

Stages of sporulation

We then wanted to monitor progress through the morphological stages of sporulation by phase contrast and fluorescence microscopy. In a first experiment, a sample from cultures of the wild type strain 630Δ*erm* was collected 24 hours after inoculation into SM, for microscopic examination following staining with the lipophilic membrane dye FM4-64 and with the DNA marker DAPI. Cells representative of several distinctive morphological classes are shown on Figure 3.1C (top). Cells with straight asymmetrically positioned septa (class *a*) and cells with curved spore membranes (*i.e.*, at intermediate stages in the engulfment sequence; class *b*), both showing intense staining of the forespore DNA, were readily seen (Figure 3.1C). Another class comprised cells showing strong uniform FM4-64 staining around the entire contour of the forespore (Figure 3.1C). The staining pattern suggests that the forespore is entirely surrounded by a double membrane, and therefore that the engulfment sequence was finalized. Those cells in which the forespore shows a continuous, strong FM4-64 signal, but has not yet developed partial or full refractility are considered to have just completed the engulfment process, and define class *c*. A strong, condensed DAPI signal in the forespore was also seen for this class (Figure 3.1C). Intense, uniform staining of the forespore by FM4-64 was maintained in cells carrying phase grey (partially refractile) or phase bright spores, defining

class *d* (Figure 3.1C). DAPI staining of the forespore DNA was variable for both cells with phase grey or phase bright spores in this class (data not shown). Free spores, at least some of which could be stained with DAPI, define a last morphological class (class *e*) (Figure 3.1C). The stages of sporulation discerned conform well to the sequence established for *B. subtilis* (Figure 3.1C, bottom) (Henriques and Moran, 2007; Hilbert and Piggot, 2004; Piggot and Coote, 1976). Importantly, staining of both the developing spores, following engulfment completion, and free spores by FM4-64 was observed (Figure 3.1C) (see also the following section).

Under our culturing conditions, cells belonging to each of the five morphological classes considered (*a* to *e*) were seen at all the time points examined (Figure 3.1D). This suggests that sporulation is heterogeneous, or asynchronous, in agreement with other results (Putnam *et al.*, 2013) with cells entering the sporulation pathway throughout the duration of the experiment. Surprisingly, the representation of cells at intermediate stages in development (classes *a* to *c*) decreased from hour 36 to hour 48, without a corresponding rise in later morphological classes (class *d*, phase grey/bright spores and class *e*, free spores) (Figure 3.1D). However, Live/Dead staining evidenced cell lysis, including of cells at intermediate stages of sporulation (classes *a* to *c*), from hour 36 of growth onwards (data not shown). As assessed by Live/Dead staining and fluorescence microscopy, lysis of sporulating cells was only marginal at hour 24 of growth (data not shown). Therefore, in subsequent experiments, sporulating cells were scored 24 hours following inoculation. At hour 24, the total number of sporulating cells (*i.e.*, the sum of classes *a* to *e* in Figure 3.1D) represents about 15% of the total cell population.

Spore staining by FM4-64

A characteristic of sporulation in *C. difficile* 630Δ*erm* that distinguishes it from the *B. subtilis* model is the strong staining of the

developing spore with the FM4-64 dye following engulfment completion (Figure 3.1C). Staining of the engulfed forespore with FM4-64 was unexpected, because the lipophilic dye does not label engulfed forespores of *B. subtilis* (Pogliano *et al.*, 1999).

Because *C. difficile* cells emit strong green fluorescence under long wavelength UV light (George *et al.*, 1979) it seemed possible that spore staining with FM4-64 could be at least in part an artifact caused by emission in the red channel. To investigate this, we mixed sporulating cells of *B. subtilis* (which are not auto-fluorescent) and *C. difficile* and labeled the mixed suspension with FM4-64, prior to fluorescence microscopy examination in both the green and red channels. Under our experimental conditions, sporulating cells of *C. difficile*, but not the developing spore, emit strong green fluorescence, but this signal is not detected in the red channel (Figure 3.2A, top). Conversely, the FM4-64 signal of membranes and spores does not contribute to the green fluorescence emission (Figure 3.2A, top). Additionally, engulfed spores of *B. subtilis*, observed in the same microscope field, do not stain with FM4-64 (Figure 3.2A, top) (Pogliano *et al.*, 1999). These observations confirm that the developing spore of *C. difficile* stains strongly with FM4-64, even following engulfment completion. One possibility is that the membrane of the engulfed spore is in contact with the mother cell membrane. Fluorescence microscopy images of sporulating cells of *Metabacterium polyspora* and of *Epulopiscium* spp., both of which form multiple spores inside the same mother cell, suggest that at least some of the forespores are in contact with the cell membrane, and that they also contact each other (Angert and Losick, 1998; Flint *et al.*, 2005). In *M. polyspora*, the engulfed forespores stain with FM4-64 (Angert and Losick, 1998) However, affinity of the FM dyes to the spore coats has also been reported (Angert and Losick, 1998) and thus, it is unclear whether forespore staining is due to contact of the spore membrane or another spore structure that is already present at this stage (just after engulfment completion) with the mother cell

membrane.

Strong staining of the spore with FM4-64 is maintained at later stages of sporulation, when the spore becomes phase bright. Moreover, and unlike *B. subtilis* (Pogliano *et al.*, 1999), free spores were also stained strongly (Figure 3.2A). To confirm staining of mature spores with FM4-64, we mixed *B. subtilis* and *C. difficile* spores, following which the suspension was stained with FM4-64. The *B. subtilis* spores expressed a fusion of a spore surface protein (Tgl) to CFP, to allow their identification by fluorescence microscopy, in the same microscope field. The *C. difficile* spores which, contrary to cells are not auto-fluorescent (see above), exhibited intense red fluorescence but no blue fluorescence, while those of *B. subtilis* remained unstained by the FM4-64 dye (Figure 3.2A, bottom). Thus, in *C. difficile,* both engulfed forespores as well as mature, free spores are selectively stained by FM4-64.

Because of the reported affinity of FM dyes to the spore coats (Angert and Losick, 1998) we consider the possibility that staining reflected the structure and composition of the spore surface layers. For example, recent work as reported on the presence of an exosporium surrounding the spore coats of *C. difficile* (Permpoonpattana *et al.*, 2011; Lawley *et al.*, 2009) (see also below). However, if an exosporium contributes to staining of the *C. difficile* spores by the FM4-64 dye, then its properties must differ considerably from those of the exosporial layers of an organism like *B. cereus*, whose spore do not stain with the dye (Figure 3.2B). A decoating regime applied to *C. difficile* spores, during which the spore remains phase bright, strongly reduces spore staining by FM4-64 (Figure 3.2C). Although this regime is expected to remove both the exosporium and coat layers, the experiment suggests that staining of *C. difficile* spores with FM4-64 is due, at least in part, to the structure and composition of the spore surface layers.

We suspect that staining of the engulfed forespores by the FM4-64 dye may also be related to the presence of a surface structure, already in place (at least in part) following engulfment completion. In *B. subtilis*, *B.*

cereus and *B. anthracis*, and although not necessarily detected by transmission electron microscopy (TEM) prior to the complete engulfment of the forespore with the concomitant activation of σ^K, assembly of the spore coats begins early in the mother cell line of gene expression, under the control of σ^E (Henriques and Moran, 2007; McKenney *et al.*, 2012).

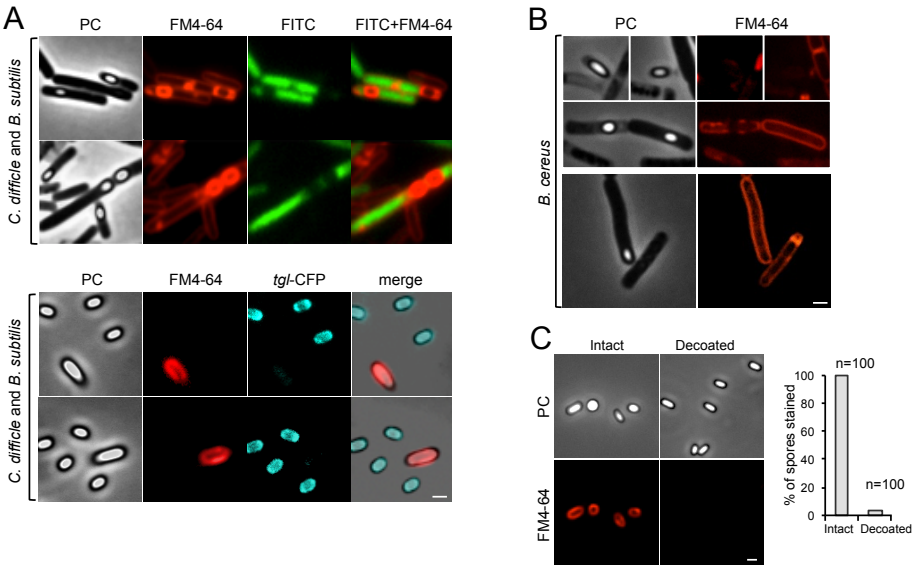


Figure 3.2. Staining of *C. difficile* spores by the membrane dye FM4-64 (A) Differential staining of *C. difficile* and *B. subtilis* spores by FM4-64. Sporulating cells (upper panel) and free spores (lower panel) of *C. difficile* 630Δ*erm* strain were collected and mixed with sporulating cells or free spores of *B. subtilis* MB24, respectively. Sporulating cells and free spores of *C. difficile* 630Δ*erm* were collected after 2 and 5 days of growth in BHI agar. *B. subtilis* MB24 cells were collected 6 and 24 hours after the onset of sporulation in DSM. After mixing, cells were stained with FM4-64, and analyzed by phase contrast (PC) and fluorescence microscopy. FITC column shows the typical autofluorescence of *C. difficile* cells. The *B. subtilis* strain used produces a CFP fusion to a coat protein (Tgl-CFP). **(B)** Staining of *B. cereus* strain ATCC4342 sporulating cells and free spores with FM4-64. **(C)** Decoating of *C. difficile* spores drastically reduces staining with FM4-64. *C. difficile* purified spores were submitted or not to a decoating treatment prior to staining with FM4-64, and analysed by fluorescence microscopy. The percentage of total spores labeled with FM4-64 was scored (total number of spores scored (n) is indicated). Scale bar, 1 μm.

Spore ultrastructure

Sporulating cells and spores of *C. difficile* were also examined using TEM. A feature noteworthy is that in some of the sporulating cells, a membrane-looking structure surrounding the forespore was noticed (Figure

3.3A and B). This membrane-looking structure appears to contact the cell membrane (Figure 3.3B). We were unable to determine when this structure was formed during sporulation, as it was only detected in cells, as those represented in Figure 3.3, at late stages in development, as judged by the presence of spores with discernible cortex and coat layers, and by the almost complete dissolution of the mother cell cytoplasm. We do not know whether the presence of this structure is related to the FM4-64 staining of engulfed forespores.

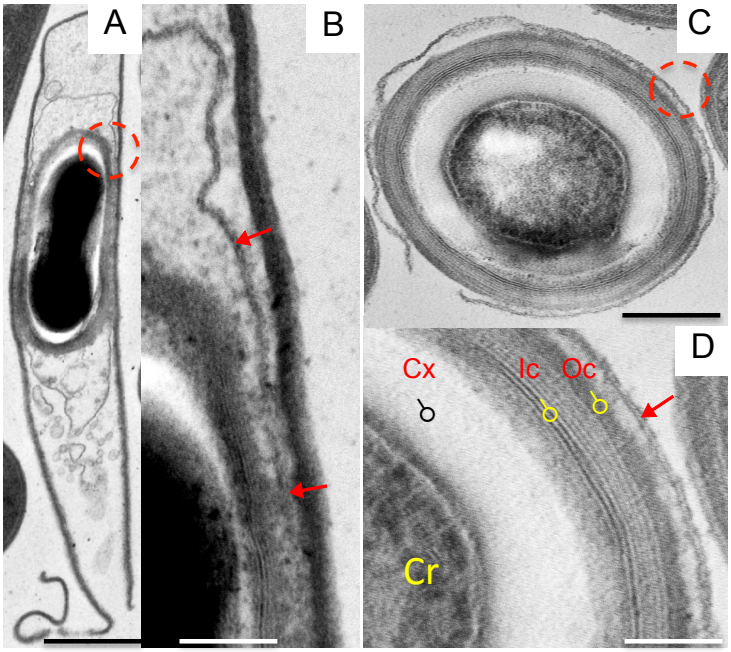


Figure 3.3. Ultrastructure of 630Δ*erm* spores. **(A)** and **(B)** show a spore within a lysed mother cell; **(C)** and **(D)** show a free spore; (B) and (D) are a magnification of the region encircled in (A) and (C). Arrows point to a layer apposed to the external surface of the spore. The main structural features of *C.difficile* spores are indicated in C/D: Cr, core; Cx, cortex; IC, inner coat; OC, outer coat. Scale bar, 1 μm (A,C) or 0.5 μm (B,D).

In most of the free spores the cortex was surrounded by a lamellar inner coat layer formed by 5-6 lamellae, covered by an electrodense outer layer (Figure 3.3C and D). In some of the free spores, an extra layer with a loose shape surrounded the spore, at a distance from the outer coat surface

(Figure 3.3C). This layer appears as an exosporium, formed by a thin basal layer, with more electrondense material closely apposed to its external surface. The basal layer appears similar to the membrane-like structure seen in sporulating cells (above). It is tempting to speculate that the membrane-like structure seen in sporulating cells represents an intermediate in the assembly of the exosporium. The exosporium-like layer was only seen by TEM in about 10% of the free spores, suggesting that this structure is labile and does not survive sample collection and processing for electron microscopy. In two recent publications, the exosporium surrounding *C. difficile* spores has a thick electrondense appearance (Permpoonpattana *et al.*, 2011; Lawley *et al.*, 2009) and thus differs from the structure herein described. It is possible that morphogenesis of the exosporial layer differs with culturing conditions.

Disruption of the genes for the sporulation sigma factors

The genes for the four cell type-specific RNA polymerase sigma factors known to control gene expression during spore differentiation in *B. subtilis* are conserved in sporeformers (Galperin *et al.*, 2012; Traag *et al.*, 2012; Miller *et al.*, 2012). Moreover, their operon structure and genomic context is also maintained (Figure 3.4A). Having defined the basic stages of spore formation and spore morphology in *C. difficile*, as well as the appropriate conditions to study the process, we went to investigate whether the function of the σ^F, σ^E, σ^G and σ^K factors in sporulation is conserved in this organism. For this purpose, each of the corresponding sig genes was disrupted using the ClosTron system (Heap *et al.*, 2007). In this system, type II introns are targeted and inserted into a gene of interest, disrupting it. As shown in Figure 3.4A, re-targeting of the intron resulted in insertion after codon 153 of the *sigF* gene, codon 151 of *sigE*, codon 182 of *sigG*, and codon 34 of the 5´-end of the split *sigK* gene, interrupted by the *skin*^{Cd} element.

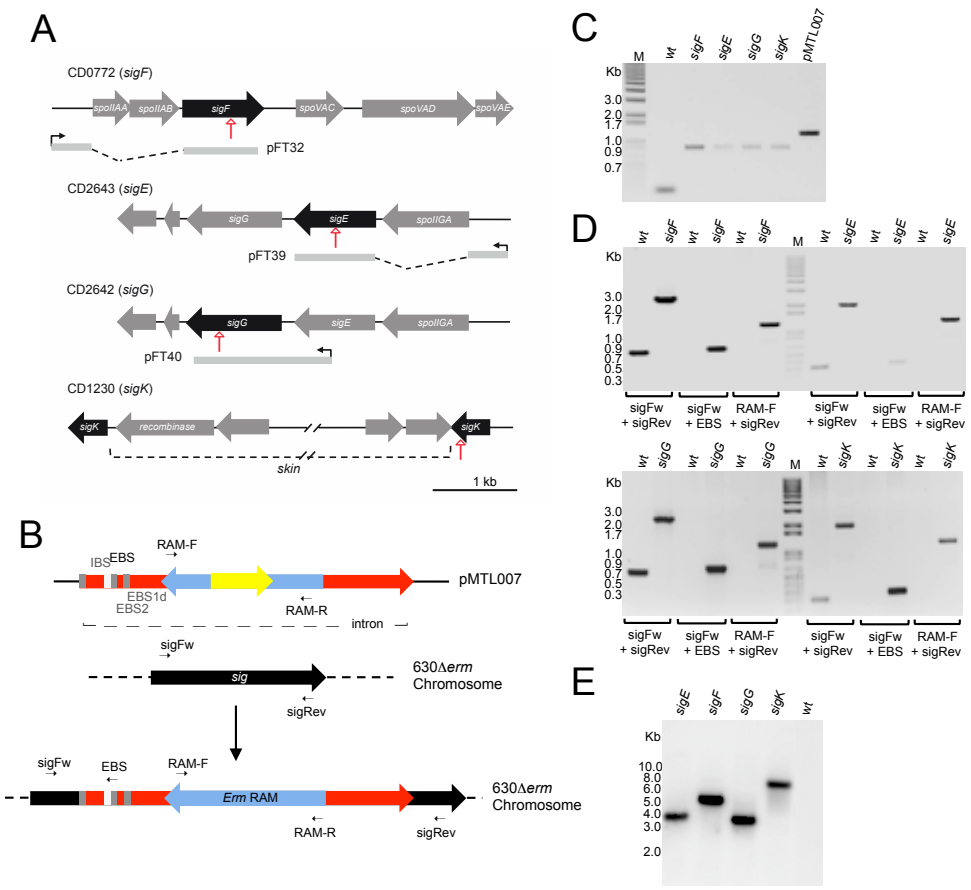


Figure 3.4. Inactivation of the *sigF*, *sigE*, *sigG* and *sigK* genes in *C. difficile* using the ClosTron system. (A) Genetic organisation of the *C. difficile* chromosome in the vicinity of *sigF*, *sigE*, *sigG* and *sigK* (interrupted by the *skin* element). The red arrow indicates the point of insertion of the re-targeted type II introns used for gene disruption. The extent of the DNA fragment present in the indicated replicative plasmids used for *in trans* complementation of the insertional mutations is shown below each of the genetic maps, except for the *sigK* gene (see Figure 3.6C). **(B)** Schematic representation of gene inactivation by a type II Intron with an associated RAM. The group II intron (bracket), originally in pMTL007 (top), carries a RAM element (yellow) interrupting an *ermB* determinant (blue). The intron was retargeted to the sig gene of interest (black; middle) by altering the IBS, EBS1 and EBS2 sequences (grey and white stripes; top). Splicing out of the td group I intron from the *ermB* gene in the RAM restores a functional marker. Primers used to confirm the integration and orientation of the type II intron are also indicated. **(C, D)** Chromosomal DNA of Em^R *C. difficile* conjugants was screened by PCR using primer pairs RAM-F/R to confirm splicing out of the group I intron (C), or with primer pairs sigFw/Rev, sigFw/EBS and RAM-F/sigRev to confirm disruption of the *sig* genes (D). pMTL007 (C) and chromosomal DNA from the 630Δ*erm* (wild type, wt) strain (D) were used as controls. **(E)** Southern blot of HindIII-digested genomic DNA from *C. difficile* 630Δ*erm* (wild type, wt), *sigF*, *sigE*, *sigG* and *sigK* mutant strains. The probe used corresponds to part of the intron sequence. The position of DNA size markers is indicate on the left side of the panel.

Correct insertion of the intron was verified, in all cases, by PCR, and Southern blot analysis showed the presence of a single intron insertion in the genome of the different mutants (Figure 3.4B through E). The *sig* mutants, along with the parental 630Δ*erm* strain, were induced to sporulate in SM, and the titer of heat resistant spores assessed after 24, 48, and 72 hours of growth. For the wild type strain 630Δ*erm*, the titer of spores was of $3x10^5$ spores/ml at hour 24, $2x10^6$ spores/ml at hour 48, and of $1x10^7$ spores/ml at hour 72 (Table 3.2). In contrast, no heat resistant spores were found, at any time point tested, for the *sigF* (AHCD533), *sigE* (AHCD532), or *sigG* (AHCD534) mutants. However, a titer of 10^3 heat resistant spores/ml of culture was found for the *sigK* mutant AHCD535 at hour 72 (Table 3.2). For complementation studies, we generated multicopy alleles of the *sig* genes, based on replicative plasmid pMTL84121 (Heap *et al.*, 2009), expressed from their native promoters (the extent of the promoter fragments is shown in Figure 3.4A). Note that for complementation of the *sigK* mutation the two halves of the gene, together with a short *skin*^{Cd} element composed only of the putative recombinase gene (*spoIVCA*, or *CD12310*) was used (Figure 3.6C; see also below for a more detailed description on the complementation of the *sigK* mutation). When measured at hour 72 of growth in SM, the heat resistant spore titer was of $1.7x10^6$ spores/ml for the wild type strain 630Δ*erm* carrying the empty vector pMTL84121. Derivatives of pMTL84121 carrying the *sigF*, *sigE*, *sigG* or *sigK* genes (the later plasmid, pFT38, with the short *skin*^{Cd} allele) restored spore formation to the *sig* mutants, as assessed by microscopy (Figure 3.5B). The same plasmids largely restored heat resistant spore formation to the *sig* mutants ($1.6x10^4$, $8.3x10^5$, $3.9x10^5$, $4.8x10^5$ spores/ml for the *sigF*, *sigE*, *sigG*, and *sigK* mutants, respectively, also measured at hour 72 of growth in SM).

Morphological characterization of the *sigF*, *sigE*, and *sigG* mutants

To establish the morphological phenotype of the various mutants we used phase contrast and fluorescence microscopy of samples collected from SM cultures at hour 24, labeled with DAPI and FM4-64 (Figure 3.5). These studies revealed that the *sigF* and *sigE* mutants were blocked at the asymmetric division stage (Figure 3.5A and C). As previously found for *B. subtilis* (Piggot and Coote, 1976; Eldar *et al.*, 2009; Eichenberger *et al.*, 2001) both mutants formed abortive disporic forms, and occasionally multiple closely located polar septa (Figure 3.5A). In addition, for the *sigF* mutant, small round cells were found, probably resulting from detachment of the forespore (Figure 3.5A). In both mutants, the DNA stained strongly in the forespore(s) and gave a diffuse signal throughout the mother cell (Figure 3.5A). TEM analysis confirmed the block at the asymmetric division stage for the two mutants (Figure 3.5D).

Cells of the *sigG* mutant completed the engulfment sequence, but did not proceed further in morphogenesis (Figure 3.5A and C). As for class *c* in the wild type (Figure 3.1B, and text above), the forespores in the *sigG* mutant stained strongly with FM4-64 (Figure 3.5A). TEM of sporulating cells of the *sigG* mutant confirmed engulfment completion, but also revealed deposition of electrodense material around the forespore protoplast (Figure 3.5D). This deposit could represent coat material. By comparison, no accumulation of electrodense coat-like material is seen by TEM around the engulfed forespore of a *B. subtilis sigG* mutant (Karmazyn-Campelli *et al.*, 1989). In this organism, coat assembly as discernible by TEM, is a late event that requires activation of σ^K in the mother cell (Henriques and Moran, 2007; McKenney *et al.*, 2012; Piggot and Coote, 1976). Importantly, activation of σ^K is triggered by σ^G, and coincides with engulfment completion (Lu *et al.*, 1990; Cutting *et al.*, 1990). Therefore, the possible accumulation of coat material in the *sigG* mutant could imply that in *C. difficile*, σ^K is active independently of σ^G. We therefore wanted to test whether coat material was deposited around the

forespore in the *C. difficile sigG* mutant. In *B. subtilis*, studies of protein localization have relied mainly on the use of translational fusions to the *gfp* gene, or its variants (e.g., Wang *et al*., 2009). However, as addressed in Chapter 2, an obstacle to the use of *gfp* or its derivatives in the anaerobe *C. difficile*, is that formation of the GFP fluorophore involves an oxidation reaction (Reid and Flynn, 1997). For this reason, we turned to the SNAP-tag reporter and constructed C-terminal fusions of the SNAP-tag to spore coat proteins CotE and CotB (Permpoonpattana *et al.*, 2011; Permpoonpattana *et al.*, 2013) in plasmid pFT58 (Table A2; see also Chapter 2). The fusions were introduced, in a replicative plasmid, in strain 630Δ*erm* and the *sigG* and *sigK* mutants, and samples from SM cultures at hour 24 were labeled with the cell-permeable fluorescent substrate TMR-Star (see Materials and Methods). Using fluorescence microscopy and fluorimaging of SDS-PAGE-resolved whole cell extracts, no accumulation of CotE-SNAP was detected in cells of a *sigK* mutant, suggesting that the *cotE* gene is under the control of σ^K (Figure 3.5E; see also below). CotB-SNAP, however, accumulated in cells of a *sigK* mutant (Figure 3.5E), but not in cells of a *sigE* mutant (data not shown), suggesting that expression of *cotB* is under the control of σ^E. Both CotE-SNAP and CotB-SNAP localized around the forespore in both wild type and in *sigG* cells (Figure 3.5E). SDS-PAGE and fluorimaging suggested instability of CotB-SNAP for which several possible proteolytic fragments were detected, all of which larger that the SNAP domain (Figure 3.5E). That no release of a labeled SNAP domain was detected for either protein implies that the localized fluorescence signal is largely due to the fusion proteins. Thus, both early (CotB) and late (CotE) coat proteins are assembled around the forespore in cells of a *sigG* mutant. In all, the results suggest that σ^K is active independently of σ^G, and thus, that the later regulatory protein is not a strict requirement for deposition of at least some coat in *C. difficile*.

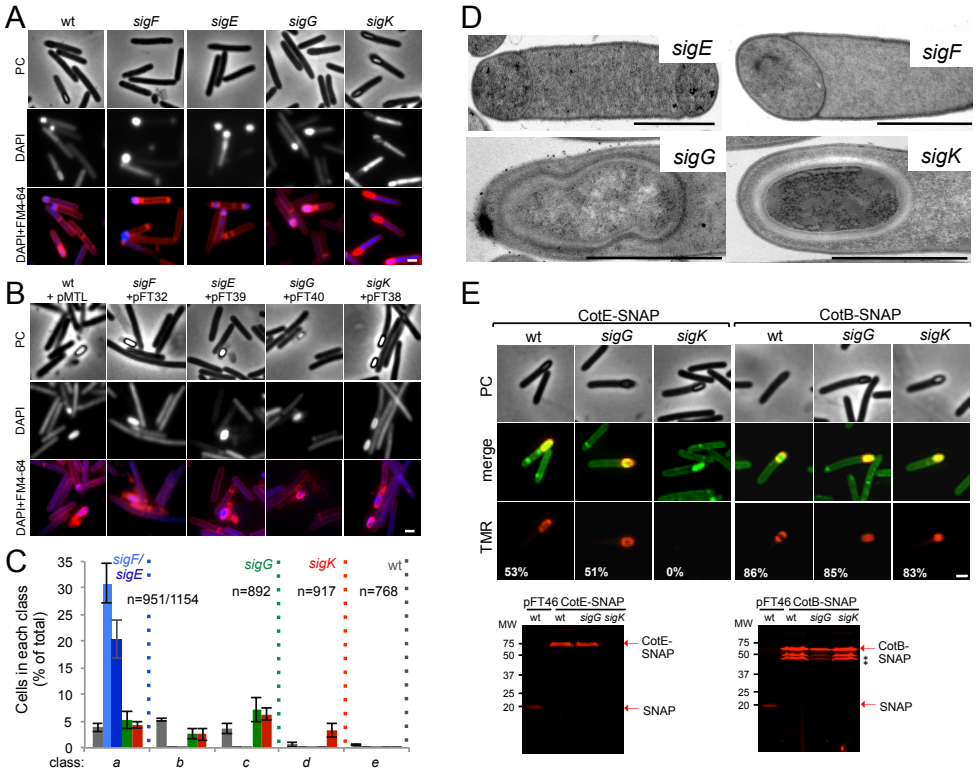


Figure 3.5. The sporulation pathway in *C. difficile* 630Δ*erm*, and the role of *sigF*, *sigE*, *sigG* and *sigK*. Phase contrast (PC) and fluorescence microscopy analysis of spore morphogenesis in the following strains: **(A)** 630Δ*erm* (wild type, wt) and the congenic *sig* mutants; **(B)** the *sig* mutants bearing the indicated plasmids or the wt strain carrying the empty vector pMTL84121. Cells were collected 24 (A) or 48h (B) after inoculation in SM broth, and stained with DAPI and FM4-64, prior to microscopic examination. **(C)** Quantification of the cells in each morphological class for the wt strain and the *sig* mutants. The data represent the average ± SD of three independent experiments. The total number of cells analysed (n) is indicated. **(D)** TEM images of *sigF*, *sigE*, *sigG* and *sigK* mutant cells. The Images are representative of the most common morphological phenotype observed for each mutant. **(E)** Fluorescence microscopy of 630Δ*erm* and *sigG* and *sigK* strains carrying CotE- and CotB-SNAP fusions. Cells were collected and labeled with the SNAP substrate TMR-Star (red channel) and the membrane dye MTG (green channel). The numbers on the bottom panel represent the percentage of cells which have completed the engulfment process that show localization of the protein fusions around the forespore (n=80-100 cells). Scale bar in panels (A, B, D, E), 1μm. Total cell extracts from the same strains were prepared and analyzed, as described in Materials and Methods. Production of the SNAP protein under the control of P_{tet} (pFT46), was used as a control. The position of the SNAP or SNAP fusion proteins is indicated by arrowheads. Asteriks indicate possible degradation products.

Functional analysis of the *sigK* gene

Phase contrast microscopy revealed the presence of some phase bright or partially phase bright spores in SM cultures of the *sigK* mutant,

although free spores were only rarely seen (Figure 3.5A). The ellipsoidal spores were often positioned slightly tilted relative to the longitudinal axis of the mother cell (Figure 3.5A and 3.6A). The appearance of phase bright spores normally correlates with synthesis of the spore cortex PG, and the development of spore heat resistance (Piggot and Coote, 1976; Vasudevan *et al.*, 2007) in line with the finding that the *sigK* mutant formed heat resistant spores (above). TEM revealed the presence of a cortex layer in cells of the *sigK* mutant, supporting the inferences drawn on the basis of the phase contrast microscopy and heat resistance assays (Figure 3.5A to E). The number of phase bright or phase grey spores by phase contrast microscopy, was 3.2% of the total number of cells scored at hour 24 of growth in liquid SM (Figure 3.5C). This is higher than the percentage of sporulation, 0.03%, measured by heat resistance (Table 3.2). Because full heat resistance requires synthesis of most of the cortex structure, this observation suggests that a large number of the spores formed have an incomplete or dysfunctional cortex. However, we cannot discard the possibility that spores of the mutant are deficient in germination. In any event, unlike in *B. subtilis*, where a *sigK* mutant is unable to form the spore cortex (Piggot and Coote, 1976) (Vasudevan *et al.*, 2007), σ^K is not obligatory for the biogenesis of this structure in *C. difficile*. In contrast, the TEM analysis did not reveal deposition of coat material around the cortex in cells of the *sigK* mutant (Figure 3.5D). Although coat assembly most likely starts early, under the control of σ^E (Henriques and Moran, 2007; McKenney *et al.*, 2012; Putnam *et al.*, 2013); above) the TEM data, together with the data on assembly of CotE (Figure 3.5E), suggest that the late stages in the assembly of the coats are under σ^K control. That free spores were only rarely seen for the *sigK* mutant, prompted us to test whether σ^K could have a role in mother cell lysis, using a Live/Dead stain and fluorescence microscopy. In the wild type strain 630Δ*erm*, development of refractility coincided with loss of viability of the mother cell (strong staining with propidium iodide) and strong staining of

the developing spore with the Syto 9 dye (Figure 3.6A) (Magge *et al.*, 2009). In contrast, the mother cell remained viable in the *sigK* mutant (strong staining with Syto 9) (Figure 3.6A and B), and the spores stained only weakly with the Syto 9 dye.

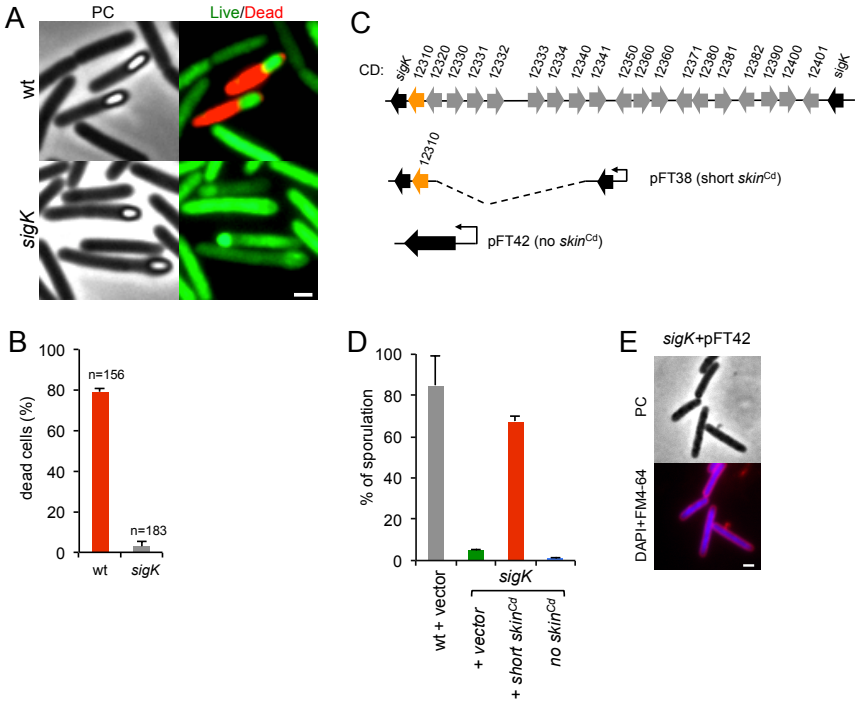


Figure 3.6. Functional analysis of the *sigK* gene. **(A)** Live/dead assay for the wild type (630Δ*erm*) and the *sigK* mutant. Shown are phase contrast and the merge between syto 9- (green) and propidium iodide- (red) stained cells collected at 24h of growth in SM broth. **(B)** Percentage of the sporulating cells of the wild type and *sigK* mutant strains (i.e., with visible spores) showing signs of mother cell lysis (red). "n" represents the total number of cells analyzed. **(C)** The *sigK-skin* region of the 630Δ*erm* chromosome and plasmids used to complement *sigK* mutant strain. Replicative plasmid pFT38 carries *sigK* interrupted by a shorter version of the *skin*^{Cd} element, which includes the gene (*CD12310*) for the recombinase (in orange). Replicative plasmid pFT42 carries an uninterrupted *sigK* gene. **(D)** Percentage of sporulation for the wt, *sigK* and *sigK* bearing either pFT38 or pFT42, 72h following inoculation in SM medium. **(E)** Fluorescence microscopy showing the phenotype of *sigK* bearing pFT42. Cells were collected at 72h of growth in SM broth, stained with DAPI and FM4-64, and viewed by phase contrast (PC) and fluorescence microscopy. Scale bar in (A) and (E), 1μm.

Lastly, our complementation analysis of the *sigK* mutant provided additional functional insight. While wild type levels of sporulation could be restored to a *sigK* mutant by a copy of the *sigK* gene bearing a deletion of all

the genes within the *skin*^{Cd} element but the recombinase gene (Figure 3.6C; see above), an uninterrupted copy of the gene, in plasmid pFT42, did not restore sporulation (Figure 3.6C and D). An earlier study has suggested that the absence of *skin*^{Cd} correlates with a sporulation defect and that a *skin*^{Cd}-allele of *sigK* is dominant over the wild type (Piggot and Coote, 1976). Our results support the view that generation of an intact *sigK* gene through SpoIVCA-mediated excision of the *skin*^{Cd} element is essential for sporulation. Moreover, we found that introduction of the multicopy *skin*-less allele in strain 630Δ*erm* blocked sporulation at an early stage, as no asymmetrically positioned septa could be seen in the transformed strain (Figure 3.6E). The results suggest that the absence of *skin*^{Cd} allows the production of active σ^K in pre-divisional cells, and that active σ^K interferes with the events leading to asymmetric septation in *C. difficile*.

Localizing the expression of the sporulation-specific *sig* genes

Having established the main features of sporulation under our culturing conditions, as well as the phenotypes associated with disruption of the *sig* genes, we next wanted to examine cell type-specific gene expression in relation to the course of morphogenesis. As a first step, we examined the expression of the genes coding for σ^F, σ^E, σ^G, and σ^K using the *SNAP*^{Cd} cassette as a transcriptional reporter. The promoter regions of *sigF*, *sigE, sigG*, and *sigK* genes were cloned in the *SNAP*^{Cd}-containing promoter probe vector pFT47 (Table A1, see also Chapter 2). The upstream boundaries of the promoter fragments fused to *SNAP*^{Cd} coincide with the 5´-end of the fragments used for the successful complementation of the various *sig* mutants (Figure 3.4A and 3.6C; see above). To monitor the production of SNAP during *C. difficile* sporulation, samples of cultures expressing each of the promoter fusions were collected at 24 h of growth in SM medium, and the cells doubly labeled with TMR-Star and the membrane dye MTG, to allow identification of the different stages of sporulation. These were defined based

on Figure 3.1C, with the addition of a class of pre-divisional cells (no signs of asymmetric division). Expression of the various P_{sig}-*SNAP*^{Cd} transcriptional fusions could thus be correlated to the stage in spore morphogenesis.

Expression of both *sigF* and *sigE* was first detected in pre-divisional cells of the wild type strain 630Δ*erm*, but not in cells of a *spo0A* mutant (Figure 3.7), consistent with previous reports (Rosenbusch *et al.*, 2012; Saujet *et al*., 2011; Fimlaid *et al*., 2013) (see also Chapter 4). Both genes continued to be expressed following asymmetric division, in the forespore and the mother cell of both the wild type, and the *sigF* or *sigE* mutants (Figure 3.7). In these experiments, complete labeling of the SNAP protein was achieved, as revealed by fluorimaging and immunobloting of SDS-PAGE resolved proteins in whole cell extracts (Figure 3.8A). Quantification of the fluorescence signal shows that while for *sigF* the average intensity did not differ much between forespores (1.8±0.5), and mother cells (1.8±0.5), it increased in both the forespore and the mother cell relative to pre-divisional cells (average signal, 1.5±0.4) ($p<0.01$). Transcription of *sigE*, in turn, was lower in the forespore (average signal, 0.8±0.3) as compared to pre-divisional cells (1.0±0.3) or the mother cell (1.0±0.3) ($p<0.0001$) (Figure 3.8B). Thus, transcription of *sigE*, seems to occur preferentially in the mother cell. Transcription of both *sigF* and *sigE* persisted in both the forespore and the mother cell until a late stage of sporulation, when the forespore becomes phase bright (Figure 3.7).

In contrast to *sigF* and *sigE*, transcription of *sigG* and *sigK* was confined to the forespore and to the mother cell, respectively (Figure 3.7). Transcription of *sigG* is detected in the forespore just after asymmetric division, consistent with the presence of a σ^F-type promoter in its regulatory region (Saujet et al., 2013) (see also Chapter 4). In agreement with this inference, expression of P_{sigG}-*SNAP*^{Cd} was not detected in cells of a *sigF* mutant (Figure 3.7). Transcription of *sigG* was detected until the development of spore refractility (Figure 3.7). Fluorimaging and immunoblot

analysis of whole cell extracts shows that under the conditions used, all of the SNAP protein detected was labeled (Figure 3.8A). In *B. sub tilis*, σ^F initiates transcription of *sigG* in the forespore (Karmazyn-Campelli *et al.*, 1989; Sun *et al.*, 1991). However, transcription of *sigG* also depends on σ^E, by an unknown mechanism (Partridge and Errington, 1993). In contrast, forespore-specific expression of P_{sigG}-*SNAP*^{Cd} was detected in most cells (82%) of a *sigE* mutant (Figure 3.7).

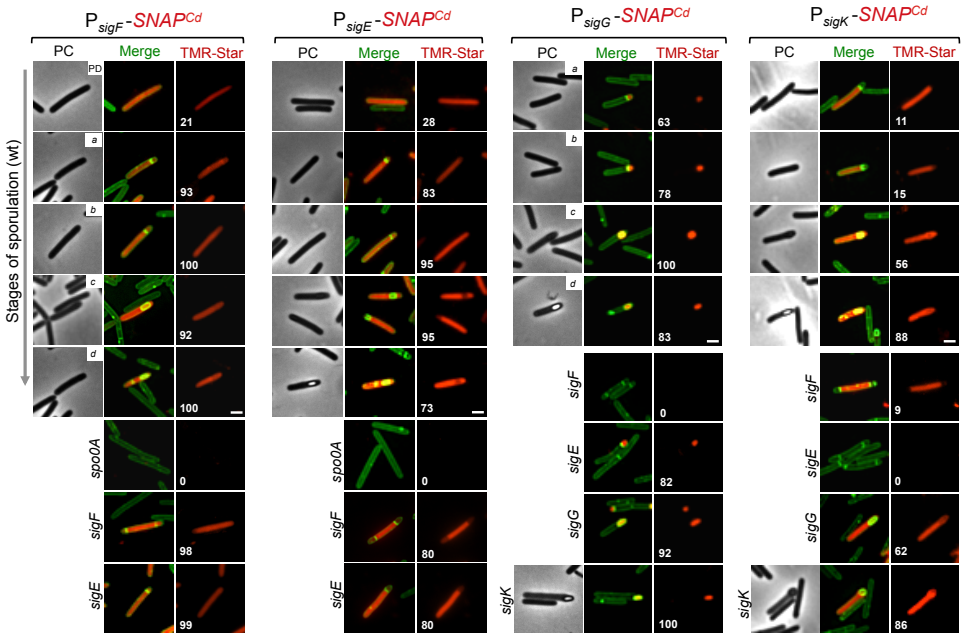


Figure 3.7. Temporal and cell type-specific expression of *sigF*, *sigE*, *sigG* and *sigK* during sporulation. Microscopy analysis of *C. difficile* cells carrying fusions of the *sigF*, *sigE*, *sigG* and *sigK* promotors to *SNAP*^{Cd} in strain 630Δ*erm* and in the indicated mutants. The merged images shows the overlap between the TMR-Star (red) and MTG (green) channels. The panels are representative of the expression patterns observed for different stages of sporulation, ordered from early to late for the wild type strains according to the morphological classes *a-d* defined in Figure 3.1, as indicated. For the *sig* strains, the morphological stage characteristic of each mutant is indicated. An extra class that accounts for pre-divisional cells (PD) was introduced for the analysis of both *sigF* and *sigE* transcription. The numbers refer to the percentage of cells at the represented stage showing SNAP fluorescence. Data shown are from one representative experiment, of three performed independently. The number of cells analysed for each class, n, is as follows: PD, n=100-150; class *a*, n=30-50; class *b*, n=50-60; class *c*, n=30-40; class *d*, n=15-25; for *sigF/E* mutants, n=80-120; for *sigG* and *sigK* mutants, n=40-50. Scale bar: 1 *μ*m.

In *B. subtilis*, the main period of *sigG* transcription takes place following engulfment completion, and relies on a positive auto-regulatory loop (Serrano *et al.*, 2011). We detected transcription of *sigG* both prior and following engulfment completion in a *sigG* mutant (Figure 3.7). However, the quantitative analysis of the SNAP-TMR signal shows an increase in the average fluorescence intensity following engulfment completion (classes *c* + *d*, 2.6±1.0 as opposed to 2.2±0.9 for classes *a* + *b*) ($p<0.01$) (Figure 3.8B). Moreover, the average fluorescence signal for engulfed forespores of a *sigG* mutant suffered a higher reduction compared to the wild type (classes *c + d,* 1.9±0.7 for the mutant as compared to 2.6±1.0 for the wild type; $p<0.01$), than did the signal for pre-engulfment forespores of the mutant (classes *a + b,* 1.7±0.5 as opposed to 2.2±0.9; $p<0.05$) (Figure 3.8B). While evidencing that σ^G contributes to transcription of its own gene both prior to and following engulfment completion, these results suggest that the auto-regulatory effect is stronger at the later stage.

In *C. difficile*, transcription of *sigK* was confined to the mother cell and detected soon after asymmetric division (Figure 3.7). Moreover, disruption of *sigE* resulted in undetected expression of P_{sigK}-*SNAP*^{Cd} (Figure 3.7). Together, the results suggest that the initial transcription of *sigK* is activated by σ^E in the mother cell, consistent with the presence of a possible σ^E-recognized promoter in the *sigK* regulatory region (see also Chapter 4). Interestingly, transcription of the *sigK* gene was also detected in a small percentage (9%) of the sporulating cells of a *sigF* mutant (Figure 3.7). This was unexpected because in *B. subtilis*, activation of σ^E in the mother cell is dependent on σ^F (Londono-Vallejo and Stragier, 1995; Karow *et al.*, 1995). This observation thus raises the possibility that the activation of σ^E in *C. difficile* is at least partially independent of σ^F (see also the following section). Transcription of *sigK* was also detected following engulfment completion, in cells carrying phase grey and phase bright spores (Figure 3.7). As shown in Figure 3.8A, all of the SNAP produced from the P_{sigK}-*SNAP*^{Cd} fusion was, under

our experimental conditions, labeled. The average intensity of the fluorescence signal from P_{sigK}-*SNAP*^{Cd} in cells prior (classes *a* + *b*, 0.6±0.1) and after engulfment completion (class *c*, 0.7±0.2) was very close. However, expression was significantly increased for those cells that carried phase bright spores (class *d*, 1.4±0.8; p<0.001) (Figure 3.8B). This suggests that the onset of the main period of *sigK* transcription coincides with the final stages in spore morphogenesis. Lastly, under our experimental conditions, we found no evidence for auto-regulation of *sigK* transcription, as expression of P_{sigK}-*SNAP*^{Cd} was not curtailed by mutation of *sigK* at any morphological stage analyzed (Figure 3.7 and data not shown).

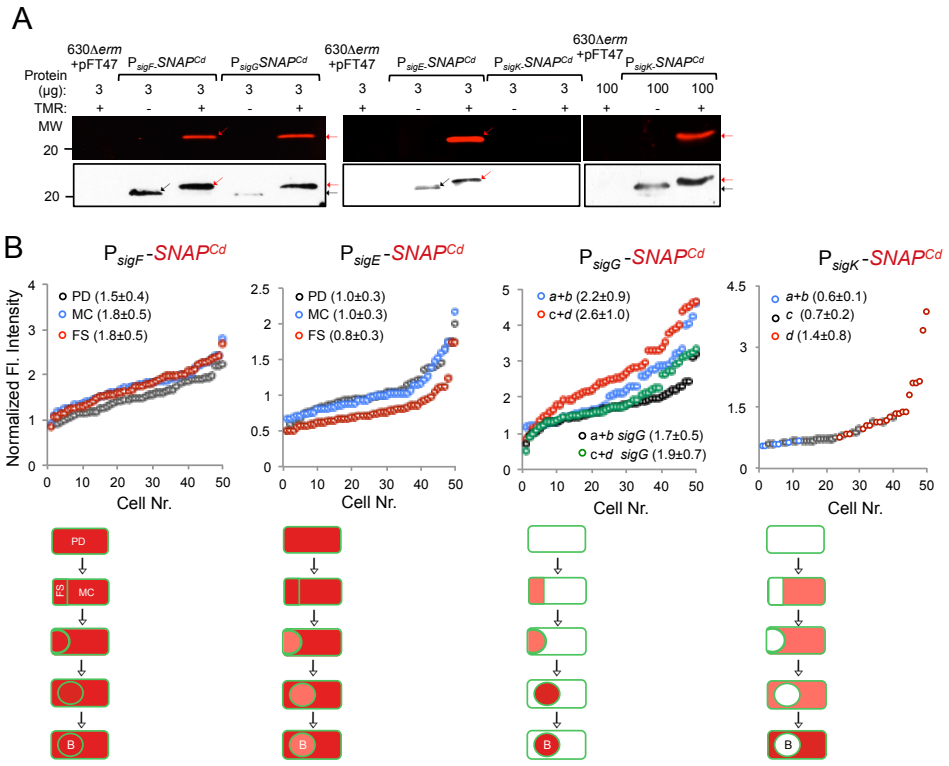


Figure 3.8. Quantitative analysis of *sigF*, *sigE*, *sigG* and *sigK* expression during sporulation. **(A)** Whole cell extracts were prepared from derivatives of strain 630Δ*erm* bearing the indicated plasmids or fusions, immediately after labeling with TMR-Star, indicated by the "+" sign (the "-" sign indicates unlabeled samples). The indicated amount of total protein was resolved by SDS-PAGE, and the gels scanned on a fluorimager (top) or subject to immunoblotting with anti-SNAP antibodies (bottom). Black and red arrows point to unlabeled or TMR-Star-labeled, respectively, SNAP. The position of molecular weight markers (in kDa) is indicated. **(B)** Quantitative analysis of the fluorescence (Fl.)

intensity in different cell types of the reporter strains for *sigF*, *sigE*, *sigG* and *sigK* transcription, as indicated. The numbers in the legend represent the average ± SD of fluorescence intensity for the cell class considered. The data shown are from one experiment, representative of three independent experiments. Schematic representation of the deduced spatial and temporal pattern of transcription (with darker red denoting increased transcription) is shown for each transcriptional fusion. PD, pre-divisional cell; MC, mother cell; FS, forespore; B, phase bright spore; *a* to *d*: sporulation classes as defined in Figure 3.1.

Localizing the activity of σ^F and σ^E

To investigate the genetic dependencies for sigma factor activity during sporulation in *C. difficile*, we used transcriptional *SNAP*^{Cd} fusions to promoters under the control of each cell type-specific sigma factor. These promoters were selected on the basis of qRT-PCR experiments and the presence on their regulatory regions, of sequences conforming well to the consensus for promoter recognition by the sporulation sigma factors of *B. subtilis* (Wang *et al.*, 2006). The *gpr* gene of *B. subtilis* codes for a spore-specific protease required for degradation of the DNA-protecting small acid-soluble spore proteins (SASP) during spore germination (Nicholson, 2000) (see also below). Even though this gene is under the dual control of σ^F and σ^G in *B. subtilis*, the *C. difficille* orthologue of *gpr* (*CD2470*) was chosen as a reporter for σ^F activity. First, qRT-PCR showed that transcription of the *C. difficille* orthologue (*CD2470*) was severely reduced in a *sigF* mutant (Figure 3.9A). Secondly, expression of a P_{gpr}-*SNAP*^{Cd} fusion, monitored by fluorescence microscopy, was confined to the forespore and detected soon after asymmetric division in 66% of the cells that were at this stage of sporulation (Figure 3.9B). Lastly, expression was eliminated by disruption of the *sigF* gene but detected in 99% of the cells of the *sigG* mutant (compared for the wild type at the same stage, *i.e.*, 95%) (Figure 3.9B). This suggests that σ^G does not contribute significantly to *gpr* expression. Forespore-specific expression of P_{gpr}-*SNAP*^{Cd} was also detected following engulfment completion (Figure 3.9B). Therefore, in spite of expression of the *sigF* gene in both the forespore and the mother cell, σ^F is active exclusively in the forespore. In these experiments, all of the SNAP protein produced from P_{gpr}-*SNAP*^{Cd} was

labeled with the TMR-Star substrate (Figure 3.10A). A quantitative analysis of the fluorescence signal from P_{gpr}-*SNAP*^{Cd} showed no significant difference between cells before (average signal for classes *a* + *b*, 2.0±0.5) or after engulfment completion (classes *c* + *d*, 1.9±0.7) (Figure 3.10B). This suggests that σ^F is active in the forespore throughout development. To monitor the activity of σ^E, we examined expression of the first gene, *spoIIIAA*, of the *spoIIIA* operon. This operon is under the control of σ^E in *B. subtilis* (Steil *et al.*, 2005; Eichenberger *et al.*, 2004; Eichenberger *et al.*, 2003) and sequences that conform well to the consensus for promoter recognition by *B. subtilis* σ^E are found just upstream of the *C. difficile spoIIIAA* gene (or *CD1192*) (see Chapter 4). The qRT-PCR experiments showed that expression of *spoIIIAA* was much more severely affected by a mutation in *sigE* than by disruption of *sigF* (Figure 3.9A). While consistent with a direct control of *spoIIIAA* by σ^E, this observation adds to the evidence suggesting that unlike in *B. subtilis* (Hilbert and Piggot, 2004; Londono-Vallejo and Stragier, 1995; Karow *et al.*, 1995), the activity of σ^E is at least partially independent on the prior activation of σ^F (as also hinted by the observation that transcription of the *sigK* gene, abolished by mutation of *sigE*, was still detected in a fraction of cells of a *sigF* mutant; above). If so, then expression of a $P_{spoIIIAA}$-*SNAP*^{Cd} fusion should be confined to the mother cell, dependent on *sigE*, but partially independent on *sigF.* $P_{spoIIIAA}$-*SNAP*^{Cd}-driven SNAP production was indeed confined to the mother cell, detected just after asymmetric division in 89% of the cells scored at this stage of sporulation, eliminated by mutation of *sigE*, but still detected (in the mother cell) in 21% of *sigF* cells (Figure 3.9B). Labeling of the SNAP protein produced from the $P_{spoIIIAA}$-*SNAP*^{Cd} fusion was quantitative (Figure 3.10A), and the quantitative analysis of the average fluorescence signal shows no significant difference in expression levels before or after engulfment completion (Figure 3.10B). $P_{spoIIIAA}$-*SNAP*^{Cd} expression persisted until late stages in development, and was still detected for cells in which phase bright spores were seen (Figure 3.9B).

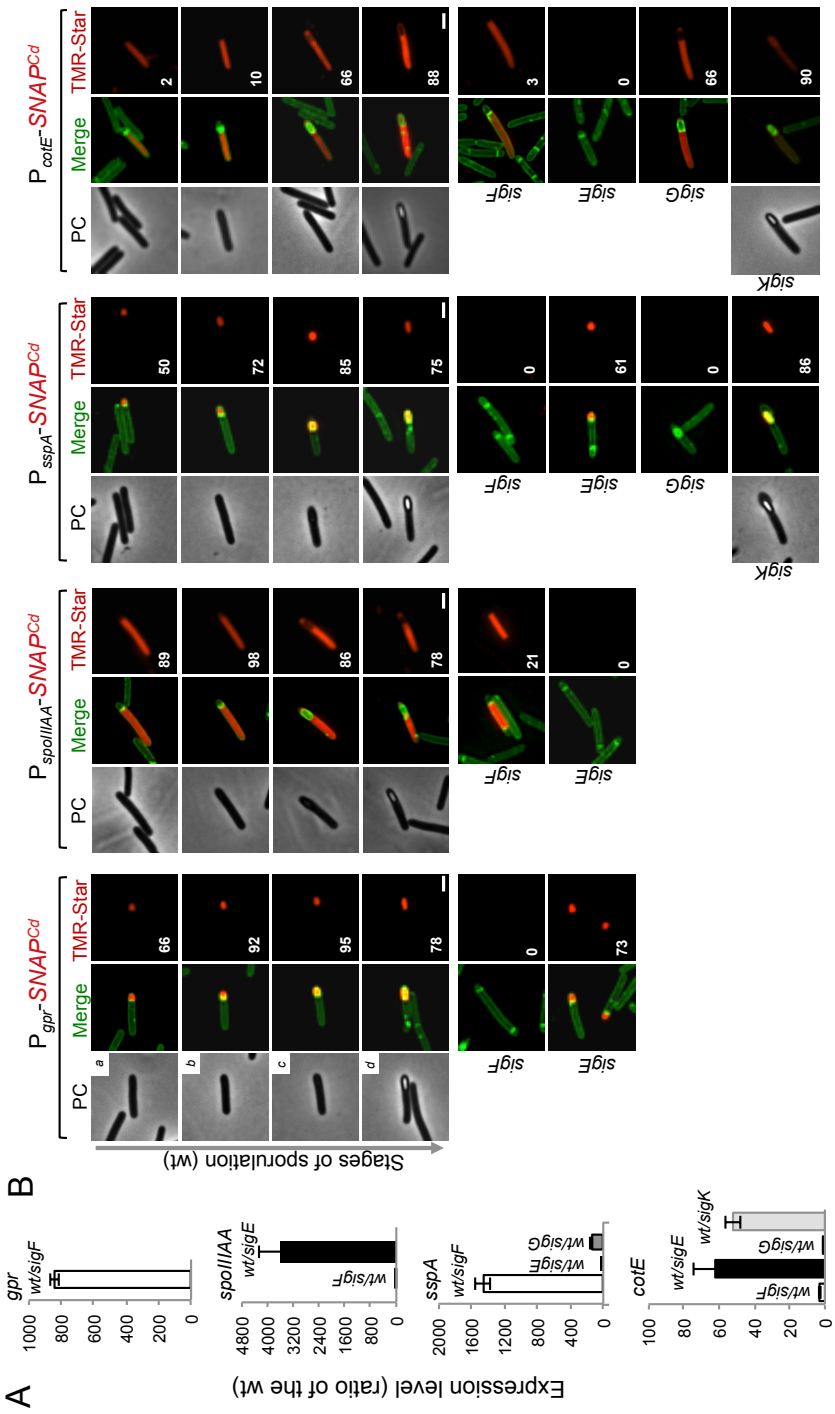

A
B
Expression level (ratio of the wt)
gpr
wt/sigF
spoIIIAA
wt/sigE
wt/sigF
sspA
wt/sigF
wt/sigE
wt/sigG
cotE
wt/sigE
wt/sigF
wt/sigG
wt/sigK
Stages of sporulation (wt)
PC
Merge
TMR-Star
sigF
sigE
sigG
sigK

Figure 3.9

Figure 3.9. Dependencies for the activation of the cell type-specific sporulation sigma factors. (A) qRT-PCR analysis of *gpr*, *spoIIIAA*, *sspA* and *cotE* gene transcription in strain 630Δ*erm* and in congenic *sigF*, *sigE*, *sigG* and *sigK* mutants. Expression is represented as the fold ratio between the wild type and the indicated mutants. Values are the average ± SD of two independent experiments. **(B)** Cell type-specific expression of transcriptional fusions of the *gpr*, *spoIIIAA*, *sspA* and *cotE* promoters to $SNAP^{Cd}$ in the wild type and in the indicated congenic mutants. Cells were collected after 24h of growth in SM medium, labeled with TMR-Star (red) and MTG (green), and imaged by phase contrast (PC) and fluorescence microscopy. Merge is the overlap between the TMR-Star (red) and MTG (green) channels. The images are ordered, and the morphological classes defined as in the legend for Figure 3.1. The numbers refer to the percentage of cells at the represented stage showing SNAP fluorescence. The data shown are from one experiment, representative of three independent experiments. The number of cells analysed for each class, n, is as follows: class *a*, 30-50; class *b*, n=50-60; class *c*, n=30-40; class *d*, n=15-25; for *sigF/E* mutants, n=80-120; for *sigG* and *sigK* mutants, n=40-50. Scale bar: 1 *μ*m.

Requirements for the activity of σ^G and σ^K

The *sspA* gene of *B. subtilis* codes for a small acid-soluble spore protein (SASP) that, together with other SASP family members, binds to and protects the spore DNA (Nicholson *et al.*, 2000). Expression of *sspA* in *B. subtilis* is controlled by σ^G (Wang *et al.*, 2006; Steil *et al.*, 2005) and a σ^G-type promoter can be recognized upstream of the *C. difficile* orthologue (*CD2688*) (see Chapter 4). Unexpectedly, a mutation in *sigF* caused a greater decrease in *sspA* transcription than disruption of *sigE* or *sigG*, in our qRT-PCR analysis (Figure 3.9A). While not excluding a contribution of σ^F to the expression of *sspA*, this result may be affected by the lack of synchronization of sporulation in the liquid SM cultures. Consistent with σ^G control of *sspA* in *C. difficile*, P_{sspA}-driven SNAP production was confined to the forespore and eliminated by disruption of *sigF* or of *sigG* (but not of *sigE* or *sigK)* (Figure 3.9B). This is in agreement with the requirement for σ^F for the transcription of *sigG* (above), and seems to exclude a contribution of σ^F for *sspA* transcription as suggested by the qRT-PCR analysis. *sspA* expression was detected in 50% of the cells that had just completed asymmetric division, but also throughout the engulfment sequence (72% of the cells), following engulfment completion (85% of the cells scored), and in cells (75%) carrying phase bright spores (Figure 3.9B). Our analysis of *sigG* transcription suggested that

it increased following engulfment completion, with a stronger auto-regulatory component than in pre-engulfed cells (above). In *B. subtilis*, continued transcription in the forespore when (upon engulfment completion) it becomes isolated from the surrounding medium, requires the activity of σ^E (Doan *et al.*, 2009; Serrano *et al.*, 2008; Meisner *et al.*, 2008; Camp and Losick, 2008; Camp and Losick, 2009). To determine whether the activity of σ^G increased following engulfment completion in a manner that required σ^E, we quantified the SNAP-TMR signal in cells expressing P_{sspA}-*SNAP*^{Cd}. Control experiments showed that all the SNAP protein produced from the P_{sspA}-SNAP^{Cd} fusion was labeled with the TMR-Star substrate (Figure 3.10A). The average intensity of the SNAP-TMR signal increased from 1.6±0.5 before engulfment completion (classes *a* + *b*) to 2.1±04, following engulfment completion (classes *c* + *d*) ($p<0.0001$) (Figure 3.10B). This result is consistent with the analysis of *sigG* transcription (above) and indicates that the activity of σ^G increases following engulfment completion. Importantly, even though *sspA* expression was found for 61% of the *sigE* mutant cells, disruption of *sigE* reduced the average fluorescence signal in the forespore (1.6±0.7) to the level seen before engulfment completion for the wild type (1.6±0.5). We conclude that disruption of *sigE* does not prevent activity of σ^G prior to engulfment completion.

Finally, to monitor the activity of σ^K, we examined expression of the *cotE* gene, coding for an abundant spore coat protein in *C. difficile* (Permpoonpattana *et al.*, 2011). This gene has no counterpart in *B. subtilis*, but as shown above, production of a CotE-SNAP translational fusion was dependent on σ^K (Figure 3.5E) and a sequence that conforms well to the consensus for σ^K promoters of *B. subtilis* can be recognized in its promoter region (see Chapter 4). qRT-PCR experiments show that disruption of the *sigE* and *sigK* genes caused a much stronger reduction in the expression of *cotE* than mutations in *sigF* or *sigG* (Figure 3.9A). While not excluding a contribution from σ^E, the qRT-PCR data are in line with the interpretation

that the main regulator of *cotE* expression is σ^K (with σ^E driving production of σ^K). We note that the reduced effect of the *sigF* mutation on *cotE* expression is in agreement with the view that σ^E production is partially independent on σ^F, as discussed above. We also note that the reduced effect of the *sigG* mutation on *cotE* expression is in agreement with the morphological analysis and the data on the assembly of the CotE-SNAP fusion (Figure 3.5D and E), suggesting σ^K-dependent deposition of coat material independently of σ^G. Fluorescence microscopy reveals that expression of P_{cotE}-*SNAP*^{Cd} is confined to the mother cell (Figure 3.9B). However, expression of P_{cotE}-*SNAP*^{Cd} was found just after asymmetric division in only 2% of the cells, and during engulfment in only 10% of the cells (Figure 3.9B). Expression increased to 66% of the cells after engulfment completion, and to 88% of the cells that showed phase bright spores (Figure 3.9B). Expression of P_{cotE}-*SNAP*^{Cd} was eliminated by disruption of *sigE,* but retained in 3% of the sporulating cells of a *sigF* mutant (Figure 3.9B). This is consistent with data presented above, also in line with the inference that the activity of σ^K is partially independent on *sigF* (Figure 3.9A and B). Moreover, 66% of the cells of a *sigG* mutant that had completed the engulfment process (as illustrated in Figure 3.9B) showed expression of the reporter fusion, again suggesting σ^K activity independently of σ^G. Interestingly, disruption of *sigK* did not abolish expression of the fusion, which was detected in 90% of the sporulating cells, but at low levels (Figure 3.9B). This raises the possibility that σ^E is responsible for the few cells that produce the reporter prior to engulfment completion.

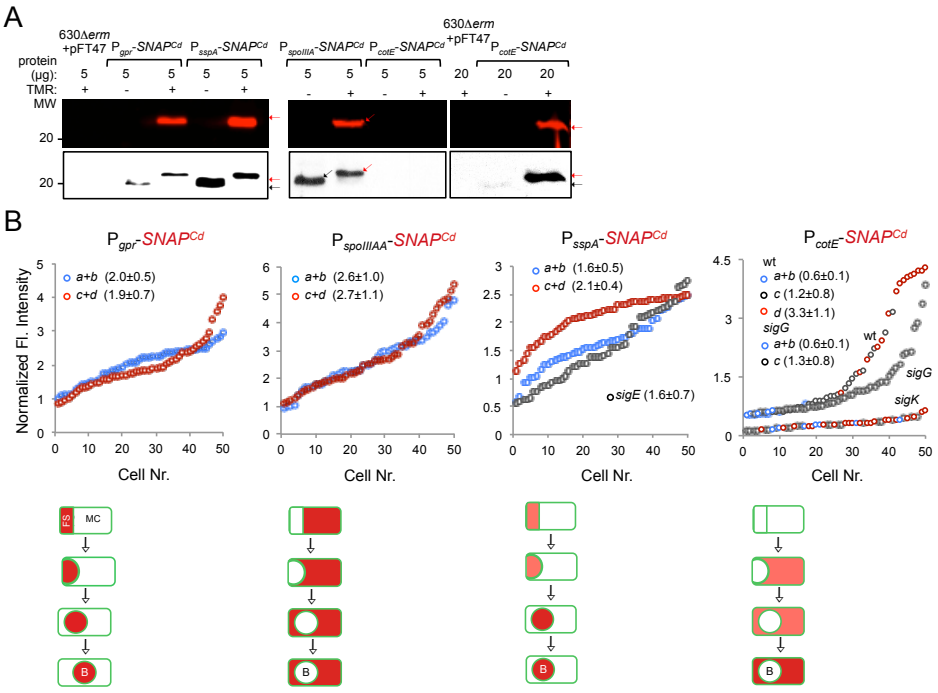


Figure 3.10. Quantitative analysis of *sigF*, *sigE*, *sigG* and *sigK* activities during sporulation. (A) SDS-PAGE gel and Western Blot analysis of extracts from 630Δ*erm* carrying fusions of *gpr*, *spoIIIAA*, *sspA* and *cotE* to SNAP. Black arrows point to unlabeled SNAP protein, while red arrows point to SNAP after TMR-Star labeling. 630Δ*erm* carrying pFT47 empty vector was used as a negative control of SNAP production. **(B)** Quantitative analysis of the SNAP fluorescence (Fl.) signal in different cell types of the reporter strains for σ^F, σ^E, σ^G and σ^K activity, as indicated. The numbers in the legend represent the average ± SD. The average fluorescence intensity (all classes included) from P_{cotE}-*SNAP*^{Cd} is 1.9±1.3 for the wild type, 1.3±0.8 for a *sigG* mutant and 0.3±0.1 for the *sigK* mutant. Data shown are from one experiment, representative of three independent experiments. Schematic representation of the deduced spatial and temporal pattern of transcription is shown for the different fusions (with darker red denoting increased transcription). PD, pre-divisional cell; MC, mother cell; FS, forespore; B, phase bright spore; *a* to *d*: sporulation classes ordered and defined as in the legend for Figure 3.1.

To test these possibilities quantitatively, we first verified that all the SNAP-tag produced from the P_{cotE}-*SNAP*^{Cd} fusion was labeled, under our experimental conditions (Figure 3.10A). The average intensity of the SNAP-TMR signal was of 0.6±0.1 for cells of the wild type strain prior to engulfment completion (classes *a* + *b*), of 1.2±0.8 for those that had just completed engulfment (class *c*), and of 3.3±1.1 for cells with phase bright spores (class *d*) (Figure 3.10B). Inactivation of *sigG* did not affect the expression level of

the fusion prior to engulfment completion (classes *a* + *b* for the *sigG* mutant, average signal, 0.6±0.1), nor did it prevent expression following engulfment completion (class *c* of the *sigG* mutant, 1.3±0.8) (Figure 3.10B). However, the average fluorescence signal for all classes of the *sigG* mutant (1.3±0.8) is significantly lower than the average for all classes of the wild type (1.9±1.3) ($p<0.001$) (Figure 3.10B). Finally, the average fluorescence signal for all cells of the *sigK* mutant was lower (0.3±0.1) than for pre-engulfment cells of the wild type (classes *a* + *b*, 0.6±0.1) ($p<0.0001$), suggesting that both σ^E and σ^K contribute to expression of the reporter fusion in these cells. Together, these data suggest that the main period of σ^K activity is delayed relative to engulfment completion, and coincides with development of spore refractility.

DISCUSSION

In this work, we analyzed the function of the four cell type-specific sigma factors of sporulation in *C. difficile*, and we studied gene expression in relation to the course of spore morphogenesis. The morphological characterization of mutants for the *sigG* genes allowed us to assign functions and to define the main periods of activity for the 4 cell type-specific sporulation sigma factors. In addition, the use of a fluorescence transcriptional reporter for single cell analysis enabled us to establish the time, cell type and dependency of transcription of the *sig* genes, as well as the time and requirements for activity of the four cell type-specific sigma factors.

Transcription of *sigF* and *sigE*, and activity of σ^F and σ^E

The cytological and TEM analysis shows that the *sigF* and *sigE* mutants are arrested just after asymmetric division. It follows that σ^F and σ^E control early stages of development in *C. difficile*, consistent with the function of these sigma factors in *B. subtilis*. Disruption of *sigE* also arrested

development just after asymmetric division in *C. perfringens* (Harry *et al.*, 2009). In contrast, disruption of either the *sigF* or *sigE* genes in *C. acetobutylicum* blocks sporulation prior to asymmetric division (Tracy *et al.*, 2011; Jones *et al.*, 2008; Jones *et al.*, 2011). In *C. difficile*, expression of both *sigF* and *sigE* commenced in predivisional cells, in line with work showing that expression of the *sigF*-containing operon (also coding for two other proteins, SpoIIAA and SpoIIAB, that control σ^F) occurs from a σ^H and Spo0A-controlled promoter, and with the observation that transcription of *sigE* is activated from a σ^A-type promoter to which Spo0A also binds (Rosenbusch *et al.*, 2012; Saujet *et al.*, 2011). In *B. subtilis*, following asymmetric septation, Spo0A becomes a cell-specific transcription factor, active predominantly in the mother cell (Fujita and Losick, 2003). This may also be the case in *C. difficile*, because transcription of *sigE* increased in the mother cell, relative to the forespore, following asymmetric division (Figure 3.8B).

In *B. subtilis*, σ^F is held in an inactive complex by the anti-sigma factor SpoIIAB (Hilbert and Piggot, 2004; Piggot and Hilbert, 2004). The reaction that releases σ^F takes place specifically in the forespore, soon after septation, and involves the anti-anti sigma factor SpoIIAA and the SpoIIE phosphatase. SpoIIAB, SpoIIAA and SpoIIE are produced in the *C. difficile* predivisional cell under Spo0A control (Galperin *et al.*, 2012; Traag *et al.*, 2012; Miller *et al.*, 2012; Abecasis *et al.*, 2013; Rosenbusch *et al.*, 2012; Saujet *et al.*, 2011). Because the activity of σ^F was confined to the forespore, we presume that the pathway leading to the forespore-specific activation of this sigma factor is also conserved. In *C. acetobutylicum*, this pathway may lead to σ^F activation in pre-divisional cells, as disruption of *sigF* or *spoIIE* blocks sporulation prior to asymmetric division (Jones *et al.*, 2011; Bi *et al.*, 2011).

In *B. subtilis*, σ^E is also synthesized in the predivisional but as an inactive pro-protein (Hilbert and Piggot, 2004; Piggot and Hilbert, 2004). Processing of pro-σ^E in the mother cell requires activation of the SpoIIGA protease by SpoIIR, a σ^F-controlled signaling protein secreted from the

forespore (Londono-Vallejo and Stragier, 1995; Karow *et al.*, 1995). Hence, the activity of σ^E requires the prior activation of σ^F. In *C. difficile*, the activity of σ^E was also restricted to the mother cell (Figure 3.9B). Because σ^E of *C. difficile* bears, like its *B. subtilis* counterpart, a pro sequence, and because the SpoIIGA protease and SpoIIR are conserved (Galperin *et al.*, 2012; Traag *et al.*, 2012; Miller *et al.*, 2012; Abecasis *et al.*, 2013), the σ^E activation pathway also seems conserved. Strikingly however, both the qRT-PCR and the SNAP labeling experiments showed that the activity of σ^E is at least partially independent on σ^F (Figure 3.9). Similar results were also obtained by Fimlaid *et al*. (2013). We do not know whether production of SpoIIR is also partially independent on σ^F. However, in *C. acetobutylicum*, in which σ^F is activated (and required) prior to asymmetric septation (Jones *et al.*, 2011; Bi *et al.*, 2011) production of SpoIIR is, at least in part, independent of σ^F (Jones *et al.*, 2011).

Production and activity of σ^G

The cytological and TEM analysis showed that a *sigG* mutant completes the engulfment sequence, suggesting that σ^G is mainly required for late stages in development, consistent with its role in *B. subtilis*. Disruption of *sigG* also causes a late morphological block in *C. acetobutylicum* (Tracy *et al.*, 2011). In *B. subtilis*, the forespore-specific transcription of *sigG* is initiated by σ^F but is delayed, relative to other σ^F-dependent genes, towards the engulfment sequence (Hilbert and Piggot, 2004; Partridge and Errington, 1993; Steil *et al.*, 2005). Moreover, the activity of σ^E, in the mother cell, is required for transcription of *sigG* (Hilbert and Piggot, 2004; Partridge and Errington, 1993). In contrast, transcription of *sigG* in *C. difficile*, was detected soon after asymmetric septation, and was not dependent on σ^E (Figure 3.7 and 3.8). Transcription of *sigG* also appears to be independent of *sigE* in *C. perfringens* (Harry *et al.*, 2009). The main period of *sigG* transcription in *B. subtilis* relies on an auto-regulatory loop activated

coincidently with engulfment completion (Serrano *et al.*, 2011). Therefore, the main period of σ^G activity coincides with engulfment completion. At least the anti-sigma factor CsfB (σ^F-controlled) appears important for impeding the σ^G auto-regulatory loop from functioning prior to engulfment completion, the main period of activity of the preceding forespore sigma factor, σ^F (Hilbert and Piggot, 2004; Piggot and Hilbert, 2004; Doan *et al.*, 2009; Chary *et al.*, 2007; Karmazyn-Campelli *et al.*, 2008). CsfB is absent from *C. difficile* as well as from other Clostridia (Galperin *et al.*, 2012; Traag *et al.*, 2012; Abecasis *et al.*, 2013). In *C. difficile*, not only is transcription of *sigG* observed soon after asymmetric division, but the activity of σ^G, is also detected prior to engulfment completion. Similar results were also obtained by other studies performed in parallel (Fimlaid *et al*., 2013). Nevertheless, our analysis indicates that σ^G activity increases following engulfment completion. In addition, our results suggest that σ^G is auto-regulatory both before, and more markedly, following engulfment completion.

A universal feature of endosporulation is the isolation of the forespore, surrounded by two membranes, from the external medium at the end of the engulfment sequence. In *B. subtilis*, the 8 mother cell proteins encoded by the *spoIIIA* operon, which localize to the forespore outer membrane, and the forespore-specific SpoIIQ protein, which localizes to the forespore inner membrane, are involved in the assembly of a specialized secretion system that links the cytoplasm of the two cells (Doan *et al.*, 2009; Serrano *et al.*, 2008; Meisner *et al.*, 2008; Camp and Losick, 2008; Camp and Losick, 2009). Recent work has shown that the SpoIIIA-SpoIIQ secretion system functions as a feeding tube required for continued macromolecular synthesis in the engulfed forespore (Camp and Losick, 2009). Mutation of *sigE* reduced the activity of σ^G but because the mutant is blocked at an early stage, we do not presently know whether σ^E is required for σ^G activity in the engulfed forespore. The SpoIIIAH and SpoIIQ proteins also facilitate forespore engulfment in *B. subtilis* (Broder and Pogliano, 2006). The *spoIIIA* operon is

conserved in sporeformers (Galperin *et al.*, 2012; Traag *et al.*, 2012; Abecasis *et al.*, 2013), and *spoIIIA* is under σ^E control in *C. difficile* (this work). A gene, *CD0125*, coding for a LytM-containing protein (as the *B. subtilis* SpoIIQ protein) may represent a non-orthologous gene replacement of *spoIIQ* (Galperin *et al.*, 2012). We do not yet know whether *spoIIIA* and *CD0125* are essential for sporulation in *C. difficile* and if so, whether they are required for engulfment and/or continued gene expression in the engulfed forespore.

Production and activity of σ^K

The TEM analysis shows that the *sigK* mutant of *C. difficile* lacks a visible coat (Figure 3.5D). However, as in *B. subtilis* (Henriques and Moran, 2007; McKenney *et al.*, 2012) assembly of the coat begins with σ^E, as suggested by the forespore localization of CotB-SNAP in cells of the *sigK* mutant, and supported by recent work on the analysis of coat morphogenetic proteins SpoIVA and SipL (Putnam *et al.*, 2013). Most likely, σ^K controls the final stages in the assembly of the spore surface structures, including the coat and exosporium. However, σ^K is not a strict requirement for the formation of heat resistant spores (Figure 3.5 and Table 3.2), and we presume that σ^E and σ^G (see above) are largely responsible for synthesis of the spore cortex. Final assembly of the coat together with the role of *C. difficile* σ^K in mother cell autolysis, are functions shared with its *B. subtilis* counterpart.

Transcription and activity of the *C. difficile sigK* gene was dependent on *sigE*, and was detected at low levels prior to engulfment completion. However, both transcription and activity increased, following engulfment completion, coincidently with the appearance of phase grey and phase bright spores. Transcription of the *sigK* and *spoIVCA* genes of *B. subtilis*, the latter coding for the recombinase that excises the *skin* element, is initiated under the control of σ^E with the assistance of the regulatory protein SpoIIID, and is delayed relative to a first wave of σ^E-directed genes (Hilbert and Piggot,

2004; Piggot and Hilbert, 2004; Steil *et al.*, 2005; Eichenberger *et al.*, 2004). SpoIIID is conserved in *C. difficile* (Abecasis *et al.*, 2013) and it may only accumulate to levels sufficient to enhance *sigK* and *spoIVCA* transcription at late stages in morphogenesis. Two observations highlight the importance of the *skin* element in *C. difficile*. First, with the exception of an asporogenous strain of *C. tetani*, the *skin* element is not present in other Clostridial species (de Hoon *et al.*, 2010; Haraldsen and Sonenshein, 2003). Second, not only a *skin*-less allele of *sigK* fails to complement a *sigK* mutation but also acts as a dominant negative mutation (Haraldsen and Sonenshein, 2003), blocking entry into sporulation (Figure 3.6D and E) (while these results seem to imply that σ^K is auto-regulatory, we did not detect auto-regulation of *sigK* in our single cell analysis). Absence of the *skin* element may allow the recruitment of σ^K for other functions. In *C. perfringens* and in *C. botulinum*, σ^K is produced in pre-divisional cells, and is involved in enterotoxin production in the first, and in cold and osmotic stress tolerance in the second (Harry *et al.*, 2009; Dahlsten *et al.*, 2013).

A key finding of the present study, and corroborated by other studies (Fimlaid *et al*., 2013; Saujet *et al*., 2013; see also Chapter 4), is that contrary to *B. subtilis*, *sigG* is not essential for the activity of σ^K. In *B. subtilis* a signaling protein, SpoIVB, secreted from the forespore activates the pro-σ^K processing protease SpoIVFB, which is kept inactive in a complex with BofA and SpoIVFA, embedded in the forespore outer membrane (Hilbert and Piggot, 2004; Higgins and Dworkin, 2012). SpoIVFB, BofA and possibly also SpoIVFA are absent from *C. difficile*, suggesting that the σ^G to σ^K pathway is absent and consistent with the lack of a pro-sequence (Galperin *et al.*, 2012; Abecasis *et al.*, 2013; Stragier, 2002). However, *C. difficile* codes for two orthologues of SpoIVB (Galperin *et al.*, 2012; Traag *et al.*, 2012; Abecasis *et al.*, 2013). Mutations that bypass the need for *sigG* or *spoIVB* in *B. subtilis* result in coat deposition, but not cortex formation, phenocopying the *sigG* mutant of *C. difficile* (Cutting *et al.*, 1990). In *B. subtilis*, SpoIVB is also

required for the engulfment-regulated proteolysis of SpoIIQ (Jiang *et al.*, 2005). The *C. difficile* SpoIVB orthologues may be involved in cortex formation and/or proteolysis of *CD0125* (above).

While the activity of σ^K did not require σ^G, our data shows that mutation of *sigG* reduced the activity of σ^K at late stages of spore morphogenesis. Because the *sigG* mutant fails to form phase grey/bright spores, we do not presently know if a forespore-mother cell signaling operates at this stage, or whether the late stages in spore morphogenesis serve as a cue for enhanced activity of σ^K.

In conclusion, we show that the main periods of activity of the four cell type-specific sigma factors of *C. difficile* are conserved, relative to the *B. subtilis* model, with σ^F and σ^E controlling early stages of development and σ^G and σ^K governing late developmental events (Figure 3.11A and B). However, the fact that the activity of σ^E was partially independent of σ^F, and that σ^G or σ^K did not require σ^E or σ^G, respectively, seems to imply a weaker connection between the forespore and mother cell lines of gene expression. In spite of the important differences in the roles of the sporulation sigma factors and the regulatory circuits leading to their activation (Figure 3.11A and B), overall, in what concerns the genetic control of sporulation, *C. difficile* seems closer to the model organism *B. subtilis* than the other Clostridial species that have been studied. Differences in the function/period of activity of the sporulation sigma factors in other Clostridial species, may be related to the coordination of solventogenesis, toxin production or other functions with sporulation (Paredes *et al.*, 2005; Harry *et al.*, 2009 ;Li and McClane, 2010) Tracy *et al.*, 2011; Jones *et al.*, 2008; Jones *et al.*, 2011; Bi *et al.*, 2011; Kirk *et al.*, 2012). We note however that the relationship between toxinogenesis and spore formation in *C. difficile* is still unclear (Deakin *et al.*, 2012; Rosenbusch et al., 2012). In general, the conclusions herein presented were in agreement with those of Fimlaid *et al*. (2013). Discrepancies will be considered in the General Discussion of this Thesis.

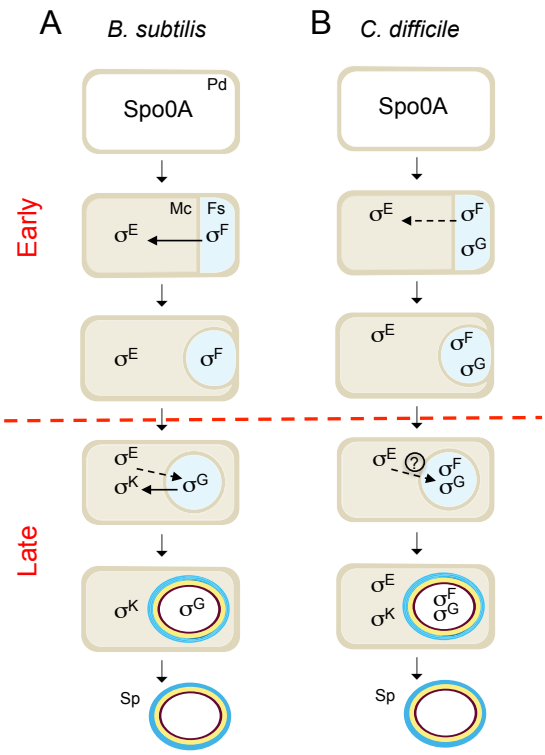


Figure 3.11. Stages and cell of σ^F, σ^E, σ^G and σ^K activity. The figure compares the main periods of activity of the 4 cell type-specific sigma factors of sporulation in *B. subtilis* **(A)** and *C. difficile* **(B)**. The figure incorporates data on the morphological analysis of sporulation in the *sigF*, *sigE*, *sigG* and *sigK* mutants (Figure 3.5, 3.7 and 3.9) and on the stage, dependencies and cell where σ^F, σ^E , σ^G and σ^K are active. Solid or broken arrows represent dependencies or partial dependencies, respectively. The representation of the *C. difficile* sigma factors indicates activity; black indicates the main period of activity. Possible cell-cell signaling pathways are show by both a broken line and a question mark. The SpoIIIA-SpoIIQ/CD0125 channel is represented in yellow. PD: predivisional cell; MC: mother cell; FS: forespore. The red horizontal broken line distinguishes early (prior to engulfment completion) from late (post-engulfment completion) development.

ACKNOWLEDGEMENTS

The author of this dissertation performed all the experiments described in this chapter except: 1) the analysis of gene transcription by qRT-PCR, which were perform by Laure Saujet and Marc Monot; 2) the steps involved in the preparation of samples for Electron Microscopy, following sample fixation, which were performed by Evelyne Couture-Tosi. We thank Catarina Fernandes for the gift of the Tgl-CFP fusion, Nigel Minton for the gift of plasmids, strains and helpful discussions, Maja Rupnik and Jorge Carneiro for helpful discussions and Tiago dos Vultos for help in the initial stages of this study. This work was supported by grant ERA-PTG/SAU/0002/2008

(ERA-NET PathoGenoMics) and PEst-OE/EQB/LA0004/2011 from the "Fundação para a Ciência e a Tecnologia" (FCT). F.P. (SFRH/BD/45459/08) was the recipient of doctoral fellowship from the FCT.

REFERENCES

Abecasis, A., Serrano, M., Alves, L., Quintais, L., Pereira-Leal, J. B. & Henriques, A. O. (2013). A genomic signature and the identification of new endosporulaion genes. *J Bacteriol* **195**, 2101-2115.

Akerlund, T., Alefjord, I., Dohnhammar, U., Struwe, J., Noren, T. & Tegmark-Wisell, K. Geographical clustering of cases of infection with moxifloxacin-resistant Clostridium difficile PCR-ribotypes 012, 017 and 046 in Sweden, 2008 and 2009. *Euro Surveill* **16**.

Angert, E. R. & Losick, R. M. (1998). Propagation by sporulation in the guinea pig symbiont Metabacterium polyspora. *Proc Natl Acad Sci U S A* **95**, 10218-23.

Balomenou, S., Fouet, A., Tzanodaskalaki, M., Couture-Tosi, E., Bouriotis, V. & Boneca, I. G. (2013). Distinct functions of polysaccharide deacetylases in cell shape, neutral polysaccharide synthesis and virulence of Bacillus anthracis. *Mol Microbiol*.

Bi, C., Jones, S. W., Hess, D. R., Tracy, B. P. & Papoutsakis, E. T. (2011). SpoIIE is necessary for asymmetric division, sporulation, and expression of sigmaF, sigmaE, and sigmaG but does not control solvent production in Clostridium acetobutylicum ATCC 824. *J Bacteriol* **193**, 5130-7.

Broder, D. H. & Pogliano, K. (2006). Forespore engulfment mediated by a ratchet-like mechanism. *Cell* **126**, 917-28.

Burns, D. A., Heap, J. T. & Minton, N. P. (2010). The diverse sporulation characteristics of Clostridium difficile clinical isolates are not associated with type. *Anaerobe* **16**, 618-22.

Burns, D. A., Heeg, D., Cartman, S. T. & Minton, N. P. (2011). Reconsidering the sporulation characteristics of hypervirulent Clostridium difficile BI/NAP1/027. *PLoS One* **6**, e24894.

Burns, D. A. & Minton, N. P. (2011). Sporulation studies in Clostridium difficile. *J Microbiol Methods* **87**, 133-8.

Camp, A. H. & Losick, R. (2008). A novel pathway of intercellular signalling in Bacillus subtilis involves a protein with similarity to a component of type III secretion channels. *Mol Microbiol* **69**, 402-17.

Camp, A. H. & Losick, R. (2009). A feeding tube model for activation of a cell-specific transcription factor during sporulation in Bacillus subtilis. *Genes Dev* **23**, 1014-24.

Carter, G. P., Rood, J. I. & Lyras, D. (2012). The role of toxin A and toxin B in the virulence of Clostridium difficile. *Trends Microbiol* **20**, 21-9.

Cartman, S. T., Heap, J. T., Kuehne, S. A., Cockayne, A. & Minton, N. P. (2010). The emergence of 'hypervirulence' in Clostridium difficile. *Int J Med Microbiol* **300**, 387-95.

Chary, V. K., Xenopoulos, P. & Piggot, P. J. (2007). Expression of the sigmaF-directed csfB locus prevents premature appearance of sigmaG activity during sporulation of Bacillus subtilis. *J Bacteriol* **189**, 8754-7.

Cutting, S., Oke, V., Driks, A., Losick, R., Lu, S. & Kroos, L. (1990). A forespore checkpoint for mother cell gene expression during development in B. subtilis. *Cell* **62**, 239-50.

Dahlsten, E., Kirk, D., Lindstrom, M. & Korkeala, H. (2013). Alternative Sigma Factor SigK Has a Role in Stress Tolerance of Group I Clostridium botulinum Strain ATCC 3502. *Appl Environ Microbiol* **79**, 3867-9.

de Hoon, M. J., Eichenberger, P. & Vitkup, D. (2010). Hierarchical evolution of the bacterial sporulation network. *Curr Biol* **20**, R735-45.

Deakin, L. J., Clare, S., Fagan, R. P., Dawson, L. F., Pickard, D. J., West, M. R., Wren, B. W., Fairweather, N. F., Dougan, G. & Lawley, T. D. (2012). The Clostridium difficile spo0A gene is a persistence and transmission factor. *Infect Immun* **80**, 2704-11.

Deneve, C., Janoir, C., Poilane, I., Fantinato, C. & Collignon, A. (2009). New trends in Clostridium difficile virulence and pathogenesis. *Int J Antimicrob Agents* **33 Suppl 1**, S24-8.

Doan, T., Morlot, C., Meisner, J., Serrano, M., Henriques, A. O., Moran, C. P., Jr. & Rudner, D. Z. (2009). Novel secretion apparatus maintains spore integrity and developmental gene expression in Bacillus subtilis. *PLoS Genet* **5**, e1000566.

Eichenberger, P., Fawcett, P. & Losick, R. (2001). A three-protein inhibitor of polar septation during sporulation in Bacillus subtilis. *Mol Microbiol* **42**, 1147-62.

Eichenberger, P., Fujita, M., Jensen, S. T., Conlon, E. M., Rudner, D. Z., Wang, S. T., Ferguson, C., Haga, K., Sato, T., Liu, J. S. & Losick, R. (2004). The program of gene transcription for a single differentiating cell type during sporulation in Bacillus subtilis. *PLoS Biol* **2**, e328.

Eichenberger, P., Jensen, S. T., Conlon, E. M., van Ooij, C., Silvaggi, J., Gonzalez-Pastor, J. E., Fujita, M., Ben-Yehuda, S., Stragier, P., Liu, J. S. & Losick, R. (2003). The sigmaE regulon and the identification of additional sporulation genes in Bacillus subtilis. *J Mol Biol* **327**, 945-72.

Eldar, A., Chary, V. K., Xenopoulos, P., Fontes, M. E., Loson, O. C., Dworkin, J., Piggot, P. J. & Elowitz, M. B. (2009). Partial penetrance facilitates developmental evolution in bacteria. *Nature* **460**, 510-4.

Fimlaid, K. A., Bond, J. P., Schutz, K. C., Putnam, E. E., Leung, J. M., Lawley, T. D. & Shen, A. (2013). Global analysis of the sporulation pathway of Clostridium difficile. *PLoS Genet* **9**, e1003660.

Flint, J. F., Drzymalski, D., Montgomery, W. L., Southam, G. & Angert, E. R. (2005). Nocturnal production of endospores in natural populations of epulopiscium-like surgeonfish symbionts. *J Bacteriol* **187**, 7460-70.

Fujita, M. & Losick, R. (2003). The master regulator for entry into sporulation in Bacillus subtilis becomes a cell-specific transcription factor after asymmetric division. *Genes Dev* **17**, 1166-74.

Galperin, M. Y., Mekhedov, S. L., Puigbo, P., Smirnov, S., Wolf, Y. I. & Rigden, D. J. (2012). Genomic determinants of sporulation in Bacilli and Clostridia: towards the minimal set of sporulation-specific genes. *Environ Microbiol* **14**, 2870-90.

George, W. L., Sutter, V. L., Citron, D. & Finegold, S. M. (1979). Selective and differential medium for isolation of Clostridium difficile. *J Clin Microbiol* **9**, 214-9.

Haraldsen, J. D. & Sonenshein, A. L. (2003). Efficient sporulation in Clostridium difficile requires disruption of the sigmaK gene. *Mol Microbiol* **48**, 811-21.

Harry, K. H., Zhou, R., Kroos, L. & Melville, S. B. (2009). Sporulation and enterotoxin (CPE) synthesis are controlled by the sporulation-specific sigma factors SigE and SigK in Clostridium perfringens. *J Bacteriol* **191**, 2728-42.

Heap, J. T., Pennington, O. J., Cartman, S. T., Carter, G. P. & Minton, N. P. (2007). The ClosTron: a universal gene knock-out system for the genus Clostridium. *J Microbiol Methods* **70**, 452-64.

Heap, J. T., Pennington, O. J., Cartman, S. T. & Minton, N. P. (2009). A modular system for Clostridium shuttle plasmids. *J Microbiol Methods* **78**, 79-85.

Henriques, A. O., Beall, B. W., Roland, K. & Moran, C. P., Jr. (1995). Characterization of cotJ, a sigma E-controlled operon affecting the polypeptide composition of the coat of Bacillus subtilis spores. *J Bacteriol* **177**, 3394-406.

Henriques, A. O. & Moran, C. P., Jr. (2007). Structure, assembly, and function of the spore surface layers. *Annu Rev Microbiol* **61**, 555-88.

Higgins, D. & Dworkin, J. (2012). Recent progress in Bacillus subtilis sporulation. *FEMS Microbiol Rev* **36**, 131-48.

Hilbert, D. W. & Piggot, P. J. (2004). Compartmentalization of gene expression during Bacillus subtilis spore formation. *Microbiol Mol Biol Rev* **68**, 234-62.

Hussain, H. A., Roberts, A. P. & Mullany, P. (2005). Generation of an erythromycin-sensitive derivative of Clostridium difficile strain 630 (630Deltaerm) and demonstration that the conjugative transposon Tn916DeltaE enters the genome of this strain at multiple sites. *J Med Microbiol* **54**, 137-41.

Jiang, X., Rubio, A., Chiba, S. & Pogliano, K. (2005). Engulfment-regulated proteolysis of SpoIIQ: evidence that dual checkpoints control sigma activity. *Mol Microbiol* **58**, 102-15.

Jones, S. W., Paredes, C. J., Tracy, B., Cheng, N., Sillers, R., Senger, R. S. & Papoutsakis, E. T. (2008). The transcriptional program underlying the physiology of clostridial sporulation. *Genome Biol* **9**, R114.

Jones, S. W., Tracy, B. P., Gaida, S. M. & Papoutsakis, E. T. (2011). Inactivation of sigmaF in Clostridium acetobutylicum ATCC 824 blocks sporulation prior to asymmetric division and abolishes sigmaE and sigmaG protein expression but does not block solvent formation. *J Bacteriol* **193**, 2429-40.

Joshi, L. T., Phillips, D. S., Williams, C. F., Alyousef, A. & Baillie, L. (2012). Contribution of spores to the ability of Clostridium difficile to adhere to surfaces. *Appl Environ Microbiol* **78**, 7671-9.

Karmazyn-Campelli, C., Bonamy, C., Savelli, B. & Stragier, P. (1989). Tandem genes encoding sigma-factors for consecutive steps of development in Bacillus subtilis. *Genes Dev* **3**, 150-7.

Karmazyn-Campelli, C., Rhayat, L., Carballido-Lopez, R., Duperrier, S., Frandsen, N. & Stragier, P. (2008). How the early sporulation sigma factor sigmaF delays the switch to late development in Bacillus subtilis. *Mol Microbiol* **67**, 1169-80.

Karow, M. L., Glaser, P. & Piggot, P. J. (1995). Identification of a gene, spoIIR, that links the activation of sigma E to the transcriptional activity of sigma F during sporulation in Bacillus subtilis. *Proc Natl Acad Sci U S A* **92**, 2012-6.

Kirk, D. G., Dahlsten, E., Zhang, Z., Korkeala, H. & Lindstrom, M. (2012). Involvement of Clostridium botulinum ATCC 3502 sigma factor K in early-stage sporulation. *Appl Environ Microbiol* **78**, 4590-6.

Lawley, T. D., Croucher, N. J., Yu, L., Clare, S., Sebaihia, M., Goulding, D., Pickard, D. J., Parkhill, J., Choudhary, J. & Dougan, G. (2009). Proteomic and genomic characterization of highly infectious Clostridium difficile 630 spores. *J Bacteriol* **191**, 5377-86.

Leighton, T. J. & Doi, R. H. (1971). The stability of messenger ribonucleic acid during sporulation in Bacillus subtilis. *J Biol Chem* **246**, 3189-95.

Li, J. & McClane, B. A. (2010). Evaluating the involvement of alternative sigma factors SigF and SigG in Clostridium perfringens sporulation and enterotoxin synthesis. *Infect Immun* **78**, 4286-93.

Londono-Vallejo, J. A. & Stragier, P. (1995). Cell-cell signaling pathway activating a developmental transcription factor in Bacillus subtilis. *Genes Dev* **9**, 503-8.

Lu, S., Halberg, R. & Kroos, L. (1990). Processing of the mother-cell sigma factor, sigma K, may depend on events occurring in the forespore during Bacillus subtilis development. *Proc Natl Acad Sci U S A* **87**, 9722-6.

Magge, A., Setlow, B., Cowan, A. E. & Setlow, P. (2009). Analysis of dye binding by and membrane potential in spores of Bacillus species. *J Appl Microbiol* **106**, 814-24.

McKenney, P. T., Driks, A. & Eichenberger, P. (2012). The Bacillus subtilis endospore: assembly and functions of the multilayered coat. *Nat Rev Microbiol* **11**, 33-44.

Meisner, J., Wang, X., Serrano, M., Henriques, A. O. & Moran, C. P., Jr. (2008). A channel connecting the mother cell and forespore during bacterial endospore formation. *Proc Natl Acad Sci U S A* **105**, 15100-5.

Miller, D. A., Suen, G., Clements, K. D. & Angert, E. R. (2012). The genomic basis for the evolution of a novel form of cellular reproduction in the bacterium Epulopiscium. *BMC Genomics* **13**, 265.

Nicholson, W. L., Munakata, N., Horneck, G., Melosh, H. J. & Setlow, P. (2000). Resistance of Bacillus endospores to extreme terrestrial and extraterrestrial environments. *Microbiol Mol Biol Rev* **64**, 548-72.

Oliva, C., Turnbough, C. L., Jr. & Kearney, J. F. (2009). CD14-Mac-1 interactions in Bacillus anthracis spore internalization by macrophages. *Proc Natl Acad Sci U S A* **106**, 13957-62.

Panessa-Warren, B. J., Tortora, G. T. & Warren, J. B. (2007). High resolution FESEM and TEM reveal bacterial spore attachment. *Microsc Microanal* **13**, 251-66.

Paredes, C. J., Alsaker, K. V. & Papoutsakis, E. T. (2005). A comparative genomic view of clostridial sporulation and physiology. *Nat Rev Microbiol* **3**, 969-78.

Paredes-Sabja, D., Cofre-Araneda, G., Brito-Silva, C., Pizarro-Guajardo, M. & Sarker, M. R. (2012). Clostridium difficile spore-macrophage interactions: spore survival. *PLoS One* **7**, e43635.

Paredes-Sabja, D. & Sarker, M. R. (2012). Adherence of Clostridium difficile spores to Caco-2 cells in culture. *J Med Microbiol* **61**, 1208-18.

Partridge, S. R. & Errington, J. (1993). The importance of morphological events and intercellular interactions in the regulation of prespore-specific gene expression during sporulation in Bacillus subtilis. *Mol Microbiol* **8**, 945-55.

Permpoonpattana, P., Phetcharaburanin, J., Mikelsone, A., Dembek, M., Tan, S., Brisson, M. C., La Ragione, R., Brisson, A. R., Fairweather, N., Hong, H. A. & Cutting, S. M. (2013). Functional characterization of Clostridium difficile spore coat proteins. *J Bacteriol* **195**, 1492-503.

Permpoonpattana, P., Tolls, E. H., Nadem, R., Tan, S., Brisson, A. & Cutting, S. M. (2011). Surface layers of Clostridium difficile endospores. *J Bacteriol* **193**, 6461-70.

Piggot, P. J. & Coote, J. G. (1976). Genetic aspects of bacterial endospore formation. *Bacteriol Rev* **40**, 908-62.

Piggot, P. J. & Hilbert, D. W. (2004). Sporulation of Bacillus subtilis. *Curr Opin Microbiol* **7**, 579-86.

Pogliano, J., Osborne, N., Sharp, M. D., Abanes-De Mello, A., Perez, A., Sun, Y. L. & Pogliano, K. (1999). A vital stain for studying membrane dynamics in bacteria: a novel mechanism controlling septation during Bacillus subtilis sporulation. *Mol Microbiol* **31**, 1149-59.

Putnam, E. E., Nock, A. M., Lawley, T. D. & Shen, A. (2013). SpoIVA and SipL are Clostridium difficile spore morphogenetic proteins. *J Bacteriol*.

Reid, B. G. & Flynn, G. C. (1997). Chromophore formation in green fluorescent protein. *Biochemistry* **36**, 6786-91.

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

Rupnik, M., Wilcox, M. H. & Gerding, D. N. (2009). Clostridium difficile infection: new developments in epidemiology and pathogenesis. *Nat Rev Microbiol* **7**, 526-36.

Sarker, M. R. & Paredes-Sabja, D. (2012). Molecular basis of early stages of Clostridium difficile infection: germination and colonization. *Future Microbiol* **7**, 933-43.

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**, 3186-96.

Saujet, L., Pereira, F. C., Serrano, M., Soutourina, O., Monot, M., Shelyakin, P. V., Gelfand, M. S., Dupuy, B., Henriques, A. O. & Martin-Verstraete, I. (2013). Genome-wide analysis of cell type-specific gene expression during spore formation in Clostridium difficile.

Schuch, R., Nelson, D. & Fischetti, V. A. (2002). A bacteriolytic agent that detects and kills Bacillus anthracis. *Nature* **418**, 884-9.

Serrano, M., Real, G., Santos, J., Carneiro, J., Moran, C. P., Jr. & Henriques, A. O. (2011). A negative feedback loop that limits the ectopic activation of a cell type-specific sporulation sigma factor of Bacillus subtilis. *PLoS Genet* **7**, e1002220.

Serrano, M., Vieira, F., Moran, C. P., Jr. & Henriques, A. O. (2008). Processing of a membrane protein required for cell-to-cell signaling during endospore formation in Bacillus subtilis. *J Bacteriol* **190**, 7786-96.

Steil, L., Serrano, M., Henriques, A. O. & Volker, U. (2005). Genome-wide analysis of temporally regulated and compartment-specific gene expression in sporulating cells of Bacillus subtilis. *Microbiology* **151**, 399-420.

Stragier, P. (2002). A gene odyssey: exploring the genomes of endospore-forming bacteria. In *Bacillus subtilis and its closest relatives: from genes to cells.* (ed. S. AL), pp. 519-525. ASM: Washington

Sun, D. X., Cabrera-Martinez, R. M. & Setlow, P. (1991). Control of transcription of the Bacillus subtilis spoIIIG gene, which codes for the forespore-specific transcription factor sigma G. *J Bacteriol* **173**, 2977-84.

Traag, B. A., Pugliese, A., Eisen, J. A. & Losick, R. (2012). Gene conservation among endospore-forming bacteria reveals additional sporulation genes in Bacillus subtilis. *J Bacteriol*.

Tracy, B. P., Jones, S. W. & Papoutsakis, E. T. (2011). Inactivation of sigmaE and sigmaG in Clostridium acetobutylicum illuminates their roles in clostridial-cell-form biogenesis, granulose synthesis, solventogenesis, and spore morphogenesis. *J Bacteriol* **193**, 1414-26.

Vasudevan, P., Weaver, A., Reichert, E. D., Linnstaedt, S. D. & Popham, D. L. (2007). Spore cortex formation in Bacillus subtilis is regulated by accumulation of peptidoglycan precursors under the control of sigma K. *Mol Microbiol* **65**, 1582-94.

Vida, T. A. & Emr, S. D. (1995). A new vital stain for visualizing vacuolar membrane dynamics and endocytosis in yeast. *J Cell Biol* **128**, 779-92.

Walk, S. T., Jain, R., Trivedi, I., Grossman, S., Newton, D. W., Thelen, T., Hao, Y., Songer, J. G., Carter, G. P., Lyras, D., Young, V. B. & Aronoff, D. M. Non-toxigenic Clostridium sordellii: clinical and microbiological features of a case of cholangitis-associated bacteremia. *Anaerobe* **17**, 252-6.

Wang, S. T., Setlow, B., Conlon, E. M., Lyon, J. L., Imamura, D., Sato, T., Setlow, P., Losick, R. & Eichenberger, P. (2006). The forespore line of gene expression in Bacillus subtilis. *J Mol Biol* **358**, 16-37.

Wilson, K. H., Kennedy, M. J. & Fekety, F. R. (1982). Use of sodium taurocholate to enhance spore recovery on a medium selective for Clostridium difficile. *J Clin Microbiol* **15**, 443-6.

Chapter 4

Genome-wide analysis of cell-type specific gene transcription during spore formation in *C. difficile*

This chapter contains data published in: Saujet, L., Pereira, F. C., Serrano, M., Monot, M., Shelyakin, P. V., Gelfand, M. S., Dupuy, B., Martin-Verstraete, I., and Henriques, A. O. (2013). Genome-wide analysis of cell type-specific gene expression during spore formation in *Clostridium difficile*. *PLoS Genet* 9: e1003756.

SUMMARY

C. difficile, a Gram positive, anaerobic, spore-forming bacterium is an emergent pathogen and the most common cause of nosocomial enteric infections in adults. Although transmission of *C. difficile* is mediated by contamination of the gut by spores, the regulatory cascade controlling spore formation remains poorly characterized. During *B. subtilis* sporulation, a cascade of four sigma factors, σ^F and σ^G in the forespore and σ^E and σ^K in the mother cell governs compartment-specific gene expression. In this work, we combined genome wide transcriptional analyses and promoters mapping to define the *C. difficile* σ^F, σ^E, σ^G and σ^K regulons. We identified about 225 genes under the control of these sigma factors: 25 in the σ^F regulon, 97 σ^E-dependent genes, 50 σ^G-governed genes and 56 genes under σ^K control. A significant fraction of genes in each regulon is of unknown function but new candidates for spore coat proteins could be proposed as being synthesized under σ^E or σ^K control and detected in a previously published spore proteome analysis. SpoIIID of *C. difficile* also plays a pivotal role in the mother cell line of expression repressing the transcription of many members of the σ^E regulon and activating *sigK* expression. Global analysis of developmental gene expression under the control of these sigma factors revealed deviations from the *B. subtilis* model regarding the communication between mother cell and forespore in *C. difficile*. We showed that the expression of the σ^E regulon in the mother cell was not strictly under the control of σ^F despite the fact that the forespore product SpoIIR was required for the processing of pro-σ^E. In addition, the σ^K regulon was not controlled by σ^G in *C. difficile* in agreement with the lack of pro-σ^K processing. This work is one key step to obtain new insights about the diversity and evolution of the sporulation process among Firmicutes.

INTRODUCTION

C. difficile is nowadays a major cause of nosocomial infections associated with antibiotic therapy and is a major burden to health care services. Transmission of *C. difficile* is further mediated by contamination of the gut by spores as demonstrated recently using a murine model for CDI (Deakin *et al.*, 2012; Lawley *et al.*, 2010). The disruption of the colonic microbiota by antimicrobial therapy precipitates colonization of the intestinal tract and *C. difficile* infection (CDI). An early event towards this colonization is the germination process that converts dormant spores into vegetative cells that multiply leading to the production of toxins (Sarker and Paredes-Sabja, 2012). Glycine and bile salts like sodium cholate or sodium taurocholate are co-germinants required to induce *C. difficile* spores germination (Sorg and Sonenshein, 2008). However, the molecular mechanisms involved in sporulation and germination are still poorly studied in *C. difficile* and our current knowledge on these processes is based mainly on data derived from the studies on *B. subtilis*.

At the onset of sporulation in *B. subtilis*, sporulating cells undergo an asymmetric division which partitions the sporangial cell into a larger mother cell and a smaller forespore (the future spore). The forespore is next wholly engulfed by the mother cell and later the dormant spore is released from the mother cell by lysis (Hilbert and Piggot, 2004). The developmental program of sporulation is mainly governed by the sequential activation of four sigma factors: σ^F, σ^E, σ^G and σ^K. Their activity is confined to the forespore for σ^F and σ^G and to the mother cell for σ^E and σ^K. Compartmentalization of gene expression is coupled to morphogenesis with σ^F and σ^E becoming active after asymmetric division and σ^G and σ^K after completion of engulfment of the

forespore by the mother cell (Stragier and Losick, 1996; Hilbert and Piggot, 2004; Higgins and Dworkin, 2012).

In *B. subtilis,* coordinated changes in gene expression underlie morphological differentiation in both the pre-divisional sporangium and later in the two cells resulting from asymmetric division, with the existence of communication between the mother cell and the forespore. The response regulator, Spo0A, and a phosphorelay involving five kinases, intermediary phosphorylated proteins and phosphatases control sporulation initiation. The alternative sigma factor, σ^H, which transcribes *spo0A* and *sigF* also controls early sporulation steps. When Spo0A-P level reaches a critical threshold, Spo0A-P activates sporulation genes including *spoIIE* as well as both the *spoIIAA-spoIIAB-sigF* and the *spoIIGA-sigE* operons encoding σ^F and σ^E, respectively. After its synthesis, σ^F is held inactive by the anti-sigma factor SpoIIAB until the phosphatase SpoIIE dephosphorylates the anti-anti sigma factor SpoIIAA leading to the release of an active σ^F from SpoIIAB after completion of asymmetric cell division (Hilbert and Piggot, 2004). σ^F then transcribes about 50 genes in the forespore (Steil *et al.*, 2005; Wang *et al.*, 2006) including *spoIIR* encoding a secretory protein required for the processing of pro-σ^E into active σ^E in the mother cell (Londono-Vallejo and Stragier, 1995). σ^E, in turn, regulates expression of mother cell specific genes (Eichenberger *et al.*, 2003; Evans *et al.*, 2003; Steil *et al.*, 2005) and together with the ancillary transcription factor SpoIIID, itself under σ^E control, activates *sigK* expression (Eichenberger *et al.*, 2004). In the forespore, σ^F also controls *sigG* transcription. However, σ^G becomes active coincidently with the completion of forespore engulfment by the mother cell. Following engulfment completion, the σ^E-controlled SpoIIIA proteins, together with the forespore-specific σ^F-controlled SpoIIQ protein, are required to maintain forespore stability and the potential for macromolecular synthesis in this cell (Camp and Losick, 2008; Higgins and Dworkin, 2012; Meisner *et al.*, 2012). The σ^G RNA polymerase holoenzyme drives transcription of 95 genes in the

forespore including those coding for proteins like SpoIVB or CtpB involved in the processing and activation of σ^K, the last factor sigma of sporulation (Higgins and Dworkin, 2012; Hilbert and Piggot, 2004).

The four sporulation-specific sigma factors are conserved in *Clostridia* and are present in *C. difficile* (Abecasis *et al.*, 2013; de Hoon *et al.*, 2010; Galperin *et al.*, 2012; Paredes *et al.*, 2005; Pereira *et al.*, 2013; Stragier, 2002). This is also the case for key genes involved in spore morphogenesis (de Hoon *et al.*, 2010; Galperin *et al.*, 2012; Stragier, 2002). A detailed study of the morphological changes during spore differentiation by *C. difficile* is described on Chapter 3 (Pereira *et al.*, 2013). The *C. difficile sigF* and *sigE* mutants are arrested just after asymmetric division while the *sigG* mutant is blocked after engulfment completion. However, and unlike in *B. subtilis*, the *sigG* mutant shows deposition of electrodense coat material around the forespore. These three mutants are unable to sporulate (Pereira *et al.*, 2013) (Chapter 3). In contrast, a *sigK* mutant of *C. difficile* still forms heat resistant spores, even though much less efficiently than the isogenic wild-type strain. Consistent with the formation of heat resistant spores, the *sigK* mutant shows accumulation of at least some cortex material, which is unlike the case for *B. subtilis* (Pereira *et al.*, 2013) (Chapter 3). However, and has also shown for *B. subtilis,* the *sigK* mutant shows no signs of coat deposition. Therefore, these studies have unraveled important deviations from the *B. subtilis* paradigm, in what concerns the functions of the sporulation sigma factors, adding to work showing that this is also the case for other *Clostridia* species (Haraldsen and Sonenshein, 2003; Harry *et al.*, 2009; Jones *et al.*, 2011; Pereira *et al.*, 2013; Stragier, 2002; Tracy *et al.*, 2011).

A global analysis of the genetic control of sporulation in a Clostridial species is also lacking. Note that although a global analysis of gene expression during *C. acetobutylicum* sporulation has been published (Jones *et al.*, 2008), the composition of the four-sigma factors regulons remains unknown. Recent work has addressed the regulatory steps during the

initiation of sporulation in *C. difficile.* Both Spo0A and σ^H are required for efficient sporulation in *C. difficile* sporulation (Saujet *et al.*, 2011; Underwood *et al.*, 2009). By contrast, the phosphorelay is absent and the sporulation initiation pathway in *C. difficile* and other *Clostridia*, may be simpler, composed by Spo0A and associated kinases (Paredes *et al.*, 2005; Steiner *et al.*, 2011; Underwood *et al.*, 2009). Both σ^H and Spo0A directly control the expression of the *spoIIAA-spoIIAB-sigF* operon while Spo0A binds *in vitro* to the promoter region of *spoIIAA*, *spoIIE* and *spoIIGA* (in an operon with *sigE*) (Rosenbusch *et al.*, 2012; Saujet *et al.*, 2011).

In this work we have combined transcriptome analysis and genome-wide transcriptional start site (TSS) mapping to define the σ^F, σ^E, σ^G and σ^K regulons. This helped us to propose candidates for new spore coat proteins. The data supports the conclusions on the main period of activity and function of the sigma factors, and unravels additional differences in the genetic control of late steps in the *C. difficile* sporulation regulatory cascade, also appear to differ relative to the *B. subtilis* paradigm (Fimlaid *et al.*, 2013; Pereira *et al.*, 2013) (Chapter 3). We show that activation of the σ^E regulon was partially independent of σ^F, even though we show that SpoIIR is the only factor involved in pro-σ^E processing, and that the σ^K regulon was not controlled by σ^G.

MATERIALS AND METHODS

Bacterial strains, growth conditions and sporulation assay. *C. difficile* strains used in this study are presented in Table 4.1. *C. difficile* strains were grown anaerobically (10 % H_2, 10 % CO_2, and 80 % N_2) in TY or in Brain Heart Infusion (BHI, Difco), which was used for conjugation. Sporulation medium (SM) (Wilson *et al.*, 1982) was used for sporulation assays. SM medium (pH at 7.4) contained per liter: 90 g Bacto tryptone, 5 g

Bacto peptone, 1 g $(NH_4)_2SO_4$, 1.5 g Tris Base. When necessary, cefoxitin (25 μg/ml), thiamphenicol (15 μg/ml) and erythromycin (2.5 μg/ml) were added to *C. difficile* cultures. *E. coli* strains were grown in Luria-Bertani (LB) broth. When indicated, ampicillin (100 μg/ml) or chloramphenicol (15 μg/ml) was added to the culture medium.

Sporulation assays were performed as follows. After 72 h of growth in SM medium, 1 ml of culture was divided into two samples. To determine the total number of cells, the first sample was serially diluted and plated on BHI with 0.1 % taurocholate (Sigma-Aldrich). Taurocholate is required for the germination of *C. difficile* spores (Sorg and Sonenshein, 2008). To determine the number of spores, the vegetative bacteria of the second sample were heat killed by incubation for 30 min at 65°C prior to plating on BHI with 0.1 % taurocholate. The percentage of sporulation was determined as the ratio of the number of spores/ml and the total number of bacteria/ml (x100).

Table 4.1. Bacterial Strains used in this study

Strains	Relevant properties	Origin
E.coli		
Dh5α		Invitrogen
TOP 10		"
BL21 (DE3)		New England Biolabs
HB101 (RP4)		Laboratory stock
C. difficile		
630Δ*erm*	630 Δ*erm*	Hussain *et al.*, 2005
CDIP3	630 Δ*erm spo0A::erm*	Saujet *et al.*, 2011
CDIP224	630 Δ*erm spoIIID::erm*	This work
CDIP227	630 Δ*erm spoVT::erm*	"
CDIP238	630 Δ*erm spoIIR::erm*	"
CDIP246	630 Δ*erm spoIIR::erm* pMTL84121-*spoIIR*	"
CDIP262	630 Δ*erm spoIIID::erm* pMTL84121-*spoIIID*	"
CDIP263	630 Δ*erm spoVT::erm* pMTL84121-*spoVT*	"
AHCD533	630 Δ*erm sigF::erm*	Chapter 3
AHCD532	630 Δ*erm sigE::erm*	"
AHCD534	630 Δ*erm sigG::erm*	"
AHCD535	630 Δ*erm sigK::erm*	"
AHCD548	630 Δ*erm sigF::erm* pMTL84121-*sigF*	"
AHCD549	630 Δ*erm sigE::erm* pMTL84121-*sigE*	"
AHCD550	630 Δ*erm sigG::erm* pMTL84121-*sigG*	"
AHCD551	630 Δ*erm sigK::erm* pMTL84121-*sigK*$^{short\ skin}$	"

Table 4.1. Bacterial Strains used in this study

Strains	Relevant properties	Origin
AHCD606	630Δ*erm* P_{spoIIR}-$SNAP^{Cd}$	This work
AHCD622	630Δ*erm spo0A::erm* P_{spoIIR}-$SNAP^{Cd}$	"
AHCD623	630Δ*erm sigF::erm* P_{spoIIR}-$SNAP^{Cd}$	"
AHCD624	630Δ*erm sigE::erm* P_{spoIIR}-$SNAP^{Cd}$	"

Construction of *C. difficile* strains. The ClosTron gene knockout system (Heap *et al.*, 2007) was used to inactivate the *spoIIID* (*CD0126*), *spoVT* (*CD3499*) and *spoIIR* (*CD3564*) genes giving strains CDIP224 (630Δ*erm CD0126*::*erm*), CDIP227 (630Δ*erm CD3499*::*erm*) and CDIP238 (630Δ*erm CD3564*::*erm*). Primers to retarget the group II intron of pMTL007 to these genes were designed by the Targetron design software (http://www.sigmaaldrich.com) (all primers used in this work are listed on Table A2). The PCR primer sets were used with the EBS universal primer and intron template DNA to generate by overlap extension PCR, a 353-bp product that would facilitate intron retargeting to *CD0126, CD3499* or *CD3564*. The PCR products were cloned between the HindIII and BsrGI sites of pMTL007 to yield pDIA6123 (pMTL007::Cdi-*CD0216*-39s), pDIA6120 (pMTL007::Cdi-*CD3499*-157a) and pMS459 (pMTL007::Cdi-*CD3564*-38a) (all plasmids used in this work are listed on Table A1). The presence and orientation of the insert in pMS459 was verified by DNA sequencing using the pMTL007-specific primers pMTL007-F and pMTL007-R. pDIA6120, pDIA6123 and pMS459 were introduced into *E. coli* HB101 (RP4) and the resulting strains subsequently mated with *C. difficile* 630Δ*erm* (Hussain *et al.*, 2005). *C. difficile* transconjugants were selected by sub-culturing on BHI agar containing Tm and Cfx and then plated on BHI agar containing Erm. Chromosomal DNA of transconjugants was isolated as previously described. PCR using the ErmRAM primers (ErmF and ErmR) confirmed that the Erm resistant phenotype was due to the splicing of the group I intron from the group II intron following integration. In order to verify the integration of the Ll.LtrB intron into the right gene targets, we performed PCR with two primers

flanking the insertion site in *CD0126* (LS184-LS185), *CD3499* (LS186-LS187) or *CD3564* (IMV649-LS113) and in one hand with the intron primer EBSu and in other hand with a primer in *CD0126* (LS184), *CD3499* (LS187) or *CD3564* (LS113).

To complement the *spoIIR* mutation, the *spoIIR* gene with its promoter (positions -169 to +791 relative to the translational start site), was amplified by PCR using oligonucleotides IMV642 and IMV641. To complement the *spoIIID* and the *spoVT* mutants, the *spoIIID* gene with its promoter (-117 to +320 from the translational start site) and the *spoVT* gene with its promoter (-165 to +731 from the translational start site), were amplified by PCR using oligonucleotides LS283 and IMV647 or IMV644 and IMV648. The PCR fragments were cloned into the XhoI and BamHI sites of pMTL84121 (Heap *et al.*, 2009) to produce plasmids pDIA6132 (*spoVT*), pDIA6133 (*spoIIID*) and pDIA6135 (*spoIIR*). Using the *E. coli* HB101 (RP4) strain containing pDIA6135 as donor, this plasmid was transferred by conjugation into the *C. difficile spoIIR* mutant giving strain CDIP246 (Table 4.1). Similarly the plasmids pDIA6132 and pDIA6133 were transferred by conjugation into the *spoVT* or *spoIIID* mutant giving strains CDIP263 and CDIP262.

To construct a P_{spoIIR}-SNAP fusion reporter strain, a 470 bp fragment encompassing the region upstream of the *spoIIR* gene was amplified by PCR using primers PspoIIR-SNAP EcoRIFw and PspoIIR-SNAP XhoIRev. The PCR product was inserted between the EcoRI and XhoI sites of pFT47 (Pereira *et al.*, 2013) (Chapter 2) to create pMS462. pMS462 was introduced into *E. coli* HB101 (RP4) and then transferred to *C. difficile* 630Δ*erm*, *spo0A*::*erm*, *sigF*::*erm* or *sigE*::*erm* strains by conjugation. Transconjugants were selected on BHI agar plates containing Tm and Cfx followed by plating on BHI agar containing Tm.

Quantitative real-time PCR. To perform transcriptional analysis for each sigma factor, we first tested the impact of their inactivation on the expression of two target genes at different times to optimize conditions. For this purpose, we used *bona fide* targets of these sigma factors in *B. subtilis* (de Hoon *et al.*, 2010): *gpr* and *spoIIR* for σ^F, *spoIIIAA* and *spoIVA* for σ^E and *sspA* and *sspB* for σ^G. For σ^K, we used *cotCD* and *cotA* encoding recently described *C. difficile* spore coat proteins (Permpoonpattana *et al.*, 2011). This preliminary test allowed us to define a time for maximal differential expression for these targets between wild-type and mutant strains for each sigma factor. To study σ^E- or σ^F-dependent control, we harvested cells of strain 630Δ*erm*, the *sigE* and the *sigF* mutants after 14 h of growth in SM medium. The strain 630Δ*erm* and the *sigG* or the *sigK* mutant were harvested after 19 h (630Δ*erm*, *sigG* mutant) and 24 h (630Δ*erm*, *sigK* mutant) of growth in SM medium. For the *spoIIR* and *spoIIID* mutants, we harvested cells after 14 h and 15 h of growth in SM medium, respectively. The culture pellets were resuspended in RNApro solution (MP Biomedicals) and RNA extracted using the FastRNA Pro Blue Kit, according to the manufacturer's instructions. The RNA quality was determined using RNA 6000 Nano Reagents (Agilent). For the *spo0A* mutant and the 630Δ*erm* strain, RNA previously obtained were used to test the impact of Spo0A inactivation on gene expression (Saujet *et al.*, 2011). Quantitative real-time PCR (qRT-PCR) analysis was performed as previously described (Saujet *et al.*, 2011). The primers used for each marker are listed in Table A2. In each sample, the quantity of cDNAs of a gene was normalized to the quantity of cDNAs of the DNApolIII gene. The relative change in gene expression was recorded as the ratio of normalized target concentrations ($\Delta\Delta^{Ct}$) (Livak and Schmittgen, 2001).

Microarray design, DNA-array hybridization and transcriptome analysis. The microarray of *C. difficile* 630 genome (GEO database

accession number GPL10556) was designed as previously described (Saujet *et al.*, 2011). Transcriptome was performed using for each condition four (630Δ*erm* compared to the mutant inactivated for each sigma factor) or two different RNA preparations (630Δ*erm* compared to *spoIIID*). Labeled DNA hybridization to microarrays and array scanning were done as previously described (Saujet *et al.*, 2011). The slides were analyzed using R and limma software (Linear Model for Microarray Data) from Bioconductor (www.bioconductor.org). We corrected background with the 'normexp' method, resulting in strictly positive values and reducing variability in the log ratios for genes with low levels of hybridization signal. Then, we normalized each slide with the 'loess' method (Smyth and Speed, 2003). To test for differential expression, we used the bayesian adjusted t-statistics and we performed a multiple testing correction of Benjamini and Hochberg based on the false discovery rate (FDR) (Benjamini, 1995). A gene was considered as differentially expressed when the p-value is < 0.05. The complete data set was deposited in the GEO database with a series record accession number GSE43202.

Purification of σ^F-His_6 and σ^E-His_6 for antibody production. The coding sequences of *sigF* and *sigE* were amplified by PCR using primer pairs CDsigF-pET28aFw/CDsigF-pET28aRev and CDsigE-pET28aFw/CDsigE-pET28aRev, respectively. The resulting fragments were introduced between the NcoI and XhoI sites of pET28a to produce pFT35 (σ^F-His_6) and pFT34 (σ^E-His_6) (Table A1). BL21(DE3) derivatives carrying pFT34 or pFT35 were grown in LB to an OD_{600nm} of 0.5 and induced with 1 mM IPTG for 4 h. The cells were then harvested by centrifugation at 4000 *g*, ressuspended in 20 mM phosphate, 1 mM phenylmethyl-sulfonyl fluoride (PMSF), 10 mM imidazole, and lysed using a French pressure cell (18000 lb/in^2). After centrifugation, the supernatant (or, for the case of σ^E-His_6, the sediment after solubilization with 8M Urea for 30 min) was loaded onto a 1 ml Histrap

column (Amersham Phamarcia Biotech). The bound proteins were eluted with a discontinuous imidazole gradient. The purified proteins were used for the production of rabbit polyclonal antibodies (Davids Biotechnologie GmbH).

Whole cell lysates and immunoblot analysis. To prepare *C. difficile* whole cell extracts, 10 ml samples were withdrawn from SM cultures at the desired times following inoculation and the cells collected by centrifugation (4000 *g*, 10 min, at 4ºC). Cells were lysed in 1 ml buffer (10 mM Tris pH 8.0, 10 mM $MgCl_2$, 0.5 mM EDTA, 0.2 mM NaCl, 10% glycerol, 1 mM PMSF) using a French pressure cell. Samples (15 μg of protein) were resolved by 12% SDS-PAGE, transferred to a nitrocellulose membrane (BioRad), and subjected to immunoblot analysis as described previously (Serrano *et al.*, 1999). Antibodies against σ^F and σ^E were used at a 1:2000 dilution. A rabbit secondary antibody conjugated to horseradish peroxidase (Sigma) was used at a 1:5000 dilution. The immunoblots were developed with enhanced chemiluminescence reagents (Amersham).

Microscopy and image analysis. 1 ml samples were withdrawn from SM cultures at the desired times following inoculation, and the cells collected by centrifugation (4000 *g*, 10 min, at 4ºC). The cells were washed with 1ml of PBS and ressuspended in 0.1 ml of PBS supplemented with the membrane dye FM4-64 (10 μg/ml) and the DNA stain DAPI (4',6-diamidino-2-phenylindole; 50 μg/ml) (Molecular Probes, Invitrogen). For SNAP staining, 1 ml samples were stained for 30 min with 50 nM SNAP-Cell TMR-Star (New England Biolabs) as described (Pereira *et al.*, 2013). Cells were washed four times by centrifugation (4000 *g*, 5 min) and ressupended in 1ml of PBS. Following washing, the cells were ressuspended in 1 ml of PBS supplemented with the membrane dye Mitotracker Green (0.5 μg/ml) (Molecular Probes, Invitrogen). Cells were mounted on 1.7% agarose coated

glass slides and imaged in a Leica DM6000B microscope as previously described (Serrano *et al.*, 2011). Phase contrast images were obtained using a Uplan F1 100x objective and fluorescence images captured with a CCD Andor Ixon^{EM} camera (Andor Technologies). Images were acquired and analyzed using the Metamorph software suite version 5.8 (Universal Imaging).

RESULTS

Overview of the transcriptome data

The *C. difficile* mutants bearing insertions in the genes coding for σ^F, σ^E, σ^G or σ^K were constructed, as described in Chapter 3 (Pereira *et al.*, 2013). The *sigF* and *sigE* mutants are blocked at stage II of sporulation. The *sigG* mutant is blocked after the completion of engulfment and the *sigK* mutant form spores lacking most of the coat material (Pereira *et al.*, 2013) (Chapter 3). Thus, the sequence of morphological events has the same overall control as observed in *B. subtilis*: σ^F and σ^E control early stages of development and are likely replaced by σ^G and σ^K later during sporulation. These studies have also shown the confinement of the activities of σ^F and σ^G to the forespore, and those of σ^E and σ^K to the mother cell (Pereira *et al.*, 2013) (Chapter 3). To investigate in more detail the regulation of sporulation in *C. difficile* and to define the regulons controlled by the four cell type-specific RNA polymerase sigma factors, we compared the expression profiles of the wild-type 630Δ*erm* strain and of the *sigF*, *sigE*, *sigG* or *sigK* mutant. We performed preliminary tests to determine conditions allowing detection of differential expression between the 630Δ*erm* strain and each sigma factor mutant (see Materials and Methods). We found 111 genes and 141 genes differentially expressed with a *P* value <0.05 between the 630Δ*erm* strain and the *sigF* mutant or the *sigE* mutant, respectively. While 92 and 19 genes were down

and upregulated in the *sigF* mutant compared to the 630Δ*erm* strain, all the 141 genes were downregulated in the *sigE* mutant. In addition, 51 and 66 genes were differentially expressed with a *P* value <0.05 between the 630Δ*erm* strain and the *sigG* or the *sigK* mutant in transcriptome, respectively. All of the presumptive *sigG*-controlled genes were downregulated in the mutant compared to the 630Δ*erm* strain. Finally, 58 and 8 genes were down and upregulated, respectively, in the *sigK* mutant relative to the 630Δ*erm* strain.

Overview of the four cell type-specific σ regulons

Genome wide determination of transcriptional star sites (TSS) for the 630Δ*erm* strain was recently performed using RNA-seq (Soutourina *et al.*, 2013). The TSS was mapped for a large set of genes found to be controlled by σ^F, σ^E, σ^G or σ^K in the transcriptome, allowing the identification of promoters likely to be recognized by the RNA polymerase when associated with each of these sigma factors (Saujet *et al.*, 2013). Consensus sequences were derived for *C. difficile* σ^F- and σ^G-controlled promoters (Figure 4.1A) and are very similar to those of *B. subtilis* (de Hoon *et al.*, 2010; Wang *et al.*, 2006) and to each other in their -10 and -35 elements, as noted for *B. subtilis* (Wang *et al.*, 2006). The *C. difficile* consensus elements specific for σ^E- and σ^K-controlled promoters are also very similar to those determined for *B. subtilis* (de Hoon *et al.*, 2010; Eichenberger *et al.*, 2004; Eichenberger *et al.*, 2003). For example, promoters recognized by σ^E and σ^K shared very similar sequences in their -10 element while the -35 elements differed, with an ATA motif for σ^E and AC for σ^K (Figure 4.1B).

We have combined transcriptome data and promoter identification to define the σ^F, σ^E, σ^G and σ^K regulons. The pool of genes likely to to be under the direct control of each sigma factor, among genes positively controlled by each of the four sigma factors, was defined and is listed in Table A3 for genes belonging to the forespore-specific σ^F and σ^G regulons,

and in Table A4 for genes belonging to the mother cell-specific σ^E and σ^K regulons.

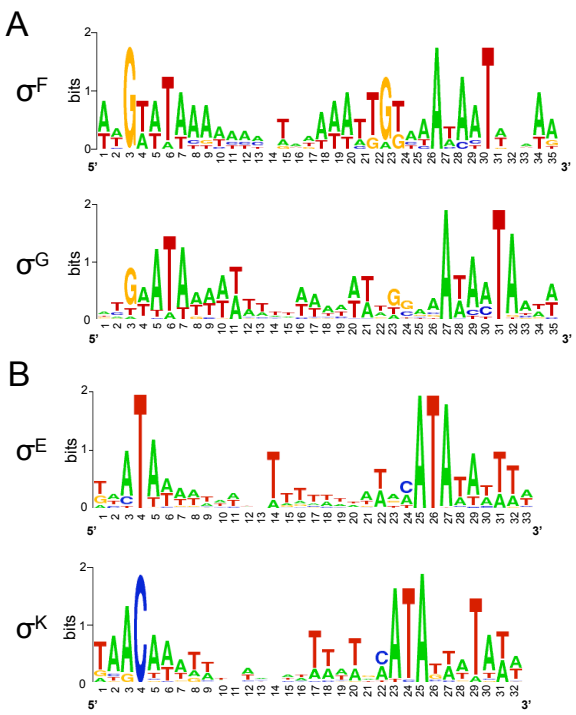


Figure 4.1. Consensus promoter sequences for σ^F- and σ^G-controlled promoters (A) or σ^E- and σ^K-controlled promoters (B) in *C. difficile*. The sequence logo was created on the WebLogo website (http://weblogo.berkeley.edu). The height of the letters is proportional to their frequency. Detailed information about the mapped promoters used to derive these consensus sequences can be found in Saujet *et al.*, 2013.

The σ^F regulon

Among the 111 genes controlled by σ^F in transcriptome, 25 were downregulated in the *sigF* mutant but not in the *sigE*, *sigG* or *sigK* mutant (Table A3). We identified 10 σ^F-dependent promoters controlling the expression of 14 genes. The σ^F regulon is the smallest of the four-compartment specific regulons as also observed in *B. subtilis* (Steil *et al.*, 2005; Wang *et al.*, 2006). Orthologs of *B. subtilis* genes involved in the regulation of sporulation, in spore morphogenesis or in germination (de Hoon *et al.*, 2010; Galperin *et al.*, 2012; Hilbert and Piggot, 2004; Stragier, 2002) were downregulated in the *sigF* mutant compared to strain 630Δ*erm*. σ^F positively controls the expression of *spoIIR*, encoding a signaling protein

that in *B. subtilis* triggers the activation of the membrane bound SpoIIGA protease that in turn cleaves the pro-σ^E protein and leads to the production of active σ^E (Hilbert and Piggot, 2004). The expression of *spoIIQ* and *spoIIP* genes, which code for proteins involved in the engulfment process, was also downregulated in the *sigF* mutant. In addition, σ^F was required to transcribe the *gpr* gene encoding a protease that is important for the degradation of small acid-soluble spore proteins (SASPs) during germination. The orthologs of these genes are members of the σ^F regulon in *B. subtilis* (de Hoon *et al.*, 2010; Steil *et al.*, 2005; Wang *et al.*, 2006). On the contrary, *gpr*, *spoIIR* and *spoIIP* are not controlled by σ^F in *C. acetobutylicum* (Jones *et al.*, 2011). σ^F was also required for the expression of genes thought to be involved in peptidoglycan metabolism: *sleB*, *CD2141* and *CD1229*. Moreover, σ^F positively controlled the expression of 5 genes encoding proteins previously detected in the spore proteome (Lawley *et al.*, 2009).

The σ^E regulon

We found 97 genes to be downregulated in a *sigE* mutant but not in a *sigG* mutant or in a *sigK* mutant (Table A4). TSS mapping allowed us to identify 47 σ^E-dependent promoters controlling the expression of 63 genes. σ^E controls the largest of the four cell type-specific sporulation regulons as also found for *B. subtilis* (Eichenberger *et al.*, 2003, Steil *et al.*, 2005). Several σ^E target genes in *C. difficile* are orthologs of genes known to control engulfment, cortex formation, initiation of spore coat assembly, the latest stages in spore maturation, and that code for proteins required for spore germination in *B. subtilis*. This is consistent with the morphological block imposed by a *sigE* insertional mutation, which arrests development just after asymmetric division (Pereira *et al.*, 2013). *spoIID,* which encodes a PG hydrolase required for the engulfment of the forespore by the mother cell was controlled by σ^E. In *B. subtilis,* SpoIID is associated to SpoIIP and SpoIIM to form the DMP machine necessary for engulfment (Eichenberger *et al.*,

2001; Morlot *et al.*, 2010). The *spoIIP* gene was only controlled by σ^F in the transcriptome analysis. However, in qRT-PCR experiments, we showed that *spoIIP* expression decreased 10-fold in a *sigE* mutant and 800-fold in a *sigF* mutant. This suggested that this gene was under the control of both sigma factors as observed in *B. subtilis* (Eichenberger *et al.*, 2003; Wang *et al.*, 2006). In addition, σ^E activated the expression of the *spoIIIAA* octacistronic operon encoding an ATPase and membrane proteins that localize to the outer forespore membrane in *B. subtilis*; together with SpoIIQ, the SpoIIIA proteins form a novel type of secretion system that allows to maintain the integrity of the forespore and its potential for macromolecular synthesis (Camp and Losick, 2008; Higgins and Dworkin; 2012; Meisner *et al.*, 2012).

In *B. subtilis*, the synthesis of the protective envelope that encases the spore, the coat and the cortex PG, is initiated under σ^E control (Eichenberger *et al.*, 2003). In *C. difficile*, the expression of *spoIVA* and *sipL* encoding an ortholog of the SpoIVA morphogenetic ATPase and a functional homolog of *B. subtilis* SpoVID was downregulated in the *sigE* mutant. These proteins are required for proper spore coat localization around the forespore in *C. difficile* (Putman *et al.*, 2013). The expression of *cotB* encoding a recently identified coat protein (Permpoonpattana *et al.*, 2011) was also specifically regulated by σ^E. Accordingly, a σ^E-dependent promoter was mapped upstream of *spoIVA, sipL* and *cotB*. Moreover, an ortholog of the *cwlD* gene, encoding a N-acetylmuramoyl-L-alanine amidase required for the synthesis of muramic *∂*-lactam, a specific cortex PG compound (Gilmore *et al.*, 2004), was controlled by σ^E. On the contrary, the second enzyme of this pathway, the PdaA deacetylase, was produced in the forespore (see also below). The synthesis of homologues for the *B. subtilis* YqfC, YqfD and YlbJ, which are involved in heat resistance and development of spore refractility in this organism (Eichenberger *et al.*, 2003; Evans *et al.*, 2003), were also controlled by σ^E in *C. difficile.* In addition, the expression of the *spmA-spmB* operon involved in

spore core dehydration and heat resistance in *B. subtilis* and *C. perfringens* (Paredes-Sabja *et al.*, 2008) was reduced in the *sigE* mutant.

Several genes under σ^E control in *B. subtilis* are required for spore germination as they are involved in modifications of the cortex or in proper assembly of the coat. In *C. difficile*, σ^E was required for the transcription of the *csp* operon encoding serine proteases. In *C. perfringens* and *C. difficile*, CspB is necessary for pro-SleC maturation to form the spore cortex lytic enzyme SleC during germination (Adams *et al.*, 2013; Cartman and Minton, 2010; Paredes-Sabja *et al.*, 2011) while CspC is a germinant receptor for taurocholate (Francis *et al.*, 2013). The promoter of the *C. difficile csp* operon matched the consensus for σ^E-dependent promoters while *sleC* was transcribed from a σ^K-dependent promoter (see below). Moreover, the *CD2833* and *CD0760* genes encoding a Ca^{2+}-transporting ATPase and a Ca^{2+}/Na^+ antiporter, respectively, were downregulated in the *sigE* mutant. These proteins are probably important for the import or export of Ca^{2+}, a key element of the sporulation and germination processes in endospore forming Firmicutes (Paredes-Sabja *et al.*, 2011).

Finally, it should be mentioned that 15 members of the σ^E regulon were detected in the spore proteome (Lawley *et al.*, 2009). This included the SpoIVA and SipL proteins, the coat protein CotB, the Csp-like proteases, the YqfC-like protein and three proteins of unknown function (CD0214, CD3522 and CD1930).

The σ^G regulon

From the transcriptome experiments, we identified 50 genes controlled by σ^G but not by σ^K (Table A3). The TSS mapping performed by RNA-seq allowed the identification of 30 σ^G-dependent promoters controlling the expression of 33 genes. Several σ^G-controlled genes encode proteins sharing similarities with proteins involved in spore resistance in *B. subtilis* (de Hoon *et al.*, 2010; Hilbert and Piggot, 2004; Setlow, 2007).

Dipicolinic acid (DPA) is a major spore component important for spore resistance. The expression of the *spoVA* operon whose products are required for the import of DPA into the forespore from the mother cell (Tovar-Rojo *et al.*, 2002), decreased in a *sigG* mutant. A *C. perfringens spoVA* mutation prevents the accumulation of DPA and reduces spore viability (Paredes-Sabja *et al.*, 2011). As observed in *B. subtilis*, we also showed the requirement of *C. difficile* σ^G for the expression of all the genes encoding SASPs; these proteins bind to the forespore chromosome protecting the DNA from damage (Setlow, 2007). Indeed, the *sspA* (*CD2688*) and *sspB* (*CD3249*) genes encoding alpha/beta-type SASP and two other genes annotated as SASPs (*CD1290* and *CD3220.1*) were expressed under the direct σ^G control. Moreover, *CD0684* encoding an ATP-dependent protease sharing similarity with FtsH, a protein involved in protein quality control and stress resistance in eubacteria, was downregulated in a *sigG* mutant. In addition, genes encoding proteins presumably involved in mitigating oxidative stress were also controlled by σ^G. These include *CD1567* (*cotG*), *CD1631* (*sodA*) and *CD2845,* encoding a catalase, a superoxide dismutase and a rubrerythrin, respectively. Interestingly, SodA and CotG have recently been detected at the *C. difficile* spore surface (Permpoonpattana *et al.*, 2013). These results suggested that proteins synthesized under the control of σ^G are located at the surface of the *C. difficile* spore. Expression of several genes required for cell wall synthesis and cortex formation is also controlled by σ^G. As an example, σ^G positively controls a gene coding for PdaA, a N-acetylmuramic acid deacetylase involved in the formation of muramic *∂*-lactam in the spore cortex PG and indirectly required for efficient spore germination (Gilmore *et al.*, 2004).

Finally, the transcriptome analysis revealed 16 σ^G-controlled genes coding for proteins of unknown function. One of the most important aspects of spore biology is the mechanism by which dormant spores sense a suitable environment for germination and trigger the process through the specific

recognition of germinants by germination receptors. These receptors are located in the spore inner membrane and are synthesized in the forespore under σ^G control in *B. subtilis* (Steil *et al.*, 2005; Wang *et al.*, 2006). No known homologs of the *B. subtilis gerA*, *gerB* and *gerK* operons as well as of the *gerA*, *gerKB*, *gerKA* and *gerKC* genes of *C. perfringens* are found in the *C. difficile* genome (Burns *et al.*, 2010a; Xiao *et al.*, 2011). Nine genes encoding probable membrane proteins were downregulated in the *sigG* mutant and these proteins may be involved in germination. 55 % of our σ^G-controlled genes encode proteins associated to the spore in proteome (Lawley *et al.*, 2009). This is in agreement with the crucial role of σ^G in the synthesis of key components protecting the spore or in preparation for germination.

The σ^K regulon

In transcriptome analysis, we found 56 genes positively controlled by σ^K (Table A4). The TSS mapping allowed us to identify 24 σ^K-dependent promoters controlling the expression of 29 genes. In *B. subtilis,* σ^K plays a crucial role in the last steps of the spore coat assembly (Henriques and Moran, 2007; Hilbert and Piggot, 2004). While the *B. subtilis* spore coat comprises over 70 different proteins, only few of them are conserved in *C. difficile* (Henriques and Moran, 2007). Recent studies have identified eight components of the *C. difficile* spore coat, CotA, CotB, CotCB, CotD, CotE, CotF, CotG and SodA (Permpoonpattana *et al.*, 2013; Permpoonpattana *et al.*, 2011). Interestingly, the expression of the *cotA* and *cotE* genes as well as the *CD0596-cotF-cotCB* and *CD2399-cotJB2-cotD* operons was downregulated in a *sigK* mutant compared to strain 630Δ*erm*. We mapped a σ^K-dependent promoter upstream of *cotA* and *cotE* while σ^K consensus elements were found upstream of *CD0596* and *CD2399*. The expression of *CD3569* encoding a YabG-like protease involved in post-translational modification of spore surface proteins in *B. subtilis* (Henriques and Moran, 2007) was also positively controlled by σ^K. Three orthologs of the *B. anthracis bclA* gene

(*bclA1*, *bclA2* and *bclA3*) are present in *C. difficile*. BclA of *B. anthracis* is a collagen-like protein that forms the hair-like nap of the exosporium, and is the immunodominant antigen of the spore surface (Steichen *et al.*, 2007). The *bclA1*, *bclA2* and *bclA3* genes were positively controlled by σ^K and sequences similar to promoters recognized by σ^K are found upstream of these genes.

An important conserved function of the mother cell is DPA production, which is controlled by σ^K in *B. subtilis* (Steil *et al.*, 2005). The expression of *dpaA* and *dpaB* encoding the dipicolinate synthase is also controlled by σ^K in *C. difficile*. We also found that *sleC* encoding the major spore cortex lytic enzyme that is essential for *C. difficile* germination (Burns *et al.*, 2010b) was a member of the σ^K regulon. Mother cell lysis is a late developmental event that in *B. subtilis* involves the σ^K-controlled production of the CwlC and CwlH N-acetylmuramic acid L-alanine amidases (Eichenberger *et al.*, 2004; Nugroho *et al.*, 1999). Two genes encoding N-acetylmuramic acid L-alanine amidases (*CD1898* and *CD2184*) were downregulated in a *sigK* mutant and are possibly involved in mother cell lysis. This is in line with previous findings (Chapter 3) where we observed no mother cell lysis in the *sigK* sporulating cells carrying phase bright spores. (Chapter 3, Figure 3.5).

Other genes controlled by σ^K have no counterparts in *B. subtilis.* Finally, the transcriptome analysis showed that genes encoding 17 proteins of unknown function were downregulated in a *sigK* mutant. Interestingly, nine of these proteins are associated to the spore (Lawley *et al.*, 2009) and we mapped a σ^K-dependent promoter upstream five of their encoding genes (*CD1063.1, CD1067, CD1133*, *CD3580* and *CD3613*). These proteins are most likely spore coat proteins but further work would be necessary to characterize their localization and their function in *C. difficile.* In conclusion, 21 out of the 57 genes controlled by σ^K were found to encode components of the spore proteome (Lawley *et al.*, 2009). The overall composition of the σ^K

regulon is in agreement with the crucial role of σ^K in the assembly of the spore surface layers, spore maturation and mother cell lysis in *C. difficile* as well as in other endospore forming Firmicutes (Henriques and Moran, 2007; Pereira *et al.*, 2013).

The forespore line of gene expression

Our global approaches allowed us to obtain new insights into the forespore line of gene expression, which is governed by the RNA polymerase sigma factors, σ^F and σ^G, and the DNA-binding transcriptional regulator, SpoVT.

Control of sigF *and* sigG *expression*

Like in *B. subtilis* and other spore formers, the *spoIIGA*, *sigE* and *sigG* genes are clustered in *C. difficile* (de Hoon *et al.*, 2010; Stragier, 2002). The *B. subtilis sigG* gene is transcribed from two promoters, one located upstream of *spoIIGA* (σ^A) and the second located upstream of *sigG* transcribed first by σ^F and later by σ^G after engulfment (Hilbert and Piggot, 2004). In both *B. subtilis* and *C. acetobutylicum*, three transcripts are detected for the *spoIIGA-sigE-sigG* locus, corresponding to a *spoIIGA-sigE-sigG*, a *spoIIGA-sigE* and a *sigG* transcript (Harris *et al.*, 2002; Hilbert and Piggot, 2004; Santangelo *et al.*, 1998). In *C. difficile*, a σ^A-dependent promoter (GTGACA and TATAAT boxes) was mapped by RNA-seq upstream of *spoIIGA* (TSS at position 3052270 in the genome of strain 630). A promoter was also mapped upstream of *sigG* both by RNA-seq and by 5'-RACE (Saujet *et al.*, 2011). The associated consensus motifs clearly correspond to those found for the forespore sigma factors (Figure 4.1A) but appear closer to that found for σ^F promoters as the -10 element contains a G at position 23 and lacks the A conserved at position 32 for σ^G-dependent promoters (Figure 4.1A). However, *sigG* expression was not downregulated in the *sigF* mutant in the transcriptome analysis as well as by qRT-PCR,

under our conditions. In *B. subtilis*, *sigG* expression also appears not to be regulated in transcriptome analyses comparing *sigF*$^{+}$ and *sigF*$^{-}$ strains (Steil *et al.*, 2005; Wang *et al.*, 2006). By contrast, *sigG* expression is downregulated in a *sigF* mutant in *C. acetobutylicum* (Jones *et al.*, 2011) suggesting some differences in the control of *sigG* expression. In *C. difficile,* the absence of control of *sigG* expression by σ^F strongly suggests that *sigG* is transcribed from at least two promoters, one in front of *spoIIGA* recognized by σ^A, and one just upstream of *sigG* recognized by σ^F and/or by σ^G. Studies with a P*sigG-SNAP*^{Cd} transcriptional fusion indicate that this second promoter responsible for the forespore-specific transcription of *sigG* is dependent on σ^F (Chapter 3). A more complete analysis of the fine-tuning of *spoIIGA-sigE-sigG* locus expression will require further work.

It is worth noting that the expression of many *sigG*-controlled genes (82 % of σ^G regulon members) including those for key sporulation proteins like SpoVA, SspA, SspB and PdaA was strongly reduced in a *sigF* mutant in transcriptome (Table A3). The absence of detection of the complete σ^G regulon in the *sigF* mutant might be due to the timing of sampling (14 h) that is probably not optimal for *sigG* targets. Moreover, inactivation of *sigF* also eliminated fluorescence from a P_{sspA}-*SNAP*^{Cd} fusion (Chapter 3). Thus, σ^F directly or indirectly controls expression of the σ^G regulon.

The role of SpoVT in sporulation

In *B. subtilis*, two transcriptional regulators participate in the forespore regulatory network, RfsA and SpoVT. SpoVT controls the synthesis of about half of the members of the σ^G regulon (Wang *et al.*, 2006). RfsA is absent from the genome of *C. difficile* and several *Clostridia* while an ortholog of SpoVT (55 % identity with SpoVT of *B. subtilis*) is present. Interestingly, we found that *spoVT* expression was controlled by both σ^F and σ^G in transcriptome (Table A3). We confirmed by qRT-PCR a 280- and a 14-fold decrease of *spoVT* transcription in a *sigF* or in a *sigG* mutant compared to

strain 630Δ*erm*, respectively (Table A5). We also showed that the expression of *spoVT* was restored to the wild-type level in the *sigG* mutant complemented by *sigG* while its expression was 19-fold higher in the *sigF* mutant complemented by *sigF* compared to the wild-type strain (Table A6). Upstream of the *spoVT* TSS, a promoter resembling both the σ^F and σ^G consensus elements with the presence of a G at position 23 and an A at position 32 was identified. So, *spoVT* might be a direct target of both σ^F and σ^G because contrary to other σ^G-controlled genes, σ^F inactivation caused a much more important decrease of *spoVT* expression than σ^G inactivation.

To test the possible role of SpoVT in sporulation in *C. difficile*, we constructed a *spoVT* mutant (Figure 4.2). A complemented strain, CDIP263 carrying a multicopy allele of *spoVT* under the control of its native promoter was also obtained. To determine the impact of SpoVT inactivation on spore morphogenesis, samples of the strain 630Δ*erm*, of the *spoVT* mutant and of the complemented strain CDIP263 (*spoVT*::*erm*, pMTL84121-*spoVT*) were collected and labeled with the DNA stain DAPI and the lipophilic membrane dye FM4-64. Phase contrast and fluorescence microscopy experiments showed that the *spoVT* mutant was able to complete the engulfment process (Figure 4.3A). However, this mutant formed phase dark immature spores, which were not released from the sporangial cells, clearly suggesting that the cortex is absent or highly reduced. The wild-type phenotype was restored in the complemented strain CDIP263. We further tested the ability of this mutant to form spores. After 72 h of growth in SM medium, the 630Δ*erm* strain or the *spoVT* mutant containing the plasmid pMTL84121-*spoVT* produced $4x10^6$ and 10^7 heat-resistant spores/ml, respectively. In contrast, no clones were obtained after a heat treatment for the *spoVT* mutant. This was probably due to the lack of production of heat-resistant spores as suggested by the phase dark spores phenotype observed for this mutant (Figure 4.3A). In conclusion, SpoVT is required for mature spore formation in *C. difficile* but the phenotype of the *spoVT* mutant of *C. difficile* differs from

that of the same mutant of *B. subtilis*. Indeed, the *spoVT* mutant of *C. difficile* formed phase dark spores instead of phase bright spores (Bagyan *et al.*, 1996). Also, a *spoVT* mutation only reduces the ability of *B. subtilis* to form spores (Bagyan *et al.*, 1996) while in *C. difficile* it completely prevented spore formation.

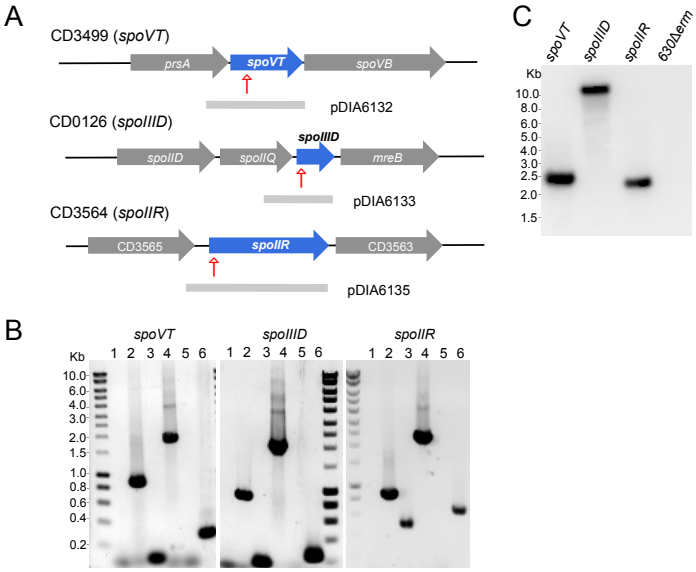


Figure 4.2. **Inactivation of the *spoIIR, spoIIID and spoVT* genes in *C. difficile* using the ClosTron system**. **(A)** Genetic organisation of the *C. difficile* chromosome in the vicinity of *spoVT*, *spoIIID and spoIIR*. The red arrow indicates the point of insertion of the re-targeted type II introns used for gene disruption. The extent of the DNA fragment present in the indicated replicative plasmids used for *in trans* complementation of the insertional mutations is shown below each of the genetic maps**. (B)** Chromosomal DNA of Em^R *C. difficile* conjugants and of strain 630Δ*erm* were screened by PCR using primer pairs RAM-F/R (lanes 1 and 2), using primers flanking the insertion site (lanes 3 and 4) and using the intron primer EBSu in one hand and a primer in *spoVT*, *spoIIID* or *spoIIR* on other hand (lanes 5 and 6). Chromosomal DNA from the 630Δ*erm* strain corresponded to lanes 1, 3 and 5 while chromosomal DNA of each mutant corresponded to lanes 2, 4 and 6. The smart ladder (Eurogentec) was used as a molecular weight marker. **(C)** Southern blot of HindIII-digested genomic DNA from *C. difficile* 630Δ*erm* (wild type, wt), *spoVT, spoIIID and spoIIR* mutant strains. The probe used corresponds to part of the intron sequence. The position of DNA size markers is indicate on the left side of the panel.

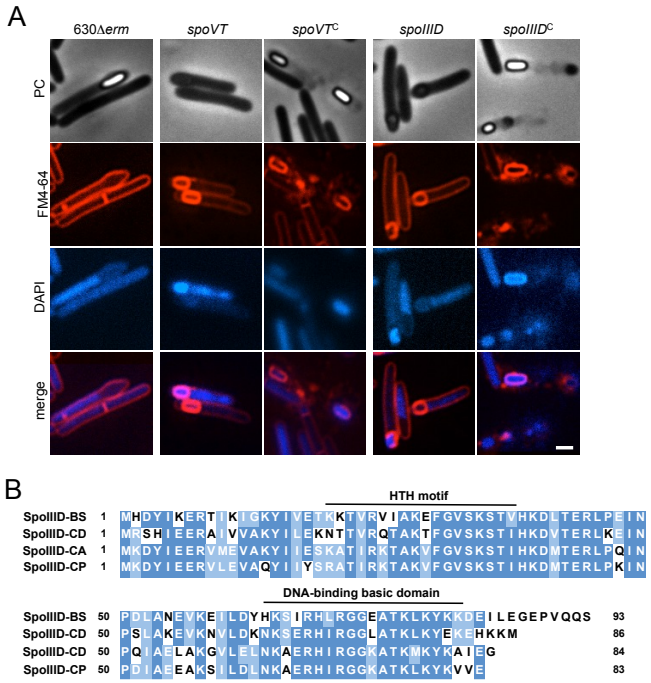


Figure 4.3. Morphological characterization of the *spoVT* and *spoIIID* mutants. (A) Cells of the *C. difficile* 630Δ*erm* strain, the *spoVT* or *spoIIID* mutant and the complemented strains, CDIP263, carrying a multicopy *spoVT* gene controlled by its native promoter or CDIP262, carrying a multicopy *spoIIID* gene controlled by its native promoter were collected after 24 h of growth in SM broth, stained with the DNA stain DAPI and the membrane dye FM4-64 and examined by phase contrast and fluorescence microscopy. Scale bar, 1 µm. **(B)** Alignment of the SpoIIID regulator. The amino acid sequences from SpoIIID of *B. subtilis* (BS), *C. difficile* (CD), *C. acetobutylicum* (CA) and *C. perfringens* (CP) were aligned using ClustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2/) and shading was created using BoxShade (http://sourceforge.net/projects/boxshade/). The two regions essential for DNA binding are indicated: an helix-turn-helix motif (HTH) and a basic region near the C-terminus of the protein (Himes *et al.*, 2010).

Then, we used qRT-PCR to measure the impact of SpoVT inactivation on the expression of selected σ^G targets after 20 h of growth. We showed that the expression of *sspA* and *sspB* decreased 70- and 12-fold, respectively in the *spoVT* mutant compared to strain 630Δ*erm* (Table 4.2). A positive effect of SpoVT on *sspA* and *sspB* expression is also observed in *B. subtilis* (Bagyan *et al.*, 1996; Ramirez-Peralta *et al.*, 2012). The decreased expression of the *sspA* and *sspB* genes is probably not sufficient to explain the phenotypes observed for the *spoVT* mutant but we did not see a reproducible differential

expression for other tested genes in our conditions. However, the expression of *spoIIR* and *gpr*, two members of the σ^F regulon, was upregulated 4- and 9-fold in a *spoVT* mutant compared to strain 630Δ*erm* earlier during sporulation (15 h) (Table 4.2). This negative effect might be due either to a direct control by SpoVT playing the role of RfsA in the negative control of σ^F targets or to an indirect effect associated with the blocked sporulation in the *spoVT* mutant.

Table 4.2. Control of expression of sporulation genes by SpoIIID and SpoVT, measured by qRT-PCR

Gene		*spoVT*/630Δ*erm*	
		15 h	**24h**
CD2470	*gpr*	9	
CD3564	*spoIIR*	4	
CD2688	*sspA*		0.014
CD3249	*sspB*		0.08
Gene		***spoIIID*/630Δ*erm***	
		15 h	**24h**
CD1230	*sigK*	0.04	$4x10^{-4}$
CD1231		NR	NR
CD1613	*cotA*	NR	$3x10^{-4}$
CD0598	*cotCB*		$3x10^{-4}$
CD1433	*cotE*		$6x10^{-4}$
CD2401	*cotD*		$8x10^{-4}$
CD0551	*sleC*		$4x10^{-4}$
CD0332	*bclA1*		0.01
CD3230	*bclA2*		0.03
CD3349	*bclA3*	0.07	$5x10^{-4}$
CD3580			0.03
CD1067		0.0003	$3x10^{-4}$
CD1133			0.03

No other genes encoding transcriptional regulators or proteins with DNA binding motifs were regulated by σ^F or σ^G in our transcriptome study (Table A3). Even though we cannot exclude that other regulatory proteins are at play, SpoVT is probably the key, if not the only, ancillary regulator of forespore gene expression. Further work will be required to analyze more precisely the morphology of the *spoVT* mutant and to understand at the

molecular level the complex role of SpoVT in the regulatory network that controls sporulation.

The mother cell line of gene expression

As discussed above, *sigE* transcription is probably initiated at the σ^A-dependent promoter located upstream of *spoIIGA,* which forms an operon with *sigE* and maybe also with *sigG*. In *B. subtilis,* σ^E positively controls the expression of *spoIIID* and *gerR* encoding regulatory proteins. GerR acts as a repressor of certain early σ^E-controlled genes and as an activator of some σ^K-dependent genes (Cangiano *et al.*, 2010). SpoIIID is an ambivalent regulator, acting as a repressor as well as an activator of a late class of σ^E-controlled genes (Eichenberger *et al.*, 2004). SpoIIID and σ^E jointly activate the expression of *sigK,* and of proteins required for the processing and activation of σ^K, and excision of the *skin* element. σ^K in turn drives production of another ancillary regulator, GerE (Eichenberger *et al.*, 2004). In *C. difficile*, SpoIIID is present while GerR and GerE are absent. The SpoIIID protein of *C. difficile* shares 64 % identity with SpoIIID from *B. subtilis* with conservation of the HTH motif and of the basic DNA binding motif located near the C-terminal of the protein (Figure 4.3B) (Himes *et al.*, 2010).

Characterization of the SpoIIID regulon

The expression of *spoIIID* was strongly curtailed in a *sigE* mutant (25- and 800-fold in transcriptome and qRT-PCR) whereas the *spoIIID* expression was restored when a plasmid pMTL84121 containing the *sigE* gene with its promoter was introduced into the *sigE* mutant (Table A6). A TSS corresponding to a σ^E-controlled promoter (ATA-N_{16}-CATATATA) was mapped upstream of *spoIIID* as observed in *B. subtilis*. It is interesting to note that in *C. perfringens* the expression of *spoIIID* is not controlled by σ^E (Harry *et al.*, 2009).

In *B. subtilis,* SpoIIID positively or negatively regulates almost half of the σ^E target genes (Eichenberger *et al.*, 2004). To test the possible role of SpoIIID in the mother cell regulatory network in *C. difficile*, we constructed a *spoIIID* mutant (Figure 4.2) and compared the expression profiles of the 630Δ*erm* strain and of the *spoIIID* mutant after 15h of growth in SM medium. We found 96 genes differentially expressed with a p value <0.05 between the 630Δ*erm* strain and the *spoIIID* mutant (Table A7). 12 and 84 genes were down and upregulated in the *spoIIID* mutant, respectively. We then performed qRT-PCR on a subset of genes regulated by SpoIIID in our transcriptome. The qRT-PCR results confirmed the microarrays data for the tested genes (Table A5). First, 47 genes that are not under the control of sporulation sigma factors in our transcriptome analyses were regulated by SpoIIID either positivelly (10 genes) or negativelly (37 genes) (Table A7). Most of these genes encode proteins involved in metabolism. Among the genes under the negative control of *spoIIID*, 28 were *bona fide* members of the σ^E-regulon (Table A4). Therefore, a principal function of SpoIIID is to inhibit the transcription of about 30% of the genes transcribed by σ^E as also observed in *B. subtilis* (Eichenberger *et al.*, 2004). SpoIIID repressed the expression of the *spoIIIA* operon and of the *spoIVA*, *sipL, cotB* and *spm* genes. In *B. subtilis, spoIIIAA* and *spoIVA* are direct targets of SpoIIID (Eichenberger *et al.*, 2004). By contrast, the *spoIID* gene, which is a direct SpoIIID target in *B. subtilis,* was not regulated by SpoIIID in transcriptome and qRT-PCR experiments in our conditions. In addition, the expression of genes encoding a transmembrane signaling protein (CD2445), an iron hydrogenase (CD3258), a bacterioferritin (CD2865), three polysaccharide deacetylases (CD3257, CD3248 and CD1319) and of several genes of unknown function was positively controlled by σ^E and negatively regulated by SpoIIID (Tables A4 and A7). By using the consensus sequence recognized by SpoIIID in *B. subtilis* (Eichenberger *et al.*, 2004), we identified a putative SpoIIID binding motif upstream of 11 of these genes (Table A4) suggesting that they might be

direct SpoIIID targets in *C. difficile*. However, the characterization of direct and indirect SpoIIID targets and the experimental identification of the DNA binding motif of SpoIIID will require further investigations.

Interestingly, SpoIIID was also an activator of the expression of two genes belonging to the σ^K regulon, *bclA3* and *CD1067* in our transcriptome experiment (Tables A4 and A7). To determine if *sigK* itself and other members of the σ^K regulon were under SpoIIID control, we performed qRT-PCR experiments (Table 4.2) using RNA extracted after 15 h of growth or after 24 h, a time where the σ^K targets are more highly expressed. We found that *sigK* expression was 25-fold and 250-fold downregulated in the *spoIIID* mutant after 15 h and 24 h of growth, respectively. This clearly indicated that the expression of *sigK* was positively controlled by SpoIIID in *C. difficile* as observed in *B. subtilis* (Eichenberger *et al.*, 2004). In *B. subtilis*, SpoIIID is also required for the expression of *spoIVCA* encoding the site-specific recombinase involved in the excision of the *skin* element that creates a functional *sigK* gene. In contrast, we did not observe a control of *CD1231* encoding the specific recombinase of the *skin* element by SpoIIID but also by σ^E both in transcriptome and qRT-PCR experiments. This result strongly suggests a different mechanism of control for the *skin* excision in *C. difficile* compared to *B. subtilis*. We also showed that eleven *bona fide* σ^K targets such as *cotA*, *cotCB*, *cotD*, *cotE*, *sleC*, *bclA1*, *bclA2*, *bclA3*, *CD3580*, *CD1067* and *CD1133* were downregulated 30- to 1000-fold in the *spoIIID* mutant as compared to the parental strain after 24 h of growth (Table 4.2). The positive regulation of members of the σ^K regulon by SpoIIID might be due either to a direct binding of this regulator to their promoter regions or to an indirect effect mediated through the control of *sigK* transcription by SpoIIID.

To investigate the role of SpoIIID in sporulation, we examined the morphology of the *spoIIID* mutant. Samples of the 630Δ*erm* strain, of the *spoIIID* mutant and of the complemented strain CDIP262 carrying a multicopy allele of *spoIIID* under the control of its native promoter

(*spoIIID*::*erm*, pMTL84121-*spoIIID*) were collected and labeled with the DNA stain DAPI and the lipophilic membrane dye FM4-64. Phase contrast and fluorescence microscopy experiments showed that the *spoIIID* mutant completed engulfment of the forespore by the mother cell and formed partially refractile spores with irregular shapes and positioning (Figure 4.3A). The wild-type phenotype with the production of free spores was restored in the strain CDIP262 containing a copy of *spoIIID* on a plasmid (Figure 4.3A). We also tested the ability of the *spoIIID* mutant to sporulate. After 72 h of growth in SM medium, the 630Δ*erm* strain produced $9x10^5$ heat-resistant spores/ml while a titer of $3x10^2$ heat-resistant spores/ml was obtained for the *spoIIID* mutant. When we complemented the *spoIIID* mutant with a pMTL84121-*spoIIID* plasmid, the titer of heat resistant spores increased to $5x10^5$/ml. Interestingly, both the morphology and the sporulation efficiency of the *spoIIID* mutant is reminiscent of the oligosporogenous phenotype obtained for the *sigK* mutant of *C. difficile* (Pereira *et al.*, 2013) in agreement with the dependency on SpoIIID of the transcription of *sigK*. The morphology and the sporulation efficiency of the *spoIIID* mutant of *C. difficile* are similar to those of the *B. subtilis* mutant (Errington, 1993; Yoshisue *et al.*, 1995). As demonstrated in *B. subtilis* (Eichenberger *et al.*, 2004), SpoIIID in *C. difficile* plays a pivotal role in the mother cell line of gene expression switching off the transcription of many members of the σ^E regulon and switching on the expression of *sigK* and of members of the σ^K regulon.

Control of sigK *transcription*

In addition to the positive control of *sigK* expression by SpoIIID, its expression also decreased 8- and 600-fold in a *sigE* mutant, in transcriptome and qRT-PCR experiments, respectively. The introduction of a plasmid pMTL84121 containing the *sigE* gene with its promoter into the *sigE* mutant restored *sigK* expression (Table A6). Interestingly, two TSS, 67 nt (P1) and

26 nt (P2) were mapped upstream of the translational start site of *sigK*. A canonical -10 box (CATATTAT) for mother cell sigma factors is located upstream of the TSS corresponding to P1 while either a TTA sequence 15 bp upstream from the -10 box or a TTT motif with a more classical 16-18 bp spacing between the -10 and -35 elements could be proposed for a -35 motif. So, this promoter resembles a σ^E-dependent promoter (Figure 1B). Upstream of the second TSS (P2), we found a consensus for σ^K-dependent promoters with a -10 box (CATATAAT) and a -35 box (AC). We note that in *B. subtilis,* following the initial transcription of *sigK* under the command of σ^E and SpoIIID, an autoregulatory loop is established, which is responsible for about 60 % of *sigK* transcription (Oke and Losick, 1993). In *C. difficile*, *sigK* is likely first transcribed by σ^E-associated RNA polymerase at P1 and then by σ^K-associated RNA polymerase at P2 probably later during sporulation. In any event, lending support to the idea that σ^E has a crucial role in *sigK* transcription, 65 % of the σ^K-controlled genes were positively controlled by σ^E including genes encoding key components of the spore surface layers (*cotCB*, *cotA*, *cotE*, *cotD*, *sleC*, *bclA1*, *bclA2*, *bclA3*). Most of these genes were much more downregulated in the *sigK* mutant than in the *sigE* mutant (Table A4) suggesting that the timing of sampling of the *sigE* mutant (14 h) is probably not optimal for σ^K targets. This might explain the absence of detection of the complete σ^K regulon in the *sigE* mutant. Importantly, with the exception of *nrdR*, we did not identify other genes for putative transcriptional regulators that were downregulated in *sigE* or *sigK* mutant (Table A4). Overall, our data indicates that in *C. difficile*, the mother cell line of gene expression is deployed according to a hierarchical regulatory cascade simpler than in *B. subtilis* (Eichenberger *et al.*, 2004).

The regulation of mother cell sigma factors synthesis in *C. difficile* also differs from other *Clostridia*. In *C. perfringens,* a biphasic pattern of *sigK* expression (early and late in sporulation) is observed and σ^E and σ^K co-regulate the expression of each other (Harry *et al.*, 2009). In *C. botulinum*,

sigK is expressed at the onset of stationary phase and σ^K positively controls the expression of *sigF* and *spo0A* (Kirk *et al.*, 2012). In *C. difficile*, we did not observe a control of *spo0A*, *sigF* or *bona fide* members of the σ^E regulon by σ^K (Table A4). In both *C. perfringens* and *C. botulinum*, sporulation is arrested at an early stage in a *sigK* mutant and not at a late stage as observed in *B. subtilis* and *C. difficile* (Hilbert and Piggot, 2004; Pereira *et al.*, 2013). It is also worth noting that the *sigK* gene is disrupted by a *skin* element in *B. subtilis* and *C. difficile* but not in *C. acetobutylicum*, *C. perfringens* and *C. botulinum*, suggesting that excision of this element may be an important factor in controlling the timing of σ^K activity in *C. difficile*.

Communication between the forespore and the mother cell

A hallmark of sporulation in *B. subtilis* is the existence of cell-cell signaling pathways that link the forespore and mother cell-specific lines of gene expression. Because these pathways operate at critical morphological stages of sporulation, the result is the coordinated deployment of the two lines of gene expression, in close register with the course of morphogenesis (Higgins and Dworkin, 2012; Hilbert and Piggot, 2004). Indeed, σ^F is required for the activation of pro-σ^E into σ^E in the mother cell and σ^E in turn is necessary to activate σ^G in the forespore. Finally, σ^G is required for the activation of pro-σ^K into σ^K in the mother cell (Higgins and Dworkin, 2012; Hilbert and Piggot, 2004). We obtained new information on the inter-compartment signaling pathway in *C. difficile*.

Absence of a strict control of the σ^E regulon by σ^F

In *B. subtilis*, the σ^E regulon is indirectly controlled by σ^F, which is required for the proteolytical activation of pro-σ^E. Surprisingly, we did not find any global control of the σ^E regulon by σ^F in *C. difficile*. We confirmed by qRT-PCR that the expression of *bona fide* σ^E targets such as *spoIIIAA*, *spoIVA*, *spoIIID* or *CD2864* did not significantly decrease in a *sigF* mutant (Figure

4.4A). This is a major difference relative to the *B. subtilis* sporulation regulatory network (Fawcett *et al.*, 2000) and indicates that σ^F is not strictly required for σ^E functionality in *C. difficile*. One possibility is that the SpoIIGA mediated proteolytical activation of pro-σ^E is not fully dependent on σ^F. To test this hypothesis, we detected in several strains (630Δ*erm*, *spo0A*, *sigF* and *sigE* strains) the σ^F, σ^E and pro-σ^E polypeptides by Western-blotting using antibodies raised against either σ^F or σ^E (Figure 4.4B). In strain 630Δ*erm*, we observed σ^F production while two forms corresponding to σ^E and pro-σ^E were detected with an anti-σ^E antibody. In a *spo0A* mutant, neither σ^E, pro-σ^E or σ^F were detected (Figure 4.4B), in agreement with the strong decrease of expression of the *spoIIAA* operon in a *spo0A* mutant (Saujet *et al.*, 2011). As shown by qRT-PCR, *spo0A* inactivation also led to a 240-fold decrease of *sigE* expression indicating that this regulator controls pro-σ^E synthesis. It has also been shown that Spo0A directly binds with low affinity to the *spoIIGA* and *spoIIAA* promoter regions (Rosenbusch *et al.*, 2012). In a *sigE* mutant, we did not detect σ^E or pro-σ^E with an anti-σ^E antibody while we specifically detected σ^F with an anti-σ^F antibody. Interestingly, in a *sigF* mutant in which σ^F was absent, we detected pro-σ^E but also the processed σ^E form in a reduced quantity compared to the situation in a wild-type strain (Figure 4.4B). The expression of the σ^E regulon in the *sigF* mutant shows that even reduced the level of active σ^E in this mutant is sufficient to allow the transcription of σ^E-controlled genes. This clearly demonstrates that an active σ^E protein can be produced in the absence of σ^F in *C. difficile* contrary to *B. subtilis* (Jonas and Haldenwang, 1989). In *C. perfringens* and *C. acetobutylicum*, neither pro-σ^E nor σ^E protein are produced in a *sigF* mutant (Jones *et al.*, 2011; Li and McClane, 2010) suggesting a diversity in the mode of forespore control of σ^E activity among spore forming Firmicutes.

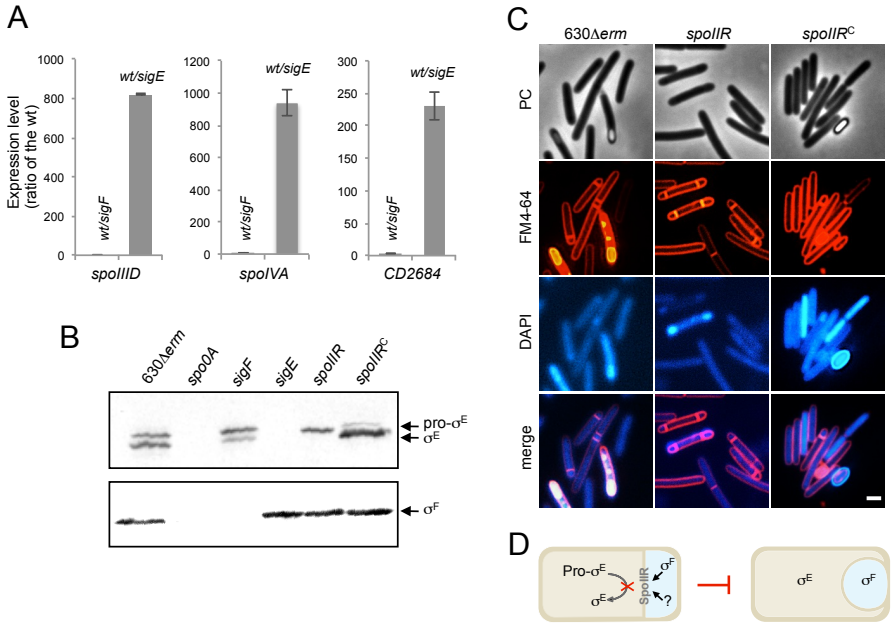


Figure 4.4. Control of σ^E targets by σ^F and σ^E and the role of SpoIIR. (A) Quantification of expression of the indicated genes by qRT-PCR. Total RNAs were extracted from *C. difficile* 630Δ*erm*, *sigF* and *sigE* mutants strains grown in SM medium for 14 hours. The expression ratio of strain 630Δ*erm*/*sigE* and 630Δ*erm*/*sigF* were indicated in white and grey, respectively. Error bars correspond to standard deviation from at least two biological replicates. **(B)** Western blot analysis showing the involvement of σ^F and SpoIIR in the control of formation of the processed σ^E protein. 15 µg of total protein extracted from 630Δ*erm* strain, the *spo0A*, *sigF*, *sigE* and *spoIIR* mutants and a *spoIIR* complemented strain were resolved by SDS-PAGE. The various samples were analysed by immunoblotting with polyclonal anti-σ^E and anti-σ^F antibodies (Materials and Methods). **(C)** Morphological characterization of a *spoIIR* mutant. Cells of the *C. difficile* 630Δ*erm* strain, the *spoIIR* mutant and the complemented strain (*spoIIR*^{C}) were collected after 24 h of growth in SM broth, stained with the DNA stain DAPI and the membrane dye FM4-64 and examined by phase contrast and fluorescence microscopy. Scale bar, 1 µm. **(D)** Schematic representation of the *spoIIR* mutant morphology. In the absence of SpoIIR, processing of pro-σ^E into σ^E is impaired (represented by a red cross). Because σ^E activation is required to start the engulfment process, the *spoIIR* mutant is blocked at the asymmetric division stage.

The role of SpoIIR in the regulatory cascade

In *B. subtilis*, σ^F drives production of the signaling protein SpoIIR, which is secreted across the forespore inner membrane into the intermembrane space, where it stimulates the SpoIIGA-dependent pro-σ^E processing in the mother cell (Higgins and Dworkin, 2012, Karow *et al.*,

1995, Londono-Vallejo and Stragier, 1995). The results exposed above raised the possibility that SpoIIR is dispensable for pro-σ^E processing in *C. difficile*, or that expression of the *spoIIR* gene is partially independent on σ^F. To determine the role of SpoIIR in *C. difficile*, we inactivated *spoIIR* (Figure 4.2). A complemented strain, CDIP246, carrying a multicopy allele of *spoIIR* under the control of its native promoter was also constructed. We first examined the morphology of the *spoIIR* mutant. Phase contrast and fluorescence microscopy experiments showed that the *spoIIR* mutant was blocked at the asymmetric septation stage and accumulated disporic cells (Figure 4.4C) as shown for the *spoIIR* mutant in *B. subtilis* (Londono-Vallejo and Stragier, 1995). The wild-type phenotype was restored in the complemented strain CDIP246. The phenotype caused by the *spoIIR* mutation in *C. difficile* phenocopied that imposed by a *sigE* mutation (Pereira *et al.*, 2013) strongly suggesting that σ^E is inactive in this mutant. In *B. subtilis*, loss of σ^E or interference with its activation leads to disporic forms (Eichenberger *et al.*, 2001; Eldar *et al.*, 2009). We also measured the sporulation efficiency of these three strains after a heat treatment. After 72 h of growth in SM medium, the strain 630Δ*erm* produced $2x10^6$ heat-resistant spores/ml. In contrast, no heat resistant spores were detected for the *spoIIR* mutant. When we complemented the *spoIIR* mutant using a pMTL84121-*spoIIR* plasmid, the titer of heat resistant spores was of $9x10^5$ spores/ml. So, inactivation of the *spoIIR* gene resulted in a complete inability of *C. difficile* to sporulate.

To test the prediction that σ^E is inactive in the *spoIIR* mutant, we compared the accumulation of σ^F, σ^E and pro-σ^E polypeptides in strain 630Δ*erm* and in the *spoIIR* mutant by immunoblotting. σ^F and pro-σ^E but not σ^E accumulated in the *spoIIR* mutant (Figure 4.4B). To independently test the impact of the *spoIIR* mutation on the activity of σ^E, the expression of selected σ^E target genes was analyzed by qRT-PCR. The expression of *spoIIIAA*, *spoIVA*, *spoIIID* and *sigK* genes decreased 10-, 8-, 16- and 23-fold in a *spoIIR* mutant compared to strain 630Δ*erm,* while the expression of *sigE* itself was

not reduced in this mutant. When the pMTL84121-*spoIIR* plasmid was introduced into the *spoIIR* mutant, the expression of *spoIIIAA*, *spoIVA*, and *spoIIID* increased 10-, 12- and 7-fold compared to strain 630Δ*erm* while the expression of the *sigK* gene was restored to the wild-type level. Therefore, both the immunoblot and the qRT-PCR studies support the idea that *spoIIR*, but not σ^F, is strictly required for the activation of σ^E.

It follows that the expression of *spoIIR* has to occur in part independently of σ^F. We first showed that the expression of *spoIIR* decreased 2.5- and 38-fold in a *sigF* mutant in transcriptome and by qRT-PCR (Table A3 and Table A5) while the introduction of pMTL84121-*sigF* into the *sigF* mutant increased *spoIIR* expression 9-fold compared to strain 630Δ*erm* (Table A6). So, the *spoIIR* transcription is under σ^F control. We also constructed a transcriptional fusion between the *spoIIR* promoter and the SNAP-tag (Pereira *et al.*, 2013; Chapter 3). This P_{spoIIR}-*SNAP*^{Cd} fusion was introduced by conjugation into strain 630Δ*erm* and a *spo0A*, *sigF* or *sigE* mutant. Labeling with the fluorescent substrate TMR allowed localization of P_{spoIIR}-*SNAP*^{Cd} expression. When examined by fluorescence microscopy, most of the cells of strain 630Δ*erm* showed fluorescence in the forespore as expected for a gene under σ^F control (Figure 4.5). *spoIIR* expression was not controlled by σ^E since a fluorescence signal was detected in the small compartments of disporic cells in a *sigE* mutant. Interestingly, a fluorescence of the P_{spoIIR}-*SNAP*^{Cd} fusion was still observed in the *sigF* mutant suggesting that some expression of *spoIIR* occurs in the absence of σ^F. In a *spo0A* mutant, the fluorescence completely disappeared and the *spoIIR* expression strongly decreased by qRT-PCR. So, the residual expression observed in the *sigF* mutant is likely Spo0A-dependent. Presumably, the expression of *spoIIR* in the *sigF* mutant allows the production of sufficient SpoIIR to trigger pro-σ^E processing. It is worth noting that *spoIIR* is not downregulated in a *sigF* mutant in *C. acetobutylicum* and only partially under σ^F control in *C. difficile* (Figure 4.5). Also interestingly, in *B. subtilis*, the forced expression of *spoIIR*

in pre-divisional cells, still allows pro-σ^E processing to occur, even in the absence of σ^F (Londono-Vallejo and Stragier, 1995, Zhang *et al.*, 1996). It is possible that the activation of σ^E in ancestral endospore formers occurred independently of the forespore. The expression of *spoIIR* under the exclusive control of σ^F may have appeared later and would allow a better coordination between the forespore and mother cell lines of gene expression.

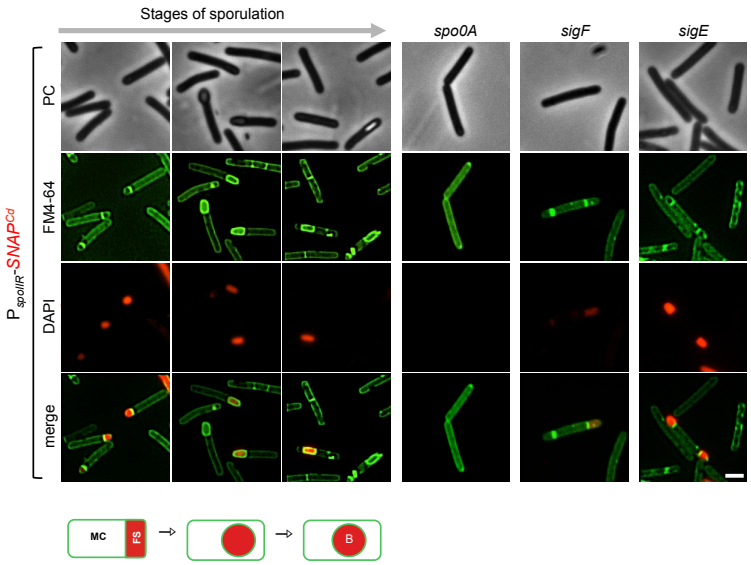


Figure 4.5. Fluorescence of a P$_{spoIIR}$-*SNAP*^{Cd} fusion in strain 630Δ*erm* and in a *spo0A*, *sigF* or *sigE* mutant. Cells of the *C. difficile* 630Δ*erm* strain, and of the *spo0A*, *sigF* and *sigE* mutants carrying a P$_{spoIIR}$-*SNAP*^{Cd} transcriptional fusion in a multicopy plasmid were collected 24 h of following inoculation in SM broth. Cells were labelled with the fluorescent TMR-Star SNAP substrate and with the membrane dye MTG and examined by phase contrast and fluorescence microscopy. A schematic representation of the deduced spatial and temporal pattern of *P$_{spoIIR}$-SNAP*^{Cd} transcription in the *wt* background is shown below the main panel. The cell membrane is represented in green. MC, mother cell; FS, forespore; B, phase bright spore. Scale bar, 1 µm.

The loose regulation of σ^G-dependent genes by σ^E

In *B. subtilis,* most of the σ^G activity occurs after engulfment completion. In addition, the expression of *sigG* target genes in the engulfed forespore depends upon σ^E activation in the mother cell at least in part through synthesis of the SpoIIIA proteins (Camp and Losick, 2008; Higgins and Dworkin, 2012; Hilbert and Piggot, 2004; Meisner *et al.*, 2012). In the

transcriptome analysis, we observed that the expression of 6 genes belonging to the σ^G regulon (*sspB*, *CD1486*, *CD1631*, *CD1880, CD2465* and *CD3551.1*) was downregulated in the *sigE* mutant (Table 4.3). We also examined by qRT-PCR if a *sigE* mutation affected the expression of several other σ^G targets. The expression of 9 additional σ^G-controlled genes decreased in a *sigE* mutant (Table 4.3). Interestingly, SpoIIID also repressed the expression of 12 members of the σ^G regulon in transcriptome (Table A7). This suggests that both σ^E and SpoIIID participated in the control of σ^G-target genes in *C. difficile*. In *B. subtilis,* σ^G activity is dependent on the SpoIIIA-SpoIIQ channel (Camp and Losick, 2008; Higgins and Dworkin, 2012; Meisner *et al.*, 2012). The SpoIIIA proteins are encoded by the octacistronic *spoIIIAA-AH* operon, which is expressed in the mother cell under the direct control of σ^E and SpoIIID (Eichenberger *et al.*, 2004; Eichenberger *et al.*, 2003; Steil *et al.*, 2005). As noted above, the *spoIIIAA* operon is a member of the σ^E and *spoIIID* regulons also in *C. difficile* (Table A4 and Table A7) and is transcribed from two promoters recognized by σ^E located upstream of the *spoIIIAA* and *spoIIIAG* cistrons, as observed in *B. subtilis* (Guillot and Moran, 2007). It thus seems possible that the control of forespore-specific gene expression by σ^E and SpoIIID involves similar mechanisms in *C. difficile*. However, we note that fluorescence from a P*sspA-SNAP*^{Cd} fusion is detected in the forespore compartment of *sigE* mutant cells (Pereira *et al.*, 2013; Chapter 3). This indicates that a strict requirement for σ^E is not observed for the expression of *C. difficile* σ^G-targets contrary to what is shown in *B. subtilis* suggesting a less tight control of σ^E on gene expression in the forespore.

Table 4.3. Control of σ^G or σ^K target genes by both σ^F and σ^E

Gene	Function	Expression ration		
Forespore σ^G-dependent		***sigF*/ 630Δ*erm***	***sigE*/ 630Δ*erm***	***sigG*/ 630Δ*erm***
transcriptome				
CD1486	Putative ribosome recycling factor	0.10	0.26	0.17
sodA	Superoxide dismutase (Mn)	0.06	0.21	0.32
CD1880	Conserved hypothetical protein	0.07	0.21	0.25
sspB	Small, acid-soluble spore protein beta	0.01	0.04	0.07
CD3551	Putative membrane protein	0.11	0.31	0.18
qRT-PCR				
spoVAC	DPA uptake protein, SpoVAC	<0.01	0.06	0.02
CD0792	Putative membrane protein, DUF81 family	<0.01	0.02	0.01
CD1213	Stage IVB sporulation protein B, peptidase	0.04	0.08	0.12
CD1290	Putative small acid-soluble spore protein	<0.01	0.05	0.01
pdaA	Putative d-lactam-biosynthetic deacteylase	<0.01	0.14	0.03
CD2375	Conserved hypothetical protein	0.03	0.24	0.05
CD2465	Putative amino acid/polyamine	0.18	0.15	0.19
CD2636	Putative membrane protein	<0.01	0.13	0.03
sspA	Small, acid-soluble spore protein alpha	<0.01	0.05	0.02
spoVT	Transcriptional regulator, SpoVT	<0.01	0.06	0.07
Mother cell σ^K-dependent		***sigF*/ 630Δ*erm***	***sigE*/ 630Δ*erm***	***sigK*/ 630Δ*erm***
transcriptome				
bclA1	Putative exosporium glycoprotein	0.43	0.21	0.08
cotCB	Spore-coat protein CotCB manganese	0.18	0.18	0.01
CD0896	Conserved hypothetical protein	0.21	0.05	0.02
CD1067	Conserved hypothetical protein	0.07	0.00	0.03
CD1581	Conserved hypothetical protein	0.03	0.01	0.01
cotA	Spore outer coat layer protein CotA	0.13	0.10	0.05
murE	UDP-N-acetylmuramyl-tripeptide	0.51	0.45	0.42
bclA2	Putative exosporium glycoprotein	0.16	0.12	0.04
bclA3	Exosporium glycoprotein BclA3	0.12	0.01	0.02
CD3620	Conserved hypothetical protein	0.35	0.34	0.04
qRT-PCR				
sleC	Spore cortex-lyic enzyme pre-pro-form	0.34	<0.01	0.01
CD1133	Conserved hypothetical protein	0.40	0.02	0.10
cotE	Spore coat protein CotE	0.32	0.01	0.01
cotD	Spore coat protein CotD manganese	0.24	0.01	<0.01

The absence of control of the σ^K regulon by σ^G

In *B. subtilis*, *sigG* regulates the expression of the σ^K regulon in the mother cell through the control of pro-σ^K processing (Higgins and Dworkin, 2012; Hilbert and Piggot, 2004). A *sigG* mutant is blocked just after engulfment completion, and does not show any signs of assembly of the surface layers around the forespore. Surprisingly, no σ^K target genes were

downregulated in a *sigG* mutant in the transcriptome analysis and this lack of effect of the *sigG* mutation was confirmed by qRT-PCR for eight σ^K-controlled genes (*cotA*, *cotCB*, *cotD*, *bclA1*, *bclA2*, *bclA3*, *CD1067*, *CD3620*). Thus, synthesis of the spore coat proteins belonging to the σ^K regulon (CotA, CotCB, CotD, CotE, CotF) is σ^G-independent. This result is in agreement with the phenotype of the *C. difficile sigG* mutant that shows deposition of some coat material around the engulfed forespore (Pereira *et al.*, 2013; Chapter 3). Interestingly, mutations that bypass the need for σ^G for pro-σ^K processing, or a pro-less allele of the *sigK* gene in *B. subtilis* allow expression of σ^K targets (Cutting *et al.*, 1990). In *C. difficile*, the pro-sequence of σ^K is absent and no homologs of SpoIVFB, the membrane-embedded protease that processes pro-σ^K in *B. subtilis* and of two other membrane proteins (SpoIVFA and BofA) that form a complex with and control the localization and activity of SpoIVFB, can be found (Abecasis *et al.*, 2013; Galperin *et al.*, 2012; Haraldsen and Sonenshein, 2003; Traag *et al.*, 2013). Thus, σ^K activity is not controlled through processing of a pre-protein in *C. difficile* and this may be related to the absence of σ^G control.

Intriguingly, *C. difficile* codes for two homologs of the *B. subtilis* signaling protein SpoIVB. CD1213 and CD0783 share 37 % and 36 % identity with SpoIVB, respectively. *CD0783* and *CD1213* belong to the σ^F and σ^G regulons, respectively (Table A3). In *B. subtilis,* SpoIVB is a protease synthesized under the control of σ^F and σ^G. Similarly, *spoIIP* is transcribed by σ^F and σ^E in *B. subtilis*, whereas two *spoIIP* genes are present in *B. anthracis* and transcribed by σ^F and σ^E, respectively (Dworkin and Losick, 2005). SpoIVB is secreted to the intermembrane space where it cleaves SpoIVFA releasing SpoIVFB and allowing pro-σ^K processing (Higgins and Dworkin, 2012; Hilbert and Piggot, 2004). However, *B. subtilis* SpoIVB has two distinct developmental functions: pro-σ^K processing control and another role essential for spore formation probably corresponding to the cleavage of SpoIIQ (Jiang *et al.*, 2005; Oke *et al.*, 1997; Cutting *et al.*, 1990). It is possible

that the *C. difficile* SpoIVB-like proteins retained the second role, which might correspond to an ancestral function.

Although *sigG* disruption has no impact on σ^K target genes, we observed that a set of σ^K target genes were downregulated in a *sigF* mutant compared to strain 630Δ*erm* in transcriptome, and these results were further confirmed by qRT-PCR (Table 4.3). This effect might be explained by the fact that the σ^K activity requires engulfment completion, a process blocked in a *sigF* mutant but not in a *sigG* mutant.

DISCUSSION

Global approaches combining transcriptome and TSS mapping allow us to have a view of the expression pattern during the early and late stages of sporulation in both forespore and mother cell and to define the four compartment-specific sigma regulons in *C. difficile*. We identified 25, 97, 50 and 56 members for the σ^F, σ^E, σ^G and σ^K regulons, respectively and in each regulon, we found key representatives of the homologous regulons of *B. subtilis* as proposed previously (de Hoon *et al.*, 2010). However, while larger than the evolutionary conserved core machinery for endosporulation, which consists of about 145 genes (Abecassis *et al.*, 2013; Galperin *et al.*, 2012), this set of *C. difficile* sporulation genes (around 225 genes) corresponds to about half the number of genes under the control of cell type-specific sigma factors of *B. subtilis* (Eichenberger *et al.*, 2003; Steil *et al.*, 2005; Wang *et al.*, 2006), and is in agreement with other results recently reported (Fimlaid *et al.*, 2013). A more dynamic study with a detailed kinetic analysis would provide additional members for each regulon but it is also probable that the presumably more ancestral type of sporulation process proposed for *Clostridia* (Abecasis *et al.*, 2013; Galperin *et al.*, 2012; Paredes *et al.*, 2005; Stragier, 2002) involves a smaller collection of proteins. This has already

been observed for the initiation of sporulation where the complex signaling transduction pathway involving in *B. subtilis* a phosphorelay that modulates Spo0A activity is replaced by a simple two-component system (Paredes *et al.*, 2005; Steiner *et al.*, 2011; Stragier, 2002). In general, the most significant variations between the *B. subtilis* and *C. difficile* sporulation process are observed at the interface with their environment: the signal transduction pathway triggering sporulation initiation, composition of the coat shell and the germination-activating pathways (de Hoon *et al.*, 2010; Paredes *et al.*, 2005; Paredes-Sabja *et al.*, 2011; Stragier, 2002). Indeed, the germination signals and germination receptors differ among spore forming Firmicutes and only a few spore coat layer proteins are conserved in *Bacilli* and *Clostridia* (Henriques and Moran, 2007; Paredes-Sabja *et al.*, 2011). In each regulon, a significant fraction of genes encodes proteins with unknown function but our work offer new insights about the role of some of them. Some σ^G-controlled membrane proteins might be involved in germination, which remain poorly characterized in *C. difficile* (Burns *et al.*, 2010a; Francis *et al.*, 2013). Three proteins of unknown function belonging to the σ^E regulon (CD0214, CD1930, CD3522) and several proteins synthesized under σ^K control (CD1063.1, CD1067, CD1133, CD3580 and CD3613) that are detected in the spore proteome (Lawley *et al.*, 2009) are probably new spore coat components. Finally, we found several oxygen detoxification proteins that might help in the long-term survival of clostridial spores as recently proposed by Galperin *et al* (Galperin *et al.*, 2012). This probably favors the dissemination of spores of strictly anaerobic *Clostridia* in aerobic environment, a crucial step for persistence and transmission of pathogenic *Clostridia* (Deakin *et al.*, 2012; Lawley *et al.*, 2010).

In this work, we also exposed important deviations from the *B. subtilis* paradigm in *C. difficile.* Both the global analysis of the program of gene expression under the control of the four cell type-specific sigma factors and the morphological characterization of the corresponding mutants

indicate that coupling between gene expression and morphogenesis is less tight in *C. difficil*e than in *B. subtilis* (Figure 4.6) (Saujet *et al.*, 2013; Pereira *et al.*, 2013; Fimlaid *et al.*, 2013). First, the σ^E regulon in the *C. difficile* mother cell is not strictly under σ^F control despite the fact that the forespore product SpoIIR is required for pro-σ^E processing. The residual *spoIIR* expression in the *sigF* mutant might be responsible for this less strict connection. Second, the tight coordination between σ^G and σ^K activities observed in *B. subtilis,* is absent in *C. difficile* since σ^K activity does not depend on σ^G as clearly shown by the morphology of a *sigK* mutant (Pereira *et al.*, 2013; Chapter 3). In the absence of a σ^K precursor, the rearrangement of the *sigK* gene associated to excision of the *skin* element may be the only way to control the timing of σ^K activity in *C. difficile* (de Hoon *et al.*, 2010; Haraldsen and Sonenshein, 2003; Stragier, 2002) and this event does not appear to be under σ^G control.

Recent data discussed above also suggest differences in the regulatory network controlling sporulation among *Clostridia*. This is especially the case for the impact of each sigma factor inactivation on the synthesis of the others and for the role of σ^K (Harry *et al.*, 2009; Kirk *et al.*, 2012; Li and McClane, 2010; Wang *et al.*, 2012). The timing of *sigK* expression, the phenotype of *sigK* mutants and/or some σ^K targets differ. In addition, with the exception of *C. difficile,* σ^K activity is controlled through processing in *Clostridia* while the insertion of a *skin* element in the *sigK* gene is found only in *C. difficile* strains (Haraldsen and Sonenshein, 2003). The rather low probability to observe orthologs in clostridial genomes for many *C. difficile* regulon members also suggests a moderate conservation of the sporulation sigma-factor regulons among *Clostridia*. This work gives new insights about the diversity and evolution of the sporulation process. The sporulation in *C. difficile* might reflect a more ancestral way of sporulation while a more sophisticated system of developmental control would have been gradually introduced during evolution. In general, the conclusions herein presented were in agreement with those of Fimlaid and co-authors

(Fimlaid *et al.*, 2013). Discrepancies will be considered in the General Discussion of this Thesis.

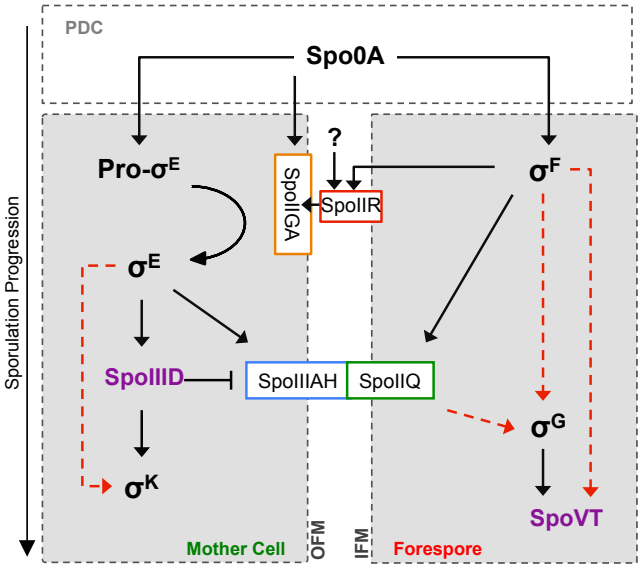


Figure 4.6. Model of the regulatory network controlling sporulation in *C. difficile*. Temporal progression of sporulation is shown from top to bottom as indicated. The lighter box (top) represents the predivisional cell (PDC), and the darker boxes represent the forespore and the mother cell, the two cell compartments formed after asymmetric division. The two parallel vertical lines represent the membranes separating the forespore and the mother cell (OFM: outer forespore membrane; IFM: inner forespore membrane). Sigma factors are shown in black and bold, with the precursor protein of σ^E indicated as pro-σ^E. Transcription factors are shown in purple and bold. Proteins associated with the membranes or located into the intermembrane space (the space between the two parallel broken lines) are in square boxes and illustrated as contacting the parallel vertical line. Black solid arrows indicate activation at the transcriptional level. Broken red arrows indicate a possible direct transcriptional activation or repression.

ACKNOWLEDGEMENTS

The author of this dissertation participated in all the experiments described in this chapter except in the microarray analysis, RNA-seq and quantitative real-time PCR experiments, which were performed by Laure Saujet, Olga Soutorina and Marc Monot. We thank Rita Tomé and Mónica Serrano for help with the SNAP labeling experiments. We thank Emilie

Camiade for helpful discussions and Isabelle Poquet for her comments on the manuscript. We are grateful to Patrick Stragier for a critical reading of the manuscript and helpful advice. This work was supported by grant ERA-PTG/SAU/0002/2008 (ERA-NET PathoGenoMics) and PEst-OE/EQB/LA0004/2011 from the "Fundação para a Ciência e a Tecnologia" (FCT). F.P. (SFRH/BD/45459/08) was the recipient of doctoral fellowship from the FCT.

REFERENCES

Abecasis, A., Serrano, M., Alves, R., Quintais, L., Pereira-Leal, J. B. & Henriques, A. (2013). A genomic signature and the identification of new endosporulation genes. . *J. Bacteriol* **195**, 2101-2115.

Adams, C. M., Eckenroth, B. E., Putnam, E. E., Doublie, S. & Shen, A. (2013). Structural and Functional Analysis of the CspB Protease Required for Clostridium Spore Germination. *PLoS Pathog* **9**, e1003165.

Bagyan, I., Hobot, J. & Cutting, S. (1996). A compartmentalized regulator of developmental gene expression in *Bacillus subtilis*. *J Bacteriol* **178**, 4500-4507.

Benjamini, Y., Hochberg Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Statist. Soc. Ser.*, 289--300.

Burns, D. A., Heap, J. T. & Minton, N. P. (2010a). Clostridium difficile spore germination: an update. *Res Microbiol* **161**, 730-4.

Burns, D. A., Heap, J. T. & Minton, N. P. (2010b). SleC is essential for germination of *Clostridium difficile* spores in nutrient-rich medium supplemented with the bile salt taurocholate. *J Bacteriol* **192**, 657-664.

Camp, A. H. & Losick, R. (2008). A novel pathway of intercellular signalling in Bacillus subtilis involves a protein with similarity to a component of type III secretion channels. *Mol Microbiol* **69**, 402-17.

Cangiano, G., Mazzone, A., Baccigalupi, L., Isticato, R., Eichenberger, P., De Felice, M. & Ricca, E. (2010). Direct and indirect control of late sporulation genes by GerR of *Bacillus subtilis*. *J Bacteriol* **192**, 3406-3413.

Cartman, S. T. & Minton, N. P. (2010). A mariner-Based Transposon System for In Vivo Random Mutagenesis of Clostridium difficile. *Appl Environ Microbiol* **76**, 1103-9.

Cutting, S., Oke, V., Driks, A., Losick, R., Lu, S. & Kroos, L. (1990). A forespore checkpoint for mother cell gene expression during development in B. subtilis. *Cell* **62**, 239-50.

de Hoon, M. J., Eichenberger, P. & Vitkup, D. (2010). Hierarchical evolution of the bacterial sporulation network. *Curr Biol* **20**, R735-745.

Deakin, L. J., Clare, S., Fagan, R. P., Dawson, L. F., Pickard, D. J., West, M. R., Wren, B. W., Fairweather, N. F., Dougan, G. & Lawley, T. D. (2012). The Clostridium difficile spo0A gene is a persistence and transmission factor. *Infect Immun* **80**, 2704-11.

Dworkin, J. & Losick, R. (2005). Developmental commitment in a bacterium. *Cell* **121**, 401-9.

Eichenberger, P., Fawcett, P. & Losick, R. (2001). A three-protein inhibitor of polar septation during sporulation in Bacillus subtilis. *Mol Microbiol* **42**, 1147-62.

Eichenberger, P., Fujita, M., Jensen, S. T., Conlon, E. M., Rudner, D. Z., Wang, S. T., Ferguson, C., Haga, K., Sato, T., Liu, J. S. & Losick, R. (2004). The program of gene transcription for a single differentiating cell type during sporulation in Bacillus subtilis. *PLoS Biol* **2**, e328.

Eichenberger, P., Jensen, S. T., Conlon, E. M., van Ooij, C., Silvaggi, J., Gonzalez-Pastor, J. E., Fujita, M., Ben-Yehuda, S., Stragier, P., Liu, J. S. & Losick, R. (2003). The sigmaE regulon and the identification of additional sporulation genes in Bacillus subtilis. *J Mol Biol* **327**, 945-72.

Eldar, A., Chary, V. K., Xenopoulos, P., Fontes, M. E., Loson, O. C., Dworkin, J., Piggot, P. J. & Elowitz, M. B. (2009). Partial penetrance facilitates developmental evolution in bacteria. *Nature* **460**, 510-4.

Errington, J. (1993). *Bacillus subtilis* sporulation: regulation of gene expression and control of morphogenesis. *Microbiol Rev* **57**, 1-33.

Evans, L., Clarkson, J., Yudkin, M. D., Errington, J. & Feucht, A. (2003). Analysis of the interaction between the transcription factor sigmaG and the anti-sigma factor SpoIIAB of Bacillus subtilis. *J Bacteriol* **185**, 4615-9.

Fawcett, P., Eichenberger, P., Losick, R. & Youngman, P. (2000). The transcriptional profile of early to middle sporulation in *Bacillus subtilis*. *Proc Natl Acad Sci U S A* **97**, 8063-8068.

Fimlaid, K. A., Bond, J. P., Schutz, K. C., Putnam, E. E., Leung, J. M., Lawley, T. D. & Shen, A. (2013). Global analysis of the sporulation pathway of Clostridium difficile. *PLoS Genet* **9**, e1003660.

Francis, M. B., Allen, C. A., Shrestha, R. & Sorg, J. A. (2013). Bile acid recognition by the Clostridium difficile germinant receptor, CspC, is important for establishing infection. *PLoS Pathog* **9**, e1003356.

Galperin, M. Y., Mekhedov, S. L., Puigbo, P., Smirnov, S., Wolf, Y. I. & Rigden, D. J. (2012). Genomic determinants of sporulation in Bacilli and Clostridia: towards the minimal set of sporulation-specific genes. *Environ Microbiol* **14**, 2870-90.

Gilmore, M., Bandyopadhyay D, Dean AM, Linnstaedt SD & DL., P. (2004). Production of muramic delta-lactam in *Bacillus subtilis* spore peptidoglycan. *J Bacteriol.* **186**, 80-89.

Gonzalez-Pastor, J. E., Hobbs, E. C. & Losick, R. (2003). Cannibalism by sporulating bacteria. *Science* **301**, 510-3.

Guillot, C. & Moran, C. P., Jr. (2007). Essential internal promoter in the spoIIIA locus of *Bacillus subtilis*. *J Bacteriol* **189**, 7181-7189.

Haraldsen, J. D. & Sonenshein, A. L. (2003). Efficient sporulation in Clostridium difficile requires disruption of the sigmaK gene. *Mol Microbiol* **48**, 811-21.

Harris, L. M., Welker, N. E. & Papoutsakis, E. T. (2002). Northern, morphological, and fermentation analysis of *spo0A* inactivation and overexpression in *Clostridium acetobutylicum* ATCC 824. *J Bacteriol* **184**, 3586-3597.

Harry, K. H., Zhou, R., Kroos, L. & Melville, S. B. (2009). Sporulation and enterotoxin (CPE) synthesis are controlled by the sporulation-specific sigma factors SigE and SigK in Clostridium perfringens. *J Bacteriol* **191**, 2728-42.

Heap, J. T., Pennington, O. J., Cartman, S. T., Carter, G. P. & Minton, N. P. (2007). The ClosTron: a universal gene knock-out system for the genus Clostridium. *J Microbiol Methods* **70**, 452-64.

Heap, J. T., Pennington, O. J., Cartman, S. T. & Minton, N. P. (2009). A modular system for Clostridium shuttle plasmids. *J Microbiol Methods* **78**, 79-85.

Henriques, A. O. & Moran, C. P., Jr. (2007). Structure, assembly, and function of the spore surface layers. *Annu Rev Microbiol* **61**, 555-88.

Higgins, D. & Dworkin, J. (2012). Recent progress in Bacillus subtilis sporulation. *FEMS Microbiol Rev* **36**, 131-48.

Hilbert, D. W. & Piggot, P. J. (2004). Compartmentalization of gene expression during Bacillus subtilis spore formation. *Microbiol Mol Biol Rev* **68**, 234-62.

Himes, P., McBryant, S. J. & Kroos, L. (2010). Two regions of *Bacillus subtilis* transcription factor SpoIIID allow a monomer to bind DNA. *J Bacteriol* **192**, 1596-1606.

Hussain, H. A., Roberts, A. P. & Mullany, P. (2005). Generation of an erythromycin-sensitive derivative of Clostridium difficile strain 630 (630Deltaerm) and demonstration that the conjugative transposon Tn916DeltaE enters the genome of this strain at multiple sites. *J Med Microbiol* **54**, 137-41.

Jiang, X., Rubio, A., Chiba, S. & Pogliano, K. (2005). Engulfment-regulated proteolysis of SpoIIQ: evidence that dual checkpoints control sigma activity. *Mol Microbiol* **58**, 102-15.

Jonas, R. M. & Haldenwang, W. G. (1989). Influence of *spo* mutations on sigma E synthesis in *Bacillus subtilis*. *J Bacteriol* **171**, 5226-5228.

Jones, S. W., Paredes, C. J., Tracy, B., Cheng, N., Sillers, R., Senger, R. S. & Papoutsakis, E. T. (2008). The transcriptional program underlying the physiology of clostridial sporulation. *Genome Biol* **9**, R114.

Jones, S. W., Tracy, B. P., Gaida, S. M. & Papoutsakis, E. T. (2011). Inactivation of sigmaF in Clostridium acetobutylicum ATCC 824 blocks sporulation prior to asymmetric division and abolishes sigmaE and sigmaG protein expression but does not block solvent formation. *J Bacteriol* **193**, 2429-40.

Karow, M. L., Glaser, P. & Piggot, P. J. (1995). Identification of a gene, spoIIR, that links the activation of sigma E to the transcriptional activity of sigma F during sporulation in Bacillus subtilis. *Proc Natl Acad Sci U S A* **92**, 2012-6.

Kirk, D. G., Dahlsten, E., Zhang, Z., Korkeala, H. & Lindstrom, M. (2012). Involvement of Clostridium botulinum ATCC 3502 sigma factor K in early-stage sporulation. *Appl Environ Microbiol* **78**, 4590-6.

Lawley, T. D., Clare, S., Deakin, L. J., Goulding, D., Yen, J. L., Raisen, C., Brandt, C., Lovell, J., Cooke, F., Clark, T. G. & Dougan, G. (2010). Use of purified Clostridium difficile spores to facilitate evaluation of health care disinfection regimens. *Appl Environ Microbiol* **76**, 6895-900.

Lawley, T. D., Croucher, N. J., Yu, L., Clare, S., Sebaihia, M., Goulding, D., Pickard, D. J., Parkhill, J., Choudhary, J. & Dougan, G. (2009). Proteomic and genomic characterization of highly infectious Clostridium difficile 630 spores. *J Bacteriol* **191**, 5377-86.

Li, J. & McClane, B. A. (2010). Evaluating the involvement of alternative sigma factors SigF and SigG in Clostridium perfringens sporulation and enterotoxin synthesis. *Infect Immun* **78**, 4286-93.

Livak, K. J. & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402-408.

Londono-Vallejo, J. A. & Stragier, P. (1995). Cell-cell signaling pathway activating a developmental transcription factor in Bacillus subtilis. *Genes Dev* **9**, 503-8.

Meisner, J., Maehigashi, T., Andre, I., Dunham, C. M. & Moran, C. P., Jr. (2012). Structure of the basal components of a bacterial transporter. *Proc Natl Acad Sci U S A* **109**, 5446-5451.

Morlot, C., Uehara, T., Marquis, K. A., Bernhardt, T. G. & Rudner, D. Z. (2010). A highly coordinated cell wall degradation machine governs spore morphogenesis in *Bacillus subtilis*. *Genes Dev* **24**, 411-422.

Nugroho, F. A., Yamamoto, H., Kobayashi, Y. & Sekiguchi, J. (1999). Characterization of a new sigma-K-dependent peptidoglycan hydrolase gene that plays a role in *Bacillus subtilis* mother cell lysis. *J Bacteriol* **181**, 6230-6237.

Oke, V. & Losick, R. (1993). Multilevel regulation of the sporulation transcription factor sigma K in Bacillus subtilis. *J Bacteriol* **175**, 7341-7.

Oke, V., Shchepetov, M. & Cutting, S. (1997). SpoIVB has two distinct functions during spore formation in *Bacillus subtilis*. *Mol Microbiol* **23**, 223-230.

Paredes, C. J., Alsaker, K. V. & Papoutsakis, E. T. (2005). A comparative genomic view of clostridial sporulation and physiology. *Nat Rev Microbiol* **3**, 969-78.

Paredes-Sabja, D., Sarker, N., Setlow, B., Setlow, P. & Sarker, M. R. (2008). Roles of DacB and spm proteins in *Clostridium perfringens* spore resistance to moist heat, chemicals, and UV radiation. *Appl Environ Microbiol* **74**, 3730-3738.

Paredes-Sabja, D., Setlow, P. & Sarker, M. R. (2011). Germination of spores of Bacillales and Clostridiales species: mechanisms and proteins involved. *Trends Microbiol* **19**, 85-94.

Pereira, F. C., Saujet, L., dos Vultos, T., Monot, M., Couture-Tosi, E., I, M.-V., Dupuy, B. & Henriques, A. O. (2013). The spore differentiation pathway in the enteric pathogen *Clostridium difficile*. *PLoS Genet* **in press**.

Permpoonpattana, P., Phetcharaburanin, J., Mikelsone, A., Dembek, M., Tan, S., Brisson, M. C., La Ragione, R., Brisson, A. R., Fairweather, N., Hong, H. A. & Cutting, S. M. (2013). Functional characterization of Clostridium difficile spore coat proteins. *J Bacteriol* **195**, 1492-503.

Permpoonpattana, P., Tolls, E. H., Nadem, R., Tan, S., Brisson, A. & Cutting, S. M. (2011). Surface layers of Clostridium difficile endospores. *J Bacteriol* **193**, 6461-70.

Putman, E., Nock, A., Lawley, T. & Shen, A. (2013). SpoIVA and SipL are *Clostridium difficile* spore morphogenic proteins. *J. Bacteriol* **195**, 1214-1225.

Ramirez-Peralta, A., Stewart, K. A., Thomas, S. K., Setlow, B., Chen, Z., Li, Y. Q. & Setlow, P. (2012). Effects of the SpoVT regulatory protein on the germination and germination protein levels of spores of *Bacillus subtilis*. *J Bacteriol* **194**, 3417-3425.

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

Santangelo, J. D., Kuhn, A., Treuner-Lange, A. & Durre, P. (1998). Sporulation and time course expression of sigma-factor homologous genes in *Clostridium acetobutylicum*. *FEMS Microbiol Lett* **161**, 157-164.

Sarker, M. R. & Paredes-Sabja, D. (2012). Molecular basis of early stages of Clostridium difficile infection: germination and colonization. *Future Microbiol* **7**, 933-43.

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**, 3186-96.

Saujet, L., Pereira, F. C., Serrano, M., Soutourina, O., Monot, M., Shelyakin, P. V., Gelfand, M. S., Dupuy, B., Henriques, A. O. & Martin-Verstraete, I. (2013). Genome-wide analysis of cell type-specific gene expression during spore formation in Clostridium difficile.

Serrano, M., Real, G., Santos, J., Carneiro, J., Moran, C. P., Jr. & Henriques, A. O. (2011). A negative feedback loop that limits the ectopic activation of a cell type-specific sporulation sigma factor of Bacillus subtilis. *PLoS Genet* **7**, e1002220.

Serrano, M., Zilhao, R., Ricca, E., Ozin, A. J., Moran, C. P., Jr. & Henriques, A. O. (1999). A *Bacillus subtilis* secreted protein with a role in endospore coat assembly and function. *J Bacteriol* **181**, 3632-3643.

Setlow, P. (2007). I will survive: DNA protection in bacterial spores. *Trends Microbiol* **15**, 172-180.

Smyth, G. K. & Speed, T. (2003). Normalization of cDNA microarray data. *Methods* **31**, 265--273.

Sorg, J. A. & Sonenshein, A. L. (2008). Bile salts and glycine as cogerminants for Clostridium difficile spores. *J Bacteriol* **190**, 2505-12.

Soutourina, O., Monot, M., Boudry, P., Saujet, L., Pichon, C., Sismeiro, O., Semenova, E., Severinov, K., Le Bouguenec, C., Coppée, J. Y., Dupuy, B. & Martin-Verstraete, I. (2013). Genome-wide identification of regulatory RNAs in the human pathogen *Clostridium difficile*. *PLoS Genet* **9**, e1003493.

Steichen, C. T., Kearney, J. F. & Turnbough, C. L., Jr. (2007). Non-uniform assembly of the *Bacillus anthracis* exosporium and a bottle cap model for spore germination and outgrowth. *Mol Microbiol* **64**, 359-367.

Steil, L., Serrano, M., Henriques, A. O. & Volker, U. (2005). Genome-wide analysis of temporally regulated and compartment-specific gene expression in sporulating cells of Bacillus subtilis. *Microbiology* **151**, 399-420.

Steiner, E., Dago, A. E., Young, D. I., Heap, J. T., Minton, N. P., Hoch, J. A. & Young, M. (2011). Multiple orphan histidine kinases interact directly with Spo0A to control the initiation of endospore formation in *Clostridium acetobutylicum*. *Mol Microbiol* **80**, 641-654.

Stragier, P. (2002). A gene odyssey: exploring the genomes of endospore-forming bacteria. ASM Press: Washington, D. C.

Stragier, P. & Losick, R. (1996). Molecular genetics of sporulation in Bacillus subtilis. *Annu Rev Genet* **30**, 297-41.

Tovar-Rojo, F., Chander, M., Setlow, B. & Setlow, P. (2002). The products of the *spoVA* operon are involved in dipicolinic acid uptake into developing spores of *Bacillus subtilis. J Bacteriol* **184**, 584-587.

Traag, B., Pugliese, A., Eisen, J. & Losick, R. (2013). Gene Conservation among Endospore-Forming Bacteria Reveals Additional Sporulation Genes in *Bacillus subtilis. J. Bacteriol* **195**, 253-260.

Tracy, B. P., Jones, S. W. & Papoutsakis, E. T. (2011). Inactivation of sigmaE and sigmaG in Clostridium acetobutylicum illuminates their roles in clostridial-cell-form biogenesis, granulose synthesis, solventogenesis, and spore morphogenesis. *J Bacteriol* **193**, 1414-26.

Underwood, S., Guan, S., Vijayasubhash, V., Baines, S. D., Graham, L., Lewis, R. J., Wilcox, M. H. & Stephenson, K. (2009). Characterization of the sporulation initiation pathway of Clostridium difficile and its role in toxin production. *J Bacteriol* **191**, 7296-305.

Wang, S. T., Setlow, B., Conlon, E. M., Lyon, J. L., Imamura, D., Sato, T., Setlow, P., Losick, R. & Eichenberger, P. (2006). The forespore line of gene expression in Bacillus subtilis. *J Mol Biol* **358**, 16-37.

Wang, Y., Li, X., Mao, Y. & Blaschek, H. P. (2012). Genome-wide dynamic transcriptional profiling in *Clostridium beijerinckii* NCIMB 8052 using single-nucleotide resolution RNA-Seq. *BMC Genomics* **13**, 102.

Wilson, K. H., Kennedy, M. J. & Fekety, F. R. (1982). Use of sodium taurocholate to enhance spore recovery on a medium selective for Clostridium difficile. *J Clin Microbiol* **15**, 443-6.

Xiao, Y., Francke, C., Abee, T. & Wells-Bennik, M. H. (2011). Clostridial spore germination versus bacilli: genome mining and current insights. *Food Microbiol* **28**, 266-74.

Yoshisue, H., Ihara, K., Nishimoto, T., Sakai, H. & Komano, T. (1995). Cloning and characterization of a *Bacillus thuringiensis* homolog of the *spoIIID* gene from *Bacillus subtilis*. *Gene* **154**, 23-29.

Zhang, L., Higgins, M. L., Piggot, P. J. & Karow, M. L. (1996). Analysis of the role of prespore gene expression in the compartmentalization of mother cell-specific gene expression during sporulation of *Bacillus subtilis*. *J Bacteriol* **178**, 2813-2817.

Chapter 5

A *C. difficile* spore surface protein with a role in host colonization

This Chapter contains data already published in: Janoir, C., Denève, C., Bouttier, S., Barbut, F., Hoys, S., Caleechum, L., Chapetón-Montes, D., Pereira, F. C., Henriques, A. O., Collignon, A., Monot, M. and Dupuy, B. 2013. Adaptive strategies and pathogenesis of *Clostridium difficile* from *in vivo* transcriptomics. *Infect Immun* 81: 3757-69.

This chapter also contains data to be published in: Pereira, F. C., Saujet, L., Vultos, T., Monot, M., Couture-Tosi, E., Janoir, C., Martin-Verstraete, I., Dupuy, B. and Henriques, A. O. A protein required for host colonization by *Clostridium difficile* is a key determinant for the assembly of a spore polar appendage. To be submitted.

SUMMARY

C. difficile is a strict anaerobic spore-forming bacterium, a major causative agent of nosocomial diseases associated to antibiotic therapy in adults, and a growing concern in the community. Spores produced by *C. difficile* are the primary cause of transmission in health care institutions. Despite the central importance of spores in the pathogenesis of *C. difficile*, our knowledge of the spore-related mechanisms involved in host colonization and infection is still incomplete. Here, we show that a 17 kDa cysteine-rich protein, termed Sp17 (CD1581) is important for colonization in a mice axenic model. We further show that Sp17 is an abundant component of the *C. difficile* spore and undergoes extensive multimerization both *in vitro* and at the spore surface. Spores of an *sp17* mutant show reduced hydrophobicity, fail to assemble an electrondense outer layer of the spore coat, and Sp17 appears to be the main component of this layer. Sp17 is also a key determinant for the formation of a spore polar appendage structure, which may serve a role in adhesion. Importantly, Sp17 also acts to delay spore germination in response to the bile salt taurocholate. Consistent with a role in the assembly of the spore coat, Sp17 is produced during spore development under the control of the late mother cell-specific regulatory protein σ^K. Sp17 is unique to *C. difficile*, suggesting that the outer coat in spores of this organism has distinctive structural and functional properties. Our study unveils a link between the spore surface and infection. Both the altered surface of the spore as well as its more prompt germination may contribute to impaired colonization ability of the mutant.

INTRODUCTION

C. difficile infection (CDI) is the cause of an intestinal disease mediated by two potent cytotoxins, TcdA and TcdB (Carter *et al.*, 2012) (Kuehne *et al.*, 2010). CDI develops in hospitalized patients undergoing antibiotic treatment because *C. difficile* can colonize the gut if the normal intestinal microbiota is disturbed (Rupnik *et al.*, 2009; Deneve *et al.*, 2009). However, *C. difficile* is also emerging as an important pathogen in the community, as well as in animals used in the feed industry (Hensgens *et al.*, 2012). The organism is an obligate anaerobe, and has the ability to form spores. Spores are extremely resilient and can accumulate and remain viable in the environment or in the host for long periods of time. Spores that remain latent in the gut are responsible for the recurrence of *C. difficile*-associated disease (CDAD) when antibiotic therapy is stopped (Rupnik *et al.*, 2009; Deneve *et al.*, 2009). At least some of the hypervirulent epidemic strains display a greater sporulation capacity *in vitro*, as well as robust toxin production (Merrigan *et al.*, 2010; Burns *et al.*, 2011).

Spores produced by species of *Bacillus* and *Clostridium* develops inside a sporangium partitioned into a smaller forespore (the future spore) and a larger mother cell (which undergoes lysis at the end of the process). Spores are metabolically inert and arguably the most resistant life form known to us (Nicholson *et al.*, 2000; Nicholson, 2004). The extreme resilience of the spore stems from its functional architecture. Transmission electron microscopy reveals an inner compartment that houses the genome, surrounded by a thick layer of peptidoglycan called the cortex. The cortex is encased within a multiprotein coat, which in *C. difficile* spores appears composed of concentric rings (Henriques and Moran, 2007; Lawley *et al.*, 2009; Permpoonpattana *et al.*, 2011). The coat is the outermost spore structure in many organisms, and confers protection against noxious chemicals and peptidoglycan-breaking enzymes. It also mediates the

response of spores to compounds able to trigger germination. In several pathogenic organisms, including *C. difficile*, the outermost spore structure is a layer called exosporium (Henriques and Moran, 2007; Lawley *et al.*, 2009; Permpoonpattana *et al.*, 2011) (Chapter 3). This structure is best characterized in spores of the pathogen *B. anthracis*, the causative agent of anthrax. It consists of a basal layer and a hair-like nap, formed by radial projections of a long collagen-like glycosylated protein called BclA (Henriques and Moran, 2007; McKenney *et al.*, 2012). The exosporium is the first line of contact of the spore with cells of the immune system, and BclA is the immunodominant protein at the spore surface. BclA mediates recognition of *B. anthracis* spores by the integrin Mac-1, with CD14 acting as a co-receptor, at the surface of professional macrophages (Oliva *et al.*, 2008; Oliva *et al.*, 2009). These interactions are key to the infectious process, as they mediate spore uptake into professional macrophages, and the conveyance of spores to sites of germination and bacterial growth in distant lymphoid organs (Oliva *et al.*, 2008). The appearance of the *C. difficile* exosporium, as observed by TEM, is distinct than that observed for *B. anthracis* and *B. cereus*, and also shows strain to strain variation (Lawley *et al.*, 2009; Permpoonpattana *et al.*, 2011; Paredes-Sabja and Sarker, 2012; Joshi *et al.*, 2012) (Chapter 3). Importantly, its composition remains to be determined. Nevertheless, it has been implicated in adhesion of spores to cells and to various surfaces, including metal, glass and plastic, all of which are of relevance in healthcare facilities (Henriques and Moran, 2007; Joshi *et al.*, 2012).

C. difficile spores are the infectious and transmissible form of the organism (Paredes-Sabja and Sarker, 2012; Deakin *et al.*, 2012). Following spore ingestion, the spore has to establish itself firmly in the colon. Spore germination in the colon allows the vegetative form of the organism to propagate. Adherence to the colonic mucosa is thought to be important for both the spores and for the vegetative cells. *C. difficile* cells have been shown

to adhere *in vitro* to Caco-2, HT-29 and Vero cells, and *in vivo* they will bind to caecal mucus of mice and to extracellular matrix proteins (Deneve *et al.*, 2009). Several virulence factors that promote adherence and intestinal colonization by *C. difficile* cells are surface-exposed proteins. These include the S-layer protein SlpA, among others (Deneve *et al.*, 2009; Merrigan *et al.*, 2013). In line with the idea that the step of adherence to the colonic mucosa could be a target for vaccination, purified S-layer proteins can activate innate and adaptive immune responses through the Toll-like receptor TLR4, and TLR4$^{-/-}$ mice are more susceptible to experimental CDI (Ryan *et al.*, 2011).

In contrast, it is not known how the spore is implanted in the colonic mucosa. Yet, the spore has to be retained long enough so that germination can be completed. Evidence suggests that in contact with cells, the smooth surface of the dormant spore first develops a series of bumps and knobs along its entire surface, and a protruding structure starts to assemble at one pole (Panessa-Warren *et al.*, 2007; Paredes-Sabja and Sarker, 2012). This appendage thickens and extends during germination, and mediates attachment to the colonic microvilli. These observations suggest that spore germination occurs in close association with the colonic epithelial cells (Panessa-Warren *et al.*, 2007; Paredes-Sabja and Sarker, 2012).

Here, we show that Sp17 is an abundant component of the *C. difficille* spore surface layers and is required for colonization in a mouse axenic model. Sp17 is a key organizer of the assembly of the outer coat, and is also required for the morphogenesis of a spore polar appendage that may be an adhesion structure. Importantly, *sp17* mutant spores are less hydrophobic, and germinate faster upon exposure to the bile salt taurocholate *in vitro*. Our study establishes a direct link between the spore surface and colonization. Decrease hydrophobicity, faster germination and an altered interaction of the spore with the colonic mucosa due to misassemble of its surface layers, may explain the strong colonization defect of *sp17* mutant.

MATERIALS AND METHODS

Strains and general techniques. Bacterial strains and their relevant properties are listed in Table 5.1. The *E. coli* strain DH5α (Bethesda Research laboratories) was used for molecular cloning. *E. coli* strain BL21(DE3) (Novagen) was used for the production of a Sp17-His_6 fusion (see below). Luria-Bertani medium was routinely used for growth and maintenance of *E. coli*. When indicated, ampicillin (100 μg/ml), chloramphenicol (15 μg /ml) or kanamycin (30 μg/ml) was added to the culture medium. The *C. difficile* strains used in this study are congenic derivatives of the wild-type strain 630Δ*erm* (Hussain *et al.*, 2005). *C. difficile* strains were grown anaerobically (5% H2, 15% CO2, 80% N2) at 37ºC in brain heart infusion (BHI) medium (Difco). When necessary, thiamphenicol (15 μg/ml) or erythromycin (5 μg/ml) was added to *C. difficile* cultures. Routine plasmid constructions were carried out using standard procedures.

Table 5.1. Bacterial Strains used in this study.

Strain	Relevant Properties	Origin
E.coli		
Dh5α		Invitrogen
BL21 (DE3)		Novagen
HB101 (RP4)		Laboratory stock
AHCD070	BL21 (DE3)/ pFT41	This work
AHCD076	HB101 (RP4)/ pMTL84121	"
AHCD113	HB101 (RP4)/ pFT52	"
AHCD127	HB101 (RP4)/ pFT57	"
B. subtilis		
BS49	CU2189::Tn*916*	P. Mullany
AHCD159	BS49 Tn*916*::pFT71	This work
C. difficile		
630Δ*erm*	*C. difficile* 630Δ*erm*	Hussain *et al.*, 2005
AHCD532	630Δ*erm sigE*::intron *ermB*	Chapter 3
AHCD533	630Δ*erm sigF*::intron *ermB*	"
AHCD534	630Δ*erm sigG*::intron *ermB*	"
AHCD535	630Δ*erm sigK*::intron *ermB*	"
AHCD543	630Δ*erm* containing pMTL84121	"

Strain	Relevant Properties	Origin
630Δ*erm sp17*	630Δ*erm sp17*::intron *ermB*	Janoir *et al*., 2013
AHCD599	630Δ*erm* containing pFT52	This work
AHCD620	AHCD535 containing pFT52	"
AHCD652	630Δ*erm* containing pFT57	"
AHCD715	630Δ*erm sp17* containing pFT71	"

Bacterial strain and plasmid construction. To complement the *sp17 C. difficile* mutant in single copy, a 1117 bp region comprising the entire coding sequence of *sp17 (CD1581*; 483 bp length) and its expected promoter region (534 bp upstream of *sp17* gene) was amplified by PCR using primers Psp17 Fw/Sp17 Rev (Table A2). The generated PCR fragment was then cloned into the SalI and XhoI sites of pMTL84121, yielding pFT43 (Table A1). The same fragment was excised from pFT43 using SalI/HindIII restriction sites and cloned into the same sites of pSMB47, yielding pFT71 (Figure 5.1A). AHCD159 was created by the integration of pFT71 into the chromosomal Tn*916* locus of *B. subtilis* strain BS49. AHCD715 is a transconjugant from the mating of *B. subtilis* strain AHCD159 and *C. difficile sp17* mutant strain, resulting in the integration of the plasmid::Tn*916* fusion of the donor strain into the *C. difficile* chromosome (Figure 5.1A and B).

To construct a P_{sp17}-*SNAP*^{Cd} fusion, a 537 bp fragment encompassing the promoter region of *sp17* was PCR-amplified using genomic DNA from strain 630Δ*erm* and primer pairs Psp17 Fw/Psp17 Rev (Table A2). This fragment was inserted between the SalI and XhoI sites of pFT47 to create pFT52 (Table A1) (see also Chapter 2). To create a C-terminal Sp17-SNAPtag protein fusion, the promoter region and coding sequence of s*p17* were amplified from pFT43 using primers Psp17 Fw and sp17 linker Rev. This fragment was joined by SOE PCR to a 574 bp PCR fragment comprising the *SNAP*^{Cd} sequence amplified from pFT47 using primer pairs SNAP-sp17 linker Fw/SNAPtag-BamHI Rev. The resulting 1591 bp joint fragment was cloned between SalI and HindIII sites of pMTL84121 to yield pFT57. The latter carries the *sp17* coding sequence (from which the stop codon was removed)

linked to the *SNAP*^{Cd} coding sequence (which in turn lacks the start codon) by a sequence coding for a 9 amino acid linker (LGGGGSAAA) (see also Chapters 2 and 3, Materials and Methods). The absence of unwanted mutations was verified by sequencing the insert in all the plasmids. Plasmids pFT52 and pFT57 bearing *SNAP*^{Cd} fusions were introduced into *E. coli* HB101 (RP4) and then transferred to *C. difficile* 630Δ*erm* and *sigK* (AHCD535) strains by conjugation (Heap *et al.*, 2007)(Table 5.1).

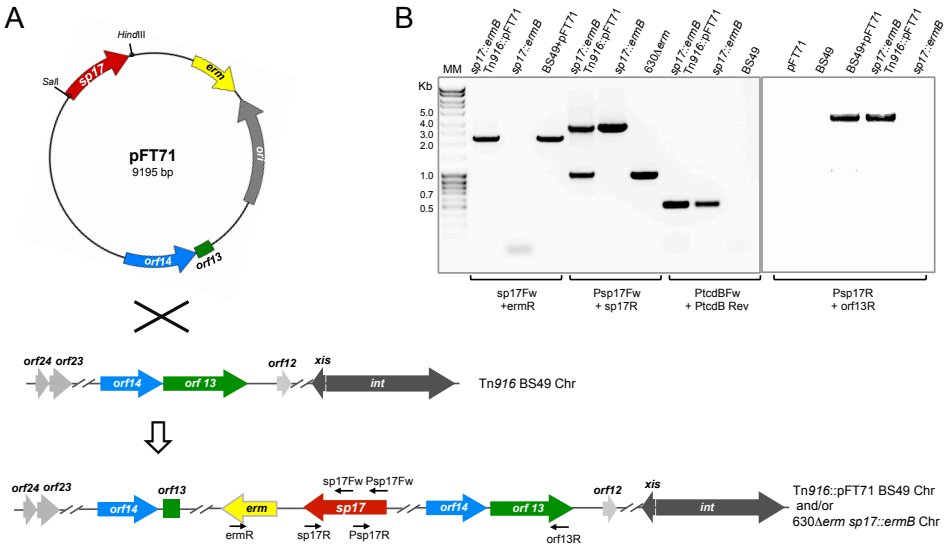


Figure 5.1. **Complementation of *sp17* mutant using the Tn*916* transposon**. **(A)** Schematic representation of the integration of plasmid pFT71 into Tn*916*. The main features of pFT71 plasmid (on top) are indicated, together with the position of the SalI/HindIII sites used to clone *sp17*. The plasmid integrates into Tn*916*, initially present in *B. subtilis* BS49 chromosome (Chr), via a single crossover homologous recombination event, represented by a cross, that occurs downstream of *orf14*. Tn*916* containing pFT71 was then transferred from BS49 to *C. difficile sp17::ermB* chromosome via bacterial conjugation (bottom). **(B)** Chromosomal DNA of *B. subtilis* BS49 transformants or of *C. difficile sp17::ermB* conjugants was screened by PCR using primer pairs sp17Fw/ermR to confirm the presence of pFT71, Psp17Fw/sp17R to confirm the presence of this second intact copy of *sp17* on *sp17::ermB* genome, and PtcdBFw/Rev primers to confirm that the transconjugants were *C.difficile*, and not *B.subtilis* (left panel). Integration of pFT71 in Tn*916* at the expected place, as depicted in A, was also confirmed by screening chromosomal DNA of *B. subtilis* BS49 transformants and of *C. difficile sp17::ermB* conjugants with primer pair Psp17R/orf13R (right panel). The position of the DNA size marker is indicated on the left.

RNA extraction, real-time quantitative RT-PCR analysis and RNA Seq. Total RNA from the strain 630Δ*erm* and the *sigF*, *sigE, sigG* and *sigK* mutants was extracted as described before (see Chapter 4, Materials and Methods). The RNA quality was determined using RNA 6000 Nano Reagents (Agilent). For quantitative RT-PCR experiments, 1 µg of total RNA was heated at 70°C for 10 min along with 1 µg of hexamer oligonucleotide primers $p(dN)_6$ (Roche). After slow cooling, cDNAs were synthesized for 2 h at 37°C with AMV Reverse Transcriptase (Promega), 20 mM dNTP mix and 40 U RNasin (Promega). The reverse transcriptase was inactivated by incubation at 85°C for 5 min. Real-time quantitative RT-PCR was performed twice in a 20 µl reaction volume containing 20 ng of cDNAs, 10 µl of FastStart SYBR Green Master mix (ROX, Roche) and 200 nM gene-specific primers in a AB7300 real-time PCR instrument (Applied Biosystems). The primers used for each marker were listed in Table A2. Amplification and detection were performed as previously described (Saujet *et al.*, 2011). In each sample, the quantity of cDNAs of a gene was normalized to the quantity of cDNAs of the DNApolIII gene. The relative transcript changes were calculated using the $2^{-\Delta\Delta Ct}$ method (Saujet *et al.*, 2011).

For RNA-seq experiment allowing transcriptional start mapping, total RNA was isolated from C. difficile 630Δ*erm* strain grown in TY medium either after 4 h and 10 h of growth or under starvation conditions that correspond to a 1 h resuspension of exponentially grown cells (6 h of growth) into PBS buffer for 1 h at 37°C (Soutourina *et al.*, 2013). The Tobacco Acid Pyrophosphatase (TAP)+/- library construction, high-throughput sequencing and data analysis were performed as described before (see Materials and methods and Chapter 4).

Antibody production. DNA fragment encoding the *sp17* gene was generated by PCR from *C. difficile* 630Δ*erm* genomic DNA using primers sp17-pET33b Fw/ sp17-pET33b Rev (Table A2). The resulting DNA fragment

was ligated into the NcoI and XhoI restriction sites of pET33b (Novagen) to produce a gene encoding a fusion protein with His6x-Tag at the C termini (Table A1). The construct was confirmed by DNA sequencing. The recombinant protein was overexpressed in BL21 (DE3) cells. Expression was induced by the addition of 1 mM IPTG at an OD_{600nm} of 0.5. Incubation was continued for 4h, and the cells were harvested by centrifugation at 4000 *g*, for 10 min, 4ºC. Cells were ressuspended in buffer containing 20mM phosphate, 1 mM PMSF, 10mM Imidazole, and lysed using a French pressure cell (18000 lb/in^2). The lysate was centrifuged for 30 min at 15000 *g* and the Sp17 recombinant protein, present mostly in the pellet fraction, was solubilized by treatment with 8M Urea for 30 min. The supernatant from the solubilized fraction containing the recombinant protein was then loaded onto a 1 ml Histrap column (Amersham Phamarcia Biotech). The bound protein was eluted with a discontinuous imidazole gradient and the fractions containing the desired purified protein were identified by SDS-PAGE. The polyclonal antibody was raised in rabbits immunized with three independent doses of 200 µg of purified recombinant protein (Eurogentec, France).

Spore production and purification. For spore production, 5 ml of BHI media was inoculated with an isolated colony of *C. difficile* and cultured overnight at 37ºC in anaerobic conditions. 100 ml of fresh BHI media was then inoculate 1:100 from the O.N. culture and incubated at 37ºC under anaerobic conditions for 7 days. Cells were collected by centrifugation at 4800xg, ressuspended in cold water and stored ON at 4ºC. Spores were then purified with a 20 to 50% Gastrografin (Schering) step gradient, as previously described (Henriques *et al.*, 1995). The pellet was ressuspended in cold water, washed 10 times, and stored at 4ºC for further use.

Spore Germination. Density-gradient-purified spores were ressuspended in PBS to a final OD_{600nm} of 1 and heat activated for 10 min at

80ºC. A 10% solution of taurocholic acid (TA) (Sigma-Aldrich) in PBS was then added, for a final concentration of 5% in TA, to induce spore germination. Germination was followed by monitoring the decrease in the OD_{600nm} of the spore suspension, using a plate reader, at 37ºC with routine agitation, until no significant changes in OD_{600nm} were detected.

Hydrophobicity assay. Hydrophobicity of *C. difficile* spores was determined using the bacterial adherence to hydrocarbons (BATH) method**.** Briefly, density-gradient-purified spores were suspended in 1 ml sterile distilled water to a final OD_{440nm} of 0.5 and mixed with 100 µl of hexadecane (Sigma-Aldrich). Adherence to hydrocarbon was measured by quantifying the drop of OD_{440nm} of the aqueous solution. The percentage of decrease in OD_{440nm} for the aqueous solution was calculated as follows: $100(OD_0-OD_f)/OD_0$, where OD_0 and OD_f refer to the initial and final OD_{440nm}, respectively.

Spore coat extraction. *C. difficile* spores were resuspended in 50 ml of extraction buffer (0.125 mM Tris-HCl pH 6.8, 5% beta-mercaptoetanol, 2% SDS, 0.025% Bromophenol Blue, 0.5 mM DTT 5% glycerol) to a final OD_{600nm} of 4, and boiled as described previously (Costa *et al.*, 2006). Proteins present in the supernatant fraction were resolved by 15% SDS-PAGE, and visualized by Coomassie brilliant blue R-250 staining. For protein identification, protein bands were excised and digested with trypsin before analysis by matrix-assisted laser desorption ionization (MALDI), performed by the Mass Spectrometry Laboratory Services at Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa (http://www.itqb.unl.pt/Services/Analytical_Services/Mass_Spectrometry).

Whole cell extracts and immunoblot analysis. To obtain *C. difficile* whole cell extracts, 10 ml of culture were centrifuged. Cells were washed

with phosphate-buffered saline (PBS), ressuspended in 1 ml French press buffer (10 mM Tris pH 8.0, 10 mM $MgCl_2$, 0.5 mM EDTA, 0.2 mM NaCl, 10% Glycerol, 1 mM PMSF) and disrupted using a French pressure cell (18000 lb/in^2).

Whole cell extracts, or spore coat extracts obtained as described above, were resolved by SDS-15 % PAGE. Proteins were then transferred to a nitrocellulose membrane (BioRad), and immunoblot analysis was conducted as described previously (Serrano *et al.*, 2011). Anti-Sp17 antibody was used at a 1:10000 dilution. Anti-CspC was used at a 1:500 dilution, and Anti-SleC, anti-CspB and anti-CspC at a 1:3000 dilution (Fimlaid *et al*., 2013). A rabbit secondary antibody conjugated to horseradish peroxidase (Sigma) was used at dilution 1:5000. The immunoblots were developed with enhanced chemiluminescence reagents (Amersham Pharmacia Biotech).

Fluorescence microscopy and image analysis. For SNAP labeling, the TMR-Star substrate was added to cells in culture samples to a final concentration of 250 nM (New England Biolabs), and the mixture incubated for 30 min in the dark. Following labeling, the cells were collected by centrifugation (4000xg for 5 min), washed four times with 1 ml of PBS, and finally ressuspended in 0.5 ml of PBS containing or not the membrane dye Mitotracker Green (0.5 μ/ml) (Molecular probes, Invitrogen).

For immunofluorescence analysis, part of the purified spores was incubated at 37°C for 1 h in a 0.1% trypsin solution (Amimed), as a control. Spores were washed three times with PBS supplemented with 2% bovine serum albumin (Sigma) and then probed with primary anti-Sp17 antibody (1:1000) and Alexa Fluor 594 goat anti-rabbit IgG secondary antibody (Molecular probes, Invitrogen) (1:500).

For phase contrast and fluorescence microscopy, cells were mounted on 1.7% agarose coated glass slides and observed on a Leica DM6000B microscope equipped with a phase contrast Uplan F1 100x objective and

captured with a CCD Andor Ixon camera (Andor Technologies). Images were acquired and analysed using the Metamorph software suite (version 5.8; Universal Imaging) and adjusted and cropped using *ImageJ* (http://rsbweb.nih.gov/ij/).

Electron Microscopy. Thin sectioning and negative-staining transmission electron microscopy of late (72h) BHI cultures or of density gradient-purified *C. difficile* spores (see above) was conducted as described previously (Balomenou *et al.*, 2013; Sylvestre *et al.*, 2002).

Atomic force microscopy (AFM). A drop (2.5µl) of a cell suspension containing highly purified C. difficilespores (approximately $1x10^7$ spores/ml) was applied on top of glass slide. Spores were air dried and observed by AFM. The AFM measurements were performed using a Digital Instruments (DI) Dimension 3100 AFM and Nanoscope IIIA controller. Tapping modeTM was used with standard Olympus Si tapping mode tips (Bruker reference OTESP) with resonance frequencies of the order of 300 kHz. The quality of AFM images obtained with the latter tips was the best. Three different types of information can be obtained from the AFM scans: surface topography, amplitude and phase. The amplitude image shows the tip oscillation and can be interpreted as a spatial derivative of the topography image along the tip scan direction. The phase image illustrates the phase difference between the tip oscillation and the excitation signal, and is a measure of the tip/surface interaction.

Animal model. All animal experiments performed for this study were carried out using 6-to-9 week old germ-free C3H mice purchased from INRA, Orléans (France). Mice were housed in sterile isolators with *ad libitum* access to food and water. Before experiment, each animal was checked for germ-free by Gram staining of faeces and inoculating faeces in BHI and

incubating broth for 48 hours, either aerobically or anaerobically. Infection was performed by oral gavage with appropriate amount of *C. difficile* vegetative cells and colonization was followed by enumeration of *C. difficile* in faeces sampled throughout time.

***In vivo* kinetics of sporulation.** Six germ-free male C3H mice were challenged by oral route with $1x10^6$ CFUs of vegetative cells of the *C. difficile* strain 630 obtained from a 9 hour culture in TY media (early stationary phase). Spores were enumerated from this suspension of gavage, and the spore rate was less than 0.1% of vegetative cells. At 5, 8, 14, 24 and 48 hours, fresh faeces were sampled from each mouse and immediately frozen at -80°C, until analysis. Preliminary experiments were performed to ensure that freezing does not alter in a significant manner the results of counts either of vegetative cells or spores. After thawing, each faeces was suspended in phosphate buffer saline (PBS) as a 10 mg/ml suspension and was further serially diluted in PBS; vegetative cells were count by plating appropriate dilutions on BHI agar plates; spores were also enumerated after a heat shock treatment (60°C, 30 min) on same medium plates containing 0.1% of taurocholate sodium salt.

Ethics statement. All animal experiments were conducted on germ-free mice, with the agreement of central animal care facilities of University Paris-Sud, in strict accordance with good animal practise as defined by the relevant french guidelines.

RESULTS

CD1581 is required for efficient colonization in a mouse axenic model

Earlier work has shown that the *CD1581* gene was strongly induced *in vivo*, upon infection of axenic mice with vegetative cells of *C. difficile* 630Δ*erm* strain (Janoir *et al.*, 2013). To test whether CD1581 has a role in colonization, the *CD1581* of gene was disrupted in this strain using the ClosTron system (Heap *et al.*, 2007; Janoir *et al.*, 2013). The ability of the mutant to colonize the ceacum was investigated by a competitive assay in a dixenic mouse model with the original wild-type strain 630Δ*erm* and the isogenic mutant strain, enabling direct fitness comparison *in vivo* (Figure 5.2A). Total cell counts, measured 38 hours post-infection, were significantly lower for the *CD1581* mutant as compared to the wt strain 630Δ*erm* (Figure 5.2B). This indicated that CD1581 might have a role in the ability of *C. difficile* to colonize the mice caecum.

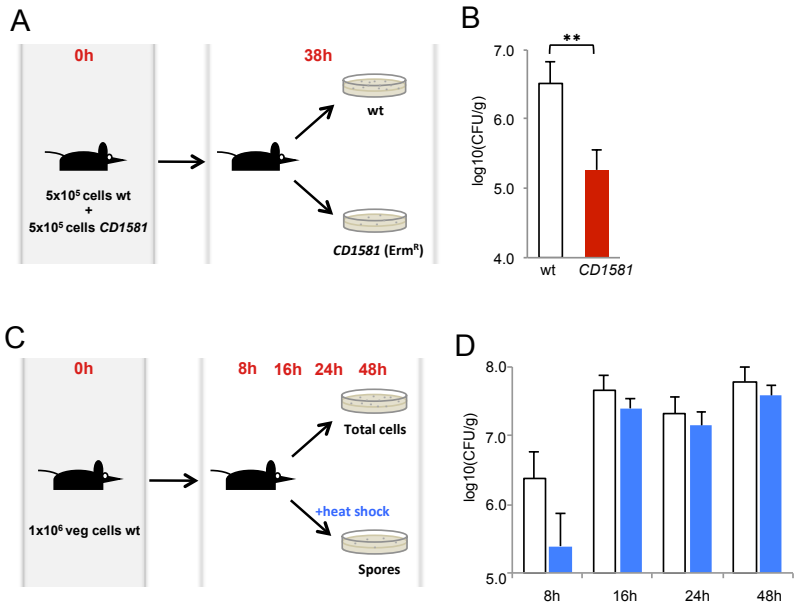


Figure 5.2. Competitive colonization assay and determination of the sporulation kinetics in germ-free mice. (A) Schematic representation of the competitive colonization assay between the parental wild-type (wt) strain and the *CD1581* mutant strain. Mice were orally challenged with $5x10^5$ cells of each strain, and the total number of bacteria from each

strain was determined by plating onto BHI agar plates, supplemented or not with erythromycin. The two strains can be distinguished due to the erythromycin resistance (Erm^R) marker carried by the mutant, but absent in the wild type. **(B)** Total bacterial counts in the faeces determined for the wt and the *CD1581* mutant 38 hours post infection. Statistically significant differences compared to the wt are indicated by asterisks (**, $P < 0.005$). **(C)** Schematic representation of the sporulation kinetics assay. Mice were orally challenged with $1x10^6$ cells of the wt strain. Total cell and spore titers were determined by plating directly on BHI agar plates supplemented with taurocholate (total cell counts) or by submitting the suspensions to a heat shock (spore counts), followed by plating. **(D)** Total cell (white bars) or spore (blue bars) titers in the faeces, assessed at 8, 16, 24 and 48 hours post-infection.

Interestingly, two studies have reported that the product of the *CD1581* gene is spore associated and is most likely part of the bacterial spore outer layers (Lawley *et al.*, 2009; Abhyankar *et al.*, 2013). Although spore formation has been shown to be asynchronous and inefficient *in vitro*, it was not known whether the kinetics of spore formation exhibits a similar pattern *in vivo*. To address this point, germ-free mice were infected with a single dose of *C. difficile* 630Δ*erm* vegetative cells, and the total cell and spore titers were monitored during colonization (Figure 5.2C). We observed that sporulation is surprisingly fast *in vivo*, with the first spores detected 8 hours after gavage. At this time, the percentage of sporulation *in vivo* is higher (10%; Figure 5.2D) than the percentage obtained *in vitro* at 48h of growth (8%; see Chapter 3). The spore titer drastically increases from hour 8 to hour 16h (from $2.5x10^5$ to $2.5x10^7$ spores/g faeces), and then remains constant (Figure 5.2D). Overall, these findings strongly suggest that *CD1581* is expressed during sporulation and is an important component of the spore surface. However, the precise localization of this protein within the outer layers that compose the spore is not yet known.

CD1581 is expressed in the mother cell chamber of developing cells, under the control of regulatory protein σ^K

In the sporeforming organisms of the *Bacillus* genus that have been examined in detail, all of the proteins that compose the spore surface structures are produced in the mother cell, under the control of σ^E and σ^K

(Henriques and Moran, 2007; McKenney *et al.*, 2012). We have shown that σ^K is essential for the assembly of the *C. difficile* spore coat layers (see Chapters 3 and 4) (Pereira *et al*., 2013; Saujet *et al*., 2013). Moreover, in Chapter 3 we have constructed and characterized *C. difficile* mutants for the four cell type-specific sigma factors that control gene expression during sporulation (Pereira *et al.*, 2013). Genome wide transcriptional analysis of these mutants was also performed (Chapter 4), allowing the identification of genes under the control of σ^F, σ^E, σ^G and σ^K (Saujet *et al.*, 2013). Among these was *CD1581*, which was assigned as a member of the σ^K regulon.

To further investigate the requirements for *CD1581* expression we mapped its promoter by RNAseq (Figure 5.3). We have identified a single possible start site for the *CD1581* mRNA, a "T" 45 nt upstream of the predicted translational start codon (Figure 5.3B). Sequences that conform well to the consensus 10 and -35 promoter elements recognized by σ^K in *C. difficile* (Saujet *et al.*, 2013) (see Chapter 4) were found upstream of the presumptive transcriptional start site (Figure 5.3B). In addition, we have used quantitative RT-PCR (qRT-PCR) to monitor the expression of *CD1581* during sporulation in SM medium of a wild type 630Δ*erm* strain as well as in congenic derivatives bearing insertional alleles of the *sigF*, *sigE*, *sigG*, and *sigK* genes. We have previously shown that this medium sports well sporulation and that activity of σ^F, σ^E, σ^G and σ^K can be detected and quantified between hours 14 and 24 of growth (see Chapters 3 and 4). Expression of *CD1581* was severely curtailed in *sigE* and *sigK* mutants (Figure 5.3C). These results suggest that *CD1581* is under the dual control of σ^E and σ^K or, most likely, that σ^E is required for the expression and activity of σ^K (Pereira *et al*., 2013; Saujet *et al.*, 2013) (see Chapters 3 and 4). Using the SNAP-tag as a fluorescent reporter system for *C. difficile*, we have recently localized the activity of σ^K during spore formation (Chapter 3) (Pereira *et al.*, 2013). σ^K is active in the mother cell compartment and its main period of activity encompasses the window from engulfment completion until spore

release to the environment (Chapter 3) (Pereira *et al.*, 2013). In line with these results, the fluorescence signal from a transcriptional fusion of *CD1581* promoter to *SNAP*^{Cd} was detected in the mother cell, at late stages in sporulation (Figure 5.3D). However, and unlike *cotE*, another σ^K target gene whose expression is detected in cells that have completed engulfment but have not yet developed spore refractility, *CD1581* expression was never detected at these earlier stages, suggesting that *CD1581* might be part of a second, late, wave of σ^K-directed gene expression. Consistent with this view, we have identified at least one other gene belonging to the σ^K regulon, *cotJC*, whose expression parallels that of *CD1581* (not shown). Thus, *CD1581* is expressed in the mother cell chamber of sporulating cells, at a late stage in sporulation, under the control of σ^K.

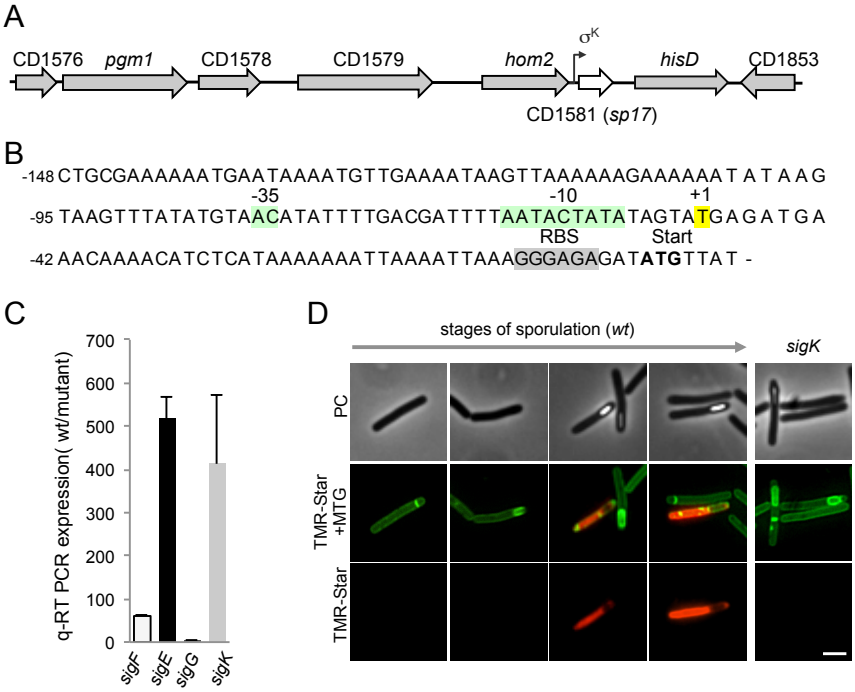


Figure 5.3. Expression of *sp17* is under the control of the sporulation factor σ^K. (A) The *sp17* (*CD1581*) region of the 630Δ*erm* chromosome. The position of a σ^K-type promoter driving expression of the *sp17* gene is indicated by the broken arrow. The region represented encompasses 11 kb of DNA. **(B)** Mapping of the *sp17* transcription start site. The figure represents the transcriptional start site (+1, yellow) and the -10 and -35 promoter elements (green) that match the consensus for σ^k binding as defined for *B. subtilis*. The ribosome binding site (RBS, shaded in grey) and the *sp17* start codon are also indicated. **(C)** Effect of mutations in the indicated genes on the level of the *sp17* transcript

as measured by qRT-PCR. The effect of the mutations is represented as the ratio of the wt/mutant. **(D)** Microscopy analysis of *C. difficile* sporulation cells carrying a fusion of the *sp17* promoter to $SNAP^{Cd}$. The cells were collected after 24h of growth in SM medium, stained with TMR-Star SNAP substrate and the membrane dye MTG, and examined by phase contrast (PC) and fluorescence microscopy. Scale bar: 1μm.

Sp17 is an abundant component of the *C. difficile* spore coat

Next, we wanted to examine whether CD1581 could be detected among the collection of proteins that can be extracted from highly purified spores produced by the wild type strain 630Δ*erm*. Different growth conditions and purification methods have been described in the literature to obtain *C. difficile* spores (Lawley *et al.*, 2009; Permpoonpattana *et al.*, 2011; Paredes-Sabja and Sarker, 2012; Heeg *et al.*, 2012; Adams *et al.*, 2013; Francis *et al.*, 2013). We found that growth in BHI liquid medium for 7 days reproducibly yields a high titer of spores. Under these culturing conditions, spore production is followed by massive vegetative cell lysis, which facilitates spore recovery. The spores were collected by centrifugation, washed extensively with water, and finally applied to a density gradient of metrizoic acid (see Materials and Methods). Using this method we were able to obtain a spore suspension essentially free of vegetative cells and debris (≥99% purity). Spore coat proteins were extracted from purified spores in the presence of SDS and the reducing agents DTT and β-mercaptoethanol, and the extracted proteins analysed by SDS-PAGE. These extraction conditions have been extensively used to analyse the composition of the coat layers of *B. subtilis*, *B. cereus* or *B. anthracis* (Henriques and Moran, 2000; Henriques and Moran, 2007). Importantly, the spores remain phase-bright following the extraction, indicating that no major disruption of the cortex and core compartments has occurred. Under these conditions, a collection of about 25 to 30 polypeptides was extracted from wild type spores and revealed by Coomassie blue staining (Figure 5.4A). The pattern obtained differs slightly from other recent reports in the number of bands and their relative abundance, which may reflect different culturing conditions,

including the time at which spores were harvested, and extraction conditions (Escobar-Cortes *et al.*, 2013; Abhyankar *et al.*, 2013; Lawley *et al.*, 2009; Paredes-Sabja and Sarker, 2012; Permpoonpattana *et al.*, 2011). Nevertheless, eight proteins were identified by mass spectrometry analysis, which were previously found to be spore surface components or part of the spore proteome (Abhyankar *et al.*, 2013; Permpoonpattana *et al.*, 2011; Lawley *et al.*, 2009) (Figure 5.4A), including NifJ (predicted molecular mass of 128 kDa, coded for by CD2682), AdhE, a bifunctional acetaldehyde CoA/alcohol dehydrogenase (95 kDa/CD2966), CdeC (42 kDa/CD1067; detected in both the 80 and 200 kDa region of the gel), CotE (82 kDa/CD1433; detected in the 40 kDa region of the gel), CotB (35 kDa/CD1511), CotA (34 kDa, CD1613), CotJC1/CotCB (21 kDa/ CD0598), and rubrerythrin (25 kDa/CD0825). Among the collection of extractable proteins is an abundant species with an apparent molecular weight of 17 kDa, identified by mass spectrometry as the product of *CD1581* (predicted molecular weight of 19 kDa) (Figure 5.4A, yellow arrows), and named Sp17 (for Spore protein with 17 kDa). To confirm the identity of this species, and to gain insight into its role on spore coat assembly, we analysed the composition of the coat layers in spores of a *sp17::ermB* (CD1581) insertional mutant (Janoir *et al.*, 2013). Sp17 was absent from the collection of proteins extracted from the mutant spores (Figure 5.4A). Also, the extractability of several other species increased in the mutant, as is the case of CotA and CotJC1 (Figure 5.4A, red dots). Presumably, these proteins are in close proximity to Sp17, or otherwise require Sp17 for proper assembly onto the coat. Importantly, complementation of *sp17::ermB* in single copy with the Tn*916* transposon carrying a copy of the native *sp17* gene under the control of the σ^K promoter restored assembly of the protein to the mutant (Figure 5.4A).

To further confirm the identity of Sp17, we overproduced an Sp17-His_6 fusion in *E. coli*, and partially purified the protein for the production of a

polyclonal antibody against Sp17. Immunoblot analysis of the wild type spore coat extracts, in comparison with the mutant, reveals that the antibody reacts with a 17 kDa species, as well as with species of 25, 37 and 75 kDa (Figure 5.4B). The anti-Sp17 antibody also cross-reacts with two additional species in the extracts from the mutant spores (Figure 5.4B). These signals may be due to the reaction of the antibody with other proteins (cysteine-rich, see below) that are known to be present at the spore surface. One of these proteins, CdeC (45 kDa), was recently found to be important for exosporium assembly (Barra-Carrasco *et al.*, 2013). Importantly, CdeC levels at the spore surface do not change upon disruption of *sp17* (Figure 5.4B, asterisk). It should be noticed that a band with the same mobility of Sp17 can also be seen in the gels of coat protein extracts of other studies but was not identified (Abhyankar *et al.*, 2013; Permpoonpattana *et al.*, 2011; Escobar-Cortes *et al.*, 2013).

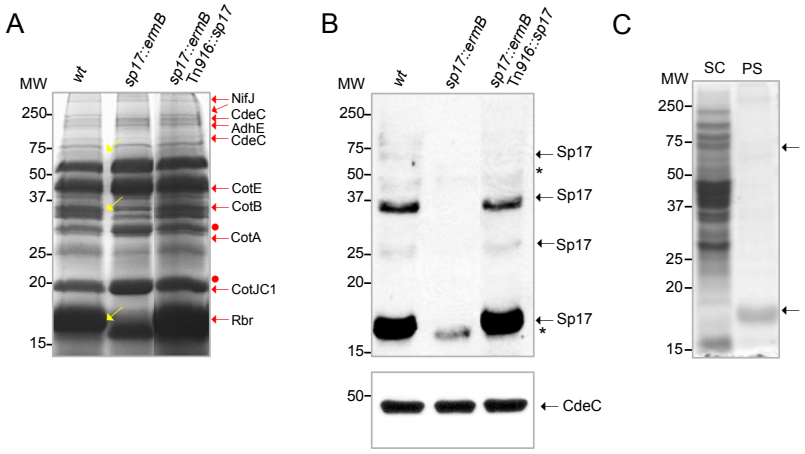


Figure 5.4. Sp17 is a spore coat protein. (A) Coomassie stained SDS-PAGE gel of the proteins extracted from highly purified spores of the wild type (wt) strain 630Δ*erm*, the *sp17::ermB* mutant, and the complemented *sp17::ermB* mutant strain carrying *sp17* a single copy in Tn*916* (Tn*916::sp17*). The yellow arrows indicate forms of the Sp17 protein. Red arrows indicate proteins that were identified by mass spectrometry (see text for details). Red dots refer to bands of proteins that are more extractable in the mutant. **(B)** Immunoblot analysis of the spore proteins resolved by SDS-PAGE in (A) with anti-Sp17 or anti-CdeC antibodies. Arrows indicate forms of the Sp17 protein and the asterisks indicate a cross-reactive species. **(C)** Proteins in extracts prepared from late sporulating cultures of the wt strain before (SC) or after spore enrichment by density gradient centrifugation (PS) were resolved by SDS-PAGE and the gel stained with Coomassie. Arrows point to forms of the Sp17 protein. The position of molecular weight (MW) markers (in kDa) is indicated.

Lastly, Sp17 was not detected in extracts of sporulating cells, prior to spore purification, even though more protein was electrophoretically resolved, indicating its enrichment in purified spores (Figure 5.4C).

Sp17 undergoes multimerization at the spore surface

In spores of the *sp17::ermB* insertional mutant, species of about 25, 35 and 75 kDa were reduced or absent, in addition to the 17 kDa product of the gene. These bands could correspond to multimeric forms of Sp17, or alternatively, to proteins whose assembly and/or maintenance at the spore surface require Sp17. During purification of the Sp17-His_6 protein for antibody production, we noticed that the formation of high molecular weight species running at about 60 and 150 kDa in a non-reducing SDS-PAGE gel (Figure 5.5A). Under reducing conditions, this species run as a single band of the expected size (17 kDa), suggesting that disulphide bonds are involved in the formation of homo-multimeric forms of Sp17 (Figure 5.5A). An analysis of the primary and secondary structures of Sp17 is in line with this idea. Sp17 appears to be an extended coiled-coil protein containing 14 cysteine residues, 7 of which clustered close to its C-terminal end (Figure 5.5B and C).

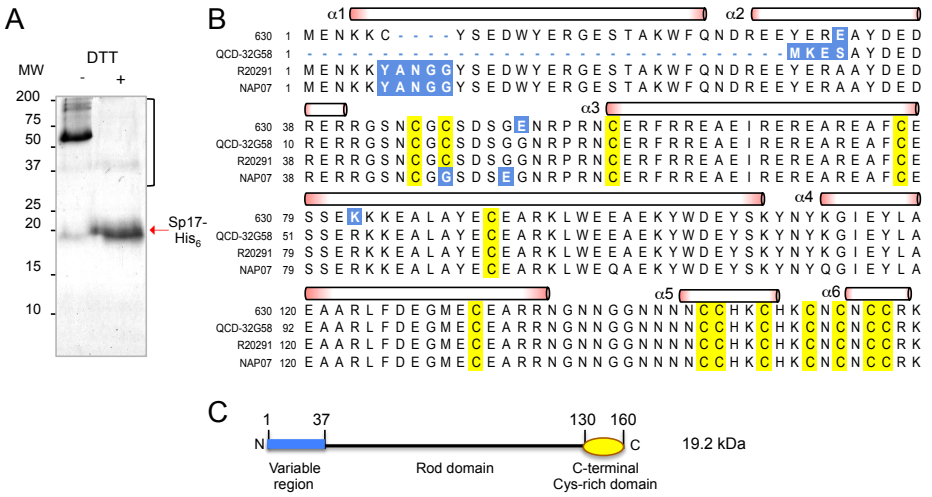


Figure 5.5. **Sp17 is a cysteine rich protein and undergoes extensive multimerization *in vitro*. (A)** An Sp17-His_6 fusion was overproduced in *E. coli* and the purified protein was resolved by SDS-PAGE in the presence or in the absence of DTT (10 mM). Coomassie

stained bands present in the brace area correspond to multimeric forms of Sp17. The red arrow indicates the position of monomeric Sp17. The position of molecular weight (MW) markers (in kDa) is shown on the left side of the panel. **(B)** Aminoacid sequence alignment of the Sp17 found in the genomes of four different *C. difficile* strains (www.ncbi.nhih.gov). The N-terminal variable protein sequence is highlighted in blue, and conserved cysteine (C) residues are shaded in yellow. Six predicted α helices are represented in red (www.psipred.org). **(C)** Representation of the different regions predicted for Sp17.

Therefore, we posit that the 17, 35 and 75 species represent the monomer, a dimer and possibly a tetramer of Sp17, respectively. It is possible that the extraction conditions do not allow the complete reduction of the cysteine cross-linked Sp17 multimers. For example, extraction of the cysteine-rich crust proteins of *B. subtilis* requires very high concentrations of β-mercaptoethanol (β-ME) (Zhang *et al.*, 1993). Alternatively, these species are hold together by other types of covalent or non-covalent interactions. We infer that Sp17 undergoes multimerization at the spore surface, at least in part through the formation of disulphide bonds.

Sp17 is surface exposed

Transmission electron microscopy studies indicate that the surface of *C. difficile* spores is organized into an inner lamellar structure and an outer electrodense layer (Lawley *et al.*, 2009; Permpoonpattana *et al.*, 2011, Paredes-Sabja *et al.*, 2012). In addition, the spore may be surrounded by and exosporium, although this structure is not always easily recognizable (Lawley *et al.*, 2009, Permpoonpattana *et al.*, 2011; Paredes-Sabja *et al.*, 2012; see also below). To gain insight into the localization of Sp17 within the coat layers, we treated purified spores of the wild-type strain 630Δ*erm* with trypsin. Studies in *B. subtilis* have shown that the spore coat layers efficiently exclude molecules greater that 14 kDa (Driks, 1999; Henriques and Moran, 2000). Following digestion, the composition of the coat layers was analysed. The results in Figure 5.6A show that the representation of Sp17 at the spore surface is greatly reduced following trypsin digestion. We take this observation as evidence that Sp17 is exposed at the spore surface. In addition, the abundance of several other coat proteins was also reduced,

suggesting that these proteins are also at least partially associated with the more external layers of the coat (Figure 5.6A).

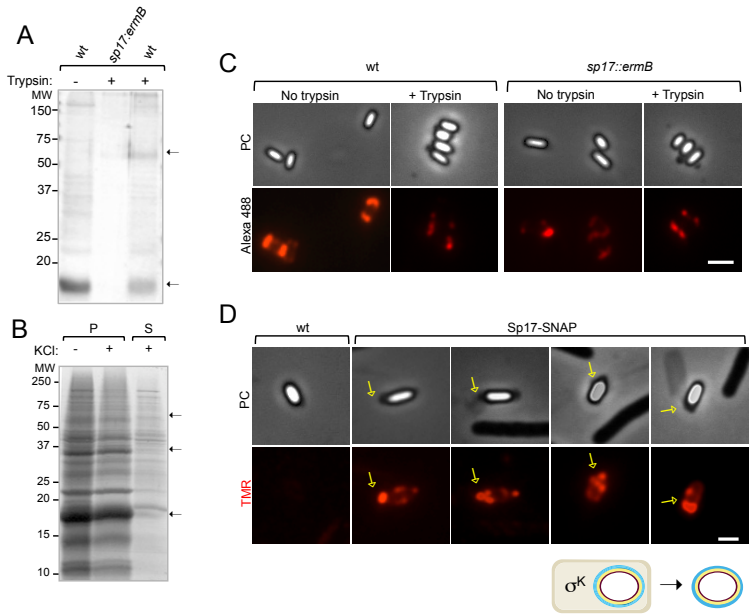


Figure 5.6. Sp17 is a spore protein and is exposed at the spore surface. (A) Coomassie stained gel of proteins extracted from wt or *sp17* mutant spores un-treated (-) or treated (+) with trypsin and resolved by SDS-PAGE. **(B)** Spore coat proteins were extracted from pellet (P) of spores not washed (-) or washed (+) with 1 M KCl, and resolved by SDS-PAGE. The supernatant fraction (S) of the spore wash was also analysed. The gel was stained with Coomassie. The position of molecular weight (MW) markers (in kDa) is shown on the left side of panels in A and B. Also in panels A and B, arrows indicate the position of forms of Sp17. **(C)** Localization of Sp17 at the spore surface of *C.difficile* spores by Immunofluorescence. Purified spores were probed with anti-Sp17 primary antibody and Alexa Fluor 488 secondary antibody. Spores from which Sp17 was removed by trypsin treatment, as well as *sp17* mutant spores, were used as a control. **(D)** Localization of Sp17 at the spore surface of *C. difficile* spores using the SNAP-tag. Spores present in a late sporulation culture of the wt strain alone or carrying an Sp17-SNAP fusion were labeled with the TMR-Star SNAP substrate and examined by microscopy. A schematic representation of σ^K-dependent Sp17 accumulation and localization around the spore is show below the main panel. Scale bar in C and D: 1μm.

Some proteins are known that accumulate in the culture medium and that can bind non-specifically, often through electrostatic interactions, with spores following their released from the mother cell (Henriques and Moran, 2007). These proteins can be detached from spores by washing with a high salt solution (Serrano *et al.*, 1999). Only a negligible part of Sp17 was removed from the spore surface by washing with a 1M solution of potassium

chloride (Figure 5.6B), strongly suggesting that Sp17 is a bona fide coat component, and not a protein that serendipitously adsorbs to the spore surface. Together, and in consonance with the σ^K-directed expression of its encoding gene, we conclude that the Sp17 protein is an abundant component of the spore surface layers.

As an independent way of assessing the localization of Sp17, we used immunofluorescence microscopy. The polyclonal anti-Sp17 antibody was bound to wild type and *sp17::ermB* mutant spores, and the antigen-antibody complex detected with a secondary anti-rabbit fluorescent antibody. We found strong decoration of 630Δ*erm* spores by the anti-Sp17 antibody, in a pattern that mimics deposition of the spore coat material (Figure 5.6C). Labeling is strongly decreased in *sp17* mutant spores and even more strongly in wild type spores treated with trypsin (Figure 5.6C). Residual labeling of the mutant spores is most likely due to the cross-reaction of the antibody with other spore cysteine rich exposed proteins as noted above, and which are also susceptible to trypsin (Figure 5.6C). We have also taken advantage of the SNAP-tag as a fluorescent reporter to localize proteins, and have analysed the localization of a translational Sp17-SNAP fusion. In agreement with the immunofluorescence analysis, labeling of spores of a wild type strain carrying pFT57 (Sp17-SNAP in a multicopy plasmid) with the TMR-Star SNAP substrate revealed that Sp17 localizes at the spore surface (Figure 5.6D). Preferential accumulation of Sp17 at one or both of the spore poles can be observed, coincidently with the appearance of an appendix-like spore structure, visible by phase contrast (Figure 5.6D; yellow arrows). SNAP-TMR signal is specific, since no fluorescence is detected in spores of the wild type strain labeled with the TMR fluorescent substrate (Figure 5.6D).

The data is consistent with the view that Sp17 is a component of the spore coat layers, and at least partially surface-exposed.

An *sp17* mutant fails to assemble the spore outer coat

The results described in the preceding sections show that Sp17 is an abundant, surface-exposed spore coat protein. The altered pattern of coat proteins extracted from spores when *sp17* is disrupted, suggest that Sp17 might be important for proper assembly of the spore coat. We next wanted to examine whether the insertional allele impacted the ultrastructure of the spore surface layers using TEM. In agreement with previous descriptions (Lawley *et al.,* 2009; Permpoonpattana *et al*., 2011; Paredes-Sabja and Sarker, 2012), *C. difficile* spores showed an inner lamellar coat layer closely apposed to the spore cortex peptidoglycan, and a thick electrodense outer coat (Figure 5.7A). The inner coat consisted of 5-8 lamellae, and has an average thickness of about 42 ± 10 nm. The outer coat showed a more amorphous organization, and an average thickness of 36 ± 14 nm (Figure 5.7A). Strikingly, in *sp17* mutant spores the outer electrodense coat layer was drastically reduced, and in most cases absent (Figure 5.7B). In contrast, the inner lamellar layer was maintained around the spore, although occasionally, in some sections around the periphery of the spore, it was also also reduced. Also occasionally, disorganized material, most likely remnants of outer coat material, was seen loosely attached to the exposed edge of the inner coat lamellae. We have previously reported that under the culturing and spore purification conditions that we used, an exosporium layer is rarely seen (Pereira *et al*., 2013; Chapter 3). This structure may be extremely labile and/or alter the density of the spores (Pereira *et al*., 2013). While we cannot presently ascertain whether Sp17 is also required for the assembly of the spore exosporium, it is clear that disrupting the *sp17* gene causes a severe block in the assembly of the outer layer of the spore coat, in agreement with the conclusion that most of the Sp17 protein is accessible to trypsin (above). Furthermore, because disruption of *sp17* does not result in the absence of other abundant proteins from spore coat extracts, we infer that Sp17 is the main component of the spore outer coat layer.

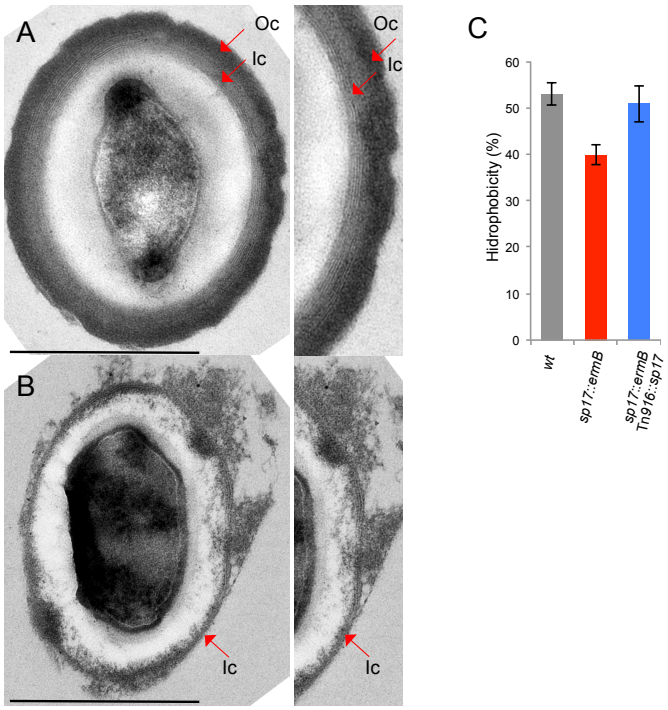


Figure 5.7. Sp17 is a key determinant for the assembly of the spore outer coat layer. Wild type **(A)** and *sp17* mutant **(B)** spores, were purified by density-gradient centrifugation and observed by TEM. The panels on the right show a magnification of a region of the images on the left. Scale bar: 1μm. **(C)** *sp17* mutant spores are less hydrophobic than wt spores. Hydrophobicity of purified spores of the wild type, *sp17* mutant and the *sp17* mutant complemented strains was determined as described in Materials and methods. Error bars are from 3 independent experiments.

The suppression of outer coat assembly caused by the *sp17* mutation is reminiscent of the phenotype of *cotE* mutants of *B. subtilis*. CotE functions as a hub protein for the assembly of the outer coat and crust layers (Zheng *et al.*, 1988; McKenney *et al.*, 2011). Moreover, spores of a *cotE* deletion mutant are extremely sensitive to cortex lytic enzymes such as lysozyme (Zheng *et al.*, 1988; Costa *et al.,* 2008). We reasoned that the impaired assembly of the outer coat could likewise render spores of the *sp17* mutant sensitive to lysozyme. However we do not detect any significant change in the ability of *sp17::ermB* spores to resist to lysozyme. Also, no differences in heat resistance were observed for *sp17::ermB* spores when compared to the wt.

Hydrophobic interactions have been shown to be important for different aspects of pathogenesis, either directly, like the adherence to host epithelial cells (Andersson *et al.*, 1998; Paredes-Sabja and Sarker, 2012; Bozue *et al.*, 2007), or indirectly, like the adherence to surfaces, making decontamination difficult (Joshi *et al.*, 2012). In the pathogen *B. anthracis*, BclA, a major component of the spore exosporium, affects spore hydrophobicity and contributes to increased adherence of the spore to other types of cells beyond macrophages (Brahmbhatt *et al.*, 2007; Bozue *et al.*, 2007). In *C.difficile* it has also been shown that the spore surface layers, such as the exosporium layer, significantly contribute to the *C. difficile* spore hydrophobicity (Escobar-Cortés *et al.*, 2013). Since Sp17 is a determinant for the assembly of the spore outer coat, we wanted to determine the impact of *sp17* disruption on spore hydrophobicity. With that purpose, we used a BATH assay (Paredes-Sabja and Sarker, 2012). The results show that 39% of *sp17::ermB* spores were present in the hydrocarbon organic phase, contrasting with 52% of the wt spores (Figure 5.7C). Complementation of the *sp17* mutant with a single copy of *sp17* in *trans* was able to recover the phenotype (50% of spores in the organic phase for the strain *sp17::ermB* Tn*916*::*sp17*) (Figure 5.7C). Our results show that Sp17 is required for proper assembly of the spore outer coat layers, and either alone or in combination with other spore surface components, contributes to spore hydrophobicity.

Sp17 is required for the assembly of a spore polar appendage

About 30% of the purified spores observed by optical microscopy of *C. difficile* 630Δ*erm* show a polar appendage (Figure 5.8A). While the spore stained brightly with the lipophilic dye FM4-64 (see also Chapter 3), the appendage, clearly visible under phase contrast optics, remained unstained (Figure 5.8A). Among the spores for which a polar appendage was observed, 22% showed a prominent, regular and partially refractive structure, whose total spore length as measured by phase contrast represents more than, or equals, 1.2 of the length of the spore as measure by FM4-64 staining. In the

remaining 7% of spores, the appendage consisted of a phase-dark structure seemingly disorganized (Figure 5.8B). Strikingly, organized, partially refractive appendages were seen in only 4% of *sp17* spores, but these were shorter in comparison with spores of the wild type. Instead, in most of the *sp17* spores (26%) showing an appendage, this structure was of the phase-dark, disorganized type (Figure 5.8A and B). The presence of organized spore appendix is restored in a sp17 mutant strain complemented with Tn*916::sp17* (18% of type A and 11% of type-B-like appendix) (Figure 5.8B). To extend these observations, density-gradient purified spores were observed by TEM (Figure 5.8C). In wild type spores, a structure was seen on one of the spore poles that appeared to result from massive accumulation of coat material at this location (Figure 5.8C). This structure was absent or was much reduced in spores of the *sp17* mutant (Figure 5.8C). Thus, Sp17 contributes to formation of a spore polar appendage.

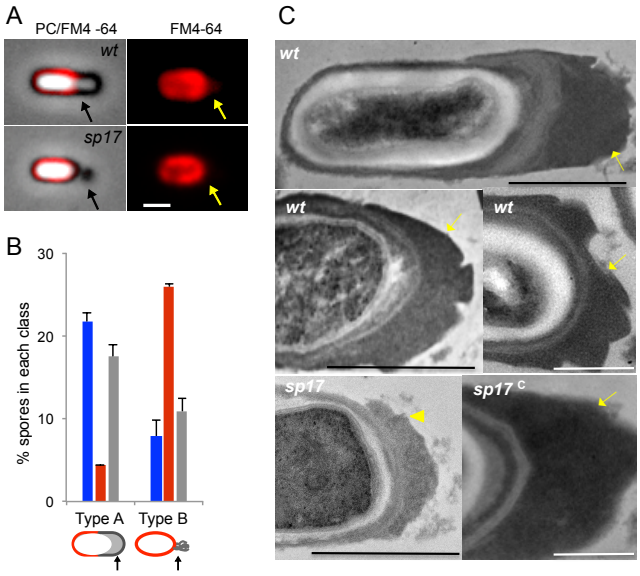


Figure 5.8. Sp17 controls the assembly of a spore polar appendage. (A) Phase contrast (PC) and fluorescence microscopy images of wild type and *sp17* mutant spores following staining with FM4-64. The arrow points to a polar structure that is partially refractile. In both spores, this structure does not stain with FM4-64. **(B)** Quantification of the types of polar structure seen for wild type (grey), *sp17::ermB* (red) and *sp17::ermB* complemented complemented in single copy (*sp17::ermB Tn916::sp17*; blue) spores. A total of 150 spores were scored; error bars are from 3 independent experiments. **(C)** TEM analysis of density-gradient purified spores of the wild type, *sp17*, and complemented

strain. Arrows point to the polar appendages (Type A) observed. Arrowheads point to the unstructured appendix (Type B) observed in *sp17* mutant spores. Scale bar, 1 μm.

We further complemented our analysis of the *C. difficile* spore surface and appendix structures using Atomic Force Microscopy. AFM has demonstrated to be a powerful tool for revealing the architectural topography of *Bacillus* and *Clostridium* spores (Plomp *et al.*, 2005; Plomp *et al.*, 2007; Ghosh *et al.*, 2008). Analysis of *C. difficile* spores by AFM revealed that the spore surface of wild type *C. difficile* spores is quite smooth (Figure 5.9). Spores showing a Type-A appendage-like structure could be observed. This appendage-like structure is the extension of the spore, and seems to be composed of the same coat material that surrounds the remainder spore (Figure 5.9). In contrast, the surface of *sp17* mutant spores often reveals the presence of bumps. Spores that show an unstructured deposition of coat material at one end are frequently observed, as is the presence of detached material (Figure 5.9). Interestingly, the AFM amplitude images also revealed that a less rigid zone is found along the longitudinal axis of the spore, in both wild type and *sp17* mutant spores (Figure 5.9, black arrows).

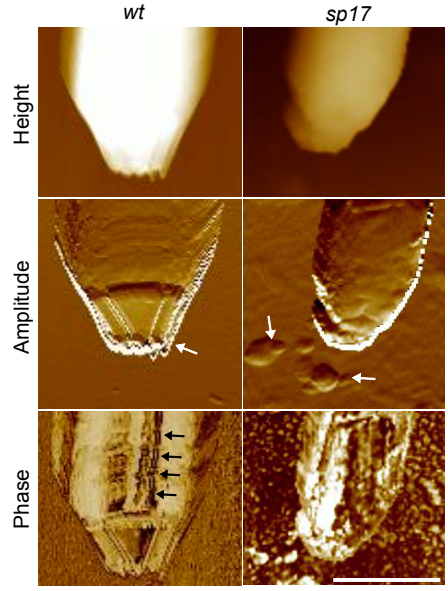


Figure 5.9. Atomic Force Microscopy (AFM) of *C. difficile* spores. Spores of the 630Δ*erm* wild type strain or its congenic *sp17* mutant derivative were air dried on a glass slide and observed by AFM. Height, Amplitude and Phase images are shown for a single spore of each strain. In the amplitude panel, white arrows point to the polar, structured appendix (Type A) frequently observed in wt spores, or for the more unstructured polar region (Type B appendage) characteristic of the mutant. Note the presence of material loosely attached to the polar region of *sp17* spores. AFM phase images of wild type spores also reveal the presence of a less rigid zone along the longitudinal axis of the spore (black arrows). Scale bar, 1 μm.

Germination is altered in *sp17* mutant spores

It takes about 48-72 h for *C. difficile* to attain the maximum spore titer, but efficient spore purification is only achieved with 7 days-old cultures (Pereira *et al*., 2013; Chapter 3). This is due in part because during this period extensive lysis of sporulating cells takes place, permitting the release of spores from the mother cell and their subsequent recovery by density gradient centrifugation (see Material and Methods). Moreover, extensive lysis of non-sporulating cells also takes place (Pereira *et al*., 2013; Chapter 3). Therefore, one possibility is that a fraction of the spores in the late BHI cultures initiate germination, and that formation of the polar appendage results from a re-organization of the spore surface layers during this process. If so, addition of a germinant such as the bile salt taurocholate (TA) (Wilson *et al*., 1982; Sorg and Sonenshein, 2008; Burns *et al*., 2010) to a suspension of density-gradient purified spores could increase the fraction of spores displaying the appendage. To test this possibility, we monitored spore germination by measuring the drop in the absorbance of a heat-activated spore suspension, following addition of the bile acid taurocholate (TA), a potent *C. difficile* spore germinant (Wilson *et al*., 1982; Sorg and Sonenshein, 2008). In parallel, samples were collected at various times for observation by phase contrast and electron microscopy. Interestingly, as measured by the drop in absorbance of the spore suspension, we found that *sp17* mutant spores germinate faster than wild type spores (Figure 5.10A). Germination was accompanied by phase darkening of the spores. Spore phase-darkening most likely results from re-hydration of the core compartment in turn resulting from spore cortex degradation, as established for *B. subtilis*. We note that spores of the complemented strain show germination at a rate more similar to the observed for the wild type spores (Figure 5.10A).

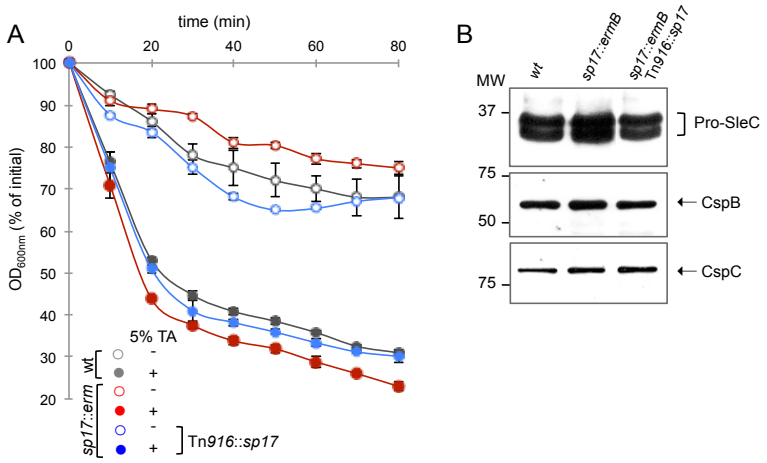


Figure 5.10. Sp17 influences spore germination. (A) Germination curves for spores of the wild type strain, *sp17::ermB*, and the *sp17::ermB Tn916::sp17* complemented strain. Germination of purified spores was followed by the decrease in OD_{600nm} over time. **(B)** Immunoblot analysis of coat proteins extracted from purified spores and resolved by SDS-PAGE. Anti-SleC, anti-CspB and anti-CspC antibodies were used. **(C)** Phase contrast (PC) and fluorescence microscopy images of wild type spores stained with FM4-64, taken immediately after germination in the presence of O_2 and 60 min after the induction of germination under anaerobic (no O_2) or aerobic conditions (+O_2). The percentage of spores with polar structures is shown on the right side of the panels, for each condition. Scale bar, 1 μm.

The spore protein CspC was recently identified as a bile acid germination receptor. Interaction with the germinant is able to induce conformational changes in CspC, which are then transmitted to CspB, a protein required to cleave the cortex hydrolase pro-SleC, leading to spore germination (Adams *et al.*, 2013; Francis *et al.*, 2013). Therefore, all of these proteins are determinants for spore germination in *C. difficile*. The fact that germination of *sp17* spores is altered led us to test whether the levels of any of these proteins were altered. Immunoblot analysis of spore coat extracts from wild type and *sp17* mutant showed no difference in the levels of any these proteins, suggesting that other factors, like increased accessibility of the germinant to its receptor might be the main reason for the faster rate of spore germination (Figure 5.10B).

The frequency of spores of the wild type, the *sp17* mutant, or the complemented strain, showing an appendage structure, did not appear to

change as the result of TA-triggered germination, under aerobic or anaerobic conditions, as judged by phase contrast and fluorescence microscopy following FM4-64 staining. To investigate whether subtle changes could be taking place during germination, we also monitored TA-triggered germination by TEM. No major morphological changes for wild type or *sp17* spores other that visible signs of cortex degradation were seen (not shown). Clearly, TA triggers cortex hydrolysis, a critical step in spore germination. However, it seems that the appendage is not formed or re-structured in response to TA-induced germination. In any event, the results show that Sp17 acts to delay spore germination.

DISCUSSION

Spores are known to play a key role in pathogenesis. Proteins located at the spore surface, like the *B. anthracis* BclA exosporium protein, mediate spore entry into lung epithelial cells by interaction with the receptor protein integrin $\alpha 2\beta 1$, in a process assisted by the complement component C1q (Xue *et al.*, 2011). In other spore-former organisms beyond bacteria, spores also play a crucial role in virulence. In the microsporidian opportunistic human pathogens *Encephalitozoon intestinalis* and *E. hellen*, two cysteine-rich proteins, SWP1 and SWP2, localize at the spore wall exospore and are responsible, at least in part, for the immune response of immunocompetent individuals to this pathogens (Hayman *et al.*, 2001; Polonais *et al.*, 2010). *C. difficile* spores are able to resist to decontamination processes, accumulate in the hospital environment and are ingested, initiating the infection cycle. These spores are able to pass the gastric barrier and reach the intestine where they must attach to the epithelial cells in order to achieve efficient colonization (Figure 5.11A). In the intestine these spores germinate, allowing the vegetative form of the organism to propagate. The vegetative

cells will produce the two main virulent factors of *C. difficile*, the TcdA and TcdB cytotoxins. The toxins will cause damage to the colonic mucosa and eventually severe diarrhea, which allows shedding of the spores and their transmission to a new host (Figure 5.11A). Spore surface proteins are thought to confer resistance to extreme environmental conditions, mediate spore adherence, and sense exterior signals and traduce favorable signals into spore germination, all crucial aspects of the *C. difficile* spore-mediated infection process.

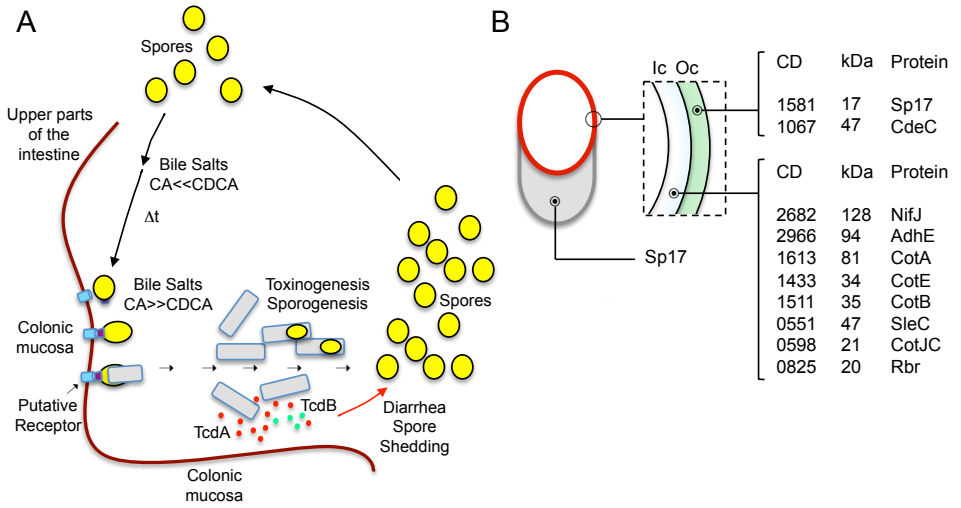


Figure 5.11. (A) Schematic representation of *C. difficile* infectious cycle with emphasis on the role of the spore. Infection starts with spore ingestion. The spore is primed to germinate in the upper, aerobic parts of the intestine, in contact with bile salts in the form of cholate, which are more abundant than deoxycholate salts (CA<<CDCA). However, spore germination may only be completed in the anaerobic colon (Δt), possibly in contact with a cognate receptor in the colonic mucosa. Following spore germination, the vegetative cells will grow, producing the two cytotoxic toxins and spores. When the colon is damaged leading to diarrhoea, the spores are shed from the organism, completing the cycle. **(B)** Localization of Sp17 and composition of the spore coat layers.

The *C. difficile* spore coat is composed of dozens of proteins and some of them have been the subjects of recent studies (Figure 5.11B). Here we have identified and characterized a novel spore coat protein, Sp17, which is the product of *sp17* gene (*CD1581*), expressed in the mother cell at later stages of sporulation under the control of the sporulation factor σ^K. We have

further demonstrated that Sp17 localizes at the cell surface of mature spores, and is required for proper assembly of the spore outer coat layer. Although we could not detect the presence of an intact exosporium in our analysis of the spore structure by TEM, we can foresee that the exosporium assembly is also impaired in the *sp17* mutant. CdeC is a spore coat component that was previously shown to be involved in the formation of the spore exosporium layer, being assembled at the spore surface as high molecular mass complexes (Barra-Carrasco et al, 2013). We could not detect the presence of these high molecular weight forms by immunoblot in any of the strains analysed. However, spores lacking Sp17 are less hydrophobic, and decreased hydrophobicity has been linked to the absence of the spore exosporium layer (Escobar-Cortes et al, 2013). Importantly, removal of the spore outer coat layers has been pointed as a cause for decreased ability of spores to adhere to intestinal epithelial cells (Paredes-Sabja and Sarker, 2012). Although we cannot conclude about the impact of *sp17* disruption on exosporium assembly, we speculate that *sp17* mutant spores lack both the outer coat and exosporium layers, and as a consequence exhibit defective adherence properties to the intestinal epithelial cells. This might explain in part the colonization defect observed for the *sp17* mutant (Janoir *et al*., 2013).

It has been suggested that upon spore germination, in contact with epithelial cells, the spore surface is re-organized and a long polar appendage is formed, that is required for adhesion and completion of the germination process (Panessa-Warren *et al*., 1997, Panessa-Warren *et al*., 2007). The only well characterized example of formation of a spore polar appendage is that of *C. taeniosporum* (Yolton *et al.*, 1972). The spore shows a polar structure formed by 12 large, flat ribbon-like appendages, at the edge of which a hair-like nap is visible (Walker *et al.*, 2007). One component of the appendage is a glycoprotein, termed GP85, which contains a collagen-like domain, and is related to BclA of *B. anthracis* (Walker *et al.*, 2007). A second BclA-like protein may also be involved. *C. difficile* has three BclA homologs, which have

been recently characterized. BclA1, like Sp17, participates in the morphogenesis of the spore surface layers, and is central to host colonization by *C. difficile* spores (Pizarro-Guajardo *et al.* 2014; Phetcharaburanin *et al*., 2014). Importantly, while the BclA proteins have orthologs in other spore formers, Sp17 is only found in *C. difficile* and is found in all sequenced genomes of the organism.

Nevertheless, Sp17 seems absolutely required for proper assembly of the spore polar appendix, which might further explains the reduced capacity of the *sp17* mutant to colonize mice (Janoir *et al.,* 2013). It has also been hypothesized that this appendage binds to a specific receptor in the colonic epithelium (Panessa-Warren *et al*., 1997; Panessa-Warren *et al*., 2007). Recently, *C. difficile* spores were cross-liked to the surface of epithelial cells, and at least four cellular proteins were found to cross-link to spores (Paredes-Sabja and Sarker, 2012). However, the identity of these proteins has not been reported, and we do not know at the present whether Sp17 might contribute to this interaction.

Lastly, another important aspect of spore-mediated colonization is the fact that while the spores may be primed to germinate upon exposure to bile salts in the upper sections of the intestine, they may only complete germination when they find appropriate adherence sites in the anaerobic colon (Figure 5.11A). *sp17* mutant spores germinate faster than the wild type. We hypothesize that faster germination could be catastrophic *in vivo* if it comes about in the aerobic sections of the intestine. Therefore, we propose that Sp17 works as a sensor, preventing spore germination in the absence of proper environment conditions.

AKNOWLEDGEMENTS

The author of this dissertation performed all the experiments described in this chapter, except the *in vivo* colonization assay in

competition, which was performed by Claire Janoir, Sandra Hoys and Bruno Dupuy; quantitative RT-PCR analysis and RNA-seq experiments, which were performed by Laure Saujet and Marc Monot; and Electron Microscopy which was performed by Evelyne Couture-Tosi. We thank Luís Viseu Melo for the help with the Atomic Force Microscopy. We would also like to thank Mónica Serrano, Isabel Gordo and Patrícia Beldade for helpful discussions, and Aimee Shen and Shona McBride for reagents. This work was supported by grant ERA-PTG/SAU/0002/2008 (ERA-NET PathoGenoMics) and PEst-OE/EQB/LA0004/2011 from the "Fundação para a Ciência e a Tecnologia" (FCT). F.P. (SFRH/BD/45459/08) was the recipient of doctoral fellowship from the FCT.

REFERENCES

Abhyankar, W., Hossain, A. H., Djajasaputra, A., Permpoonpattana, P., Ter Beek, A., Dekker, H. L., Cutting, S. M., Brul, S., de Koning, L. J. & de Koster, C. G. (2013). In pursuit of protein targets: proteomic characterization of bacterial spore outer layers. *J Proteome Res* **12**, 4507-21.

Adams, C. M., Eckenroth, B. E., Putnam, E. E., Doublie, S. & Shen, A. (2013). Structural and functional analysis of the CspB protease required for Clostridium spore germination. *PLoS Pathog* **9**, e1003165.

Andersson, A., Granum, P. E. & Ronner, U. (1998). The adhesion of Bacillus cereus spores to epithelial cells might be an additional virulence mechanism. *Int J Food Microbiol* **39**, 93-9.

Balomenou, S., Fouet, A., Tzanodaskalaki, M., Couture-Tosi, E., Bouriotis, V. & Boneca, I. G. (2013). Distinct functions of polysaccharide deacetylases in cell shape, neutral polysaccharide synthesis and virulence of Bacillus anthracis. *Mol Microbiol*.

Barra-Carrasco, J., Olguin-Araneda, V., Plaza-Garrido, A., Miranda-Cardenas, C., Cofre-Araneda, G., Pizarro-Guajardo, M., Sarker, M. R. & Paredes-Sabja, D. (2013). The Clostridium difficile exosporium cysteine (CdeC)-rich protein is required for exosporium morphogenesis and coat assembly. *J Bacteriol* **195**, 3863-75.

Bozue, J., Moody, K. L., Cote, C. K., Stiles, B. G., Friedlander, A. M., Welkos, S. L. & Hale, M. L. (2007). Bacillus anthracis spores of the bclA mutant exhibit increased

adherence to epithelial cells, fibroblasts, and endothelial cells but not to macrophages. *Infect Immun* **75**, 4498-505.

Brahmbhatt, T. N., Janes, B. K., Stibitz, E. S., Darnell, S. C., Sanz, P., Rasmussen, S. B. & O'Brien, A. D. (2007). Bacillus anthracis exosporium protein BclA affects spore germination, interaction with extracellular matrix proteins, and hydrophobicity. *Infect Immun* **75**, 5233-9.

Burns, D. A., Heeg, D., Cartman, S. T. & Minton, N. P. (2011). Reconsidering the sporulation characteristics of hypervirulent Clostridium difficile BI/NAP1/027. *PLoS One* **6**, e24894.

Carter, G. P., Rood, J. I. & Lyras, D. (2012). The role of toxin A and toxin B in the virulence of Clostridium difficile. *Trends Microbiol* **20**, 21-9.

Costa, T., Isidro, A. L., Moran, C. P., Jr. & Henriques, A. O. (2006). Interaction between coat morphogenetic proteins SafA and SpoVID. *J Bacteriol* **188**, 7731-41.

Deakin, L. J., Clare, S., Fagan, R. P., Dawson, L. F., Pickard, D. J., West, M. R., Wren, B. W., Fairweather, N. F., Dougan, G. & Lawley, T. D. (2012). The Clostridium difficile spo0A gene is a persistence and transmission factor. *Infect Immun* **80**, 2704-11.

Deneve, C., Janoir, C., Poilane, I., Fantinato, C. & Collignon, A. (2009). New trends in Clostridium difficile virulence and pathogenesis. *Int J Antimicrob Agents* **33 Suppl 1**, S24-8.

Driks, A. (1999). Bacillus subtilis spore coat. *Microbiol Mol Biol Rev* **63**, 1-20.

Escobar-Cortes, K., Barra-Carrasco, J. & Paredes-Sabja, D. (2013). Proteases and sonication specifically remove the exosporium layer of spores of Clostridium difficile strain 630. *J Microbiol Methods* **93**, 25-31.

Fimlaid, K. A., Bond, J. P., Schutz, K. C., Putnam, E. E., Leung, J. M., Lawley, T. D. & Shen, A. (2013). Global analysis of the sporulation pathway of Clostridium difficile. *PLoS Genet* **9**, e1003660.

Francis, M. B., Allen, C. A., Shrestha, R. & Sorg, J. A. (2013). Bile acid recognition by the Clostridium difficile germinant receptor, CspC, is important for establishing infection. *PLoS Pathog* **9**, e1003356.

Ghosh, S., Setlow, B., Wahome, P. G., Cowan, A. E., Plomp, M., Malkin, A. J. & Setlow, P. (2008). Characterization of spores of Bacillus subtilis that lack most coat layers. *J Bacteriol* **190**, 6741-8.

Hayman, J. R., Hayes, S. F., Amon, J. & Nash, T. E. (2001). Developmental expression of two spore wall proteins during maturation of the microsporidian Encephalitozoon intestinalis. *Infect Immun* **69**, 7057-66.

Heap, J. T., Pennington, O. J., Cartman, S. T., Carter, G. P. & Minton, N. P. (2007). The ClosTron: a universal gene knock-out system for the genus Clostridium. *J Microbiol Methods* **70**, 452-64.

Heeg, D., Burns, D. A., Cartman, S. T. & Minton, N. P. (2012). Spores of Clostridium difficile clinical isolates display a diverse germination response to bile salts. *PLoS One* **7**, e32381.

Henriques, A. O., Beall, B. W., Roland, K. & Moran, C. P., Jr. (1995). Characterization of cotJ, a sigma E-controlled operon affecting the polypeptide composition of the coat of Bacillus subtilis spores. *J Bacteriol* **177**, 3394-406.

Henriques, A. O. & Moran, C. P., Jr. (2000). Structure and assembly of the bacterial endospore coat. *Methods* **20**, 95-110.

Henriques, A. O. & Moran, C. P., Jr. (2007). Structure, assembly, and function of the spore surface layers. *Annu Rev Microbiol* **61**, 555-88.

Hensgens, M. P., Keessen, E. C., Squire, M. M., Riley, T. V., Koene, M. G., de Boer, E., Lipman, L. J. & Kuijper, E. J. (2012). Clostridium difficile infection in the community: a zoonotic disease? *Clin Microbiol Infect* **18**, 635-45.

Hussain, H. A., Roberts, A. P. & Mullany, P. (2005). Generation of an erythromycin-sensitive derivative of Clostridium difficile strain 630 (630Deltaerm) and demonstration that the conjugative transposon Tn916DeltaE enters the genome of this strain at multiple sites. *J Med Microbiol* **54**, 137-41.

Janoir, C., Deneve, C., Bouttier, S., Barbut, F., Hoys, S., Caleechum, L., Chapeton-Montes, D., Pereira, F. C., Henriques, A. O., Collignon, A., Monot, M. & Dupuy, B. (2013). Adaptive strategies and pathogenesis of Clostridium difficile from in vivo transcriptomics. *Infect Immun* **81**, 3757-69.

Joshi, L. T., Phillips, D. S., Williams, C. F., Alyousef, A. & Baillie, L. (2012). Contribution of spores to the ability of Clostridium difficile to adhere to surfaces. *Appl Environ Microbiol* **78**, 7671-9.

Kuehne, S. A., Cartman, S. T., Heap, J. T., Kelly, M. L., Cockayne, A. & Minton, N. P. (2010). The role of toxin A and toxin B in Clostridium difficile infection. *Nature* **467**, 711-3.

Lawley, T. D., Croucher, N. J., Yu, L., Clare, S., Sebaihia, M., Goulding, D., Pickard, D. J., Parkhill, J., Choudhary, J. & Dougan, G. (2009). Proteomic and genomic characterization of highly infectious Clostridium difficile 630 spores. *J Bacteriol* **191**, 5377-86.

McKenney, P. T., Driks, A. & Eichenberger, P. (2012). The Bacillus subtilis endospore: assembly and functions of the multilayered coat. *Nat Rev Microbiol* **11**, 33-44.

Merrigan, M., Venugopal, A., Mallozzi, M., Roxas, B., Viswanathan, V. K., Johnson, S., Gerding, D. N. & Vedantam, G. (2010). Human hypervirulent Clostridium difficile strains exhibit increased sporulation as well as robust toxin production. *J Bacteriol* **192**, 4904-11.

Merrigan, M. M., Venugopal, A., Roxas, J. L., Anwar, F., Mallozzi, M. J., Roxas, B. A., Gerding, D. N., Viswanathan, V. K. & Vedantam, G. (2013). Surface-layer protein A (SlpA) is a major contributor to host-cell adherence of Clostridium difficile. *PLoS One* **8**, e78404.

Nicholson, W. L. (2004). Ubiquity, longevity, and ecological roles of Bacillus spores. In *Bacterial Spore Formers: Probiotics and Emerging Application.* (ed. A. O. Henriques, E. Ricca and S. Cutting), pp. 1-15. Horizon Scientific Press: London.

Nicholson, W. L., Munakata, N., Horneck, G., Melosh, H. J. & Setlow, P. (2000). Resistance of Bacillus endospores to extreme terrestrial and extraterrestrial environments. *Microbiol Mol Biol Rev* **64**, 548-72.

Oliva, C., Turnbough, C. L., Jr. & Kearney, J. F. (2009). CD14-Mac-1 interactions in Bacillus anthracis spore internalization by macrophages. *Proc Natl Acad Sci U S A* **106**, 13957-62.

Oliva, C. R., Swiecki, M. K., Griguer, C. E., Lisanby, M. W., Bullard, D. C., Turnbough, C. L., Jr. & Kearney, J. F. (2008). The integrin Mac-1 (CR3) mediates internalization and directs Bacillus anthracis spores into professional phagocytes. *Proc Natl Acad Sci U S A* **105**, 1261-6.

Panessa-Warren, B. J., Tortora, G. T. & Warren, J. B. (2007). High resolution FESEM and TEM reveal bacterial spore attachment. *Microsc Microanal* **13**, 251-66.

Paredes-Sabja, D. & Sarker, M. R. (2012). Adherence of Clostridium difficile spores to Caco-2 cells in culture. *J Med Microbiol* **61**, 1208-18.

Pereira, F. C., Saujet, L., Tome, A. R., Serrano, M., Monot, M., Couture-Tosi, E., Martin-Verstraete, I., Dupuy, B. & Henriques, A. O. (2013). The spore differentiation pathway in the enteric pathogen Clostridium difficile. *PLoS Genet* **9**, e1003782.

Permpoonpattana, P., Tolls, E. H., Nadem, R., Tan, S., Brisson, A. & Cutting, S. M. (2011). Surface layers of Clostridium difficile endospores. *J Bacteriol* **193**, 6461-70.

Pizarro-Guajardo, M., Olguin-Araneda, V., Barra-Carrasco, J., Brito-Silva, C., Sarker, M. R. & Paredes-Sabja, D. (2014). Characterization of the collagen-like exosporium protein, BclA1, of Clostridium difficile spores. *Anaerobe* **25**, 18-30.

Plomp, M., Leighton, T. J., Wheeler, K. E., Hill, H. D. & Malkin, A. J. (2007). In vitro high-resolution structural dynamics of single germinating bacterial spores. *Proc Natl Acad Sci U S A* **104**, 9644-9.

Plomp, M., Leighton, T. J., Wheeler, K. E., Pitesky, M. E. & Malkin, A. J. (2005). Bacillus atrophaeus outer spore coat assembly and ultrastructure. *Langmuir* **21**, 10710-6.

Polonais, V., Mazet, M., Wawrzyniak, I., Texier, C., Blot, N., El Alaoui, H. & Delbac, F. (2010). The human microsporidian Encephalitozoon hellem synthesizes two spore wall polymorphic proteins useful for epidemiological studies. *Infect Immun* **78**, 2221-30.

Phetcharaburanin, J., Hong, H. A., Colenutt, C., Bianconi, I., Sempere, L., Permpoonpattana, P., Smith, K., Dembek, M., Tan, S., Brisson, M. C., Brisson, A. R., Fairweather, N. & Cutting, S. M. (2014). The Spore-Associated Protein BclA1 Affects the Susceptibility of Animals to Colonization and Infection by Clostridium difficile. *Mol Microbiol*.

Rupnik, M., Wilcox, M. H. & Gerding, D. N. (2009). Clostridium difficile infection: new developments in epidemiology and pathogenesis. *Nat Rev Microbiol* **7**, 526-36.

Ryan, A., Lynch, M., Smith, S. M., Amu, S., Nel, H. J., McCoy, C. E., Dowling, J. K., Draper, E., O'Reilly, V., McCarthy, C., O'Brien, J., Ni Eidhin, D., O'Connell, M. J., Keogh, B., Morton, C. O., Rogers, T. R., Fallon, P. G., O'Neill, L. A., Kelleher, D. & Loscher, C. E. (2011). A role for TLR4 in Clostridium difficile infection and the recognition of surface layer proteins. *PLoS Pathog* **7**, e1002076.

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**, 3186-96.

Saujet, L., Pereira, F. C., Serrano, M., Soutourina, O., Monot, M., Shelyakin, P. V., Gelfand, M. S., Dupuy, B., Henriques, A. O. & Martin-Verstraete, I. (2013). Genome-wide analysis of cell type-specific gene expression during spore formation in Clostridium difficile.

Serrano, M., Real, G., Santos, J., Carneiro, J., Moran, C. P., Jr. & Henriques, A. O. (2011). A negative feedback loop that limits the ectopic activation of a cell type-specific sporulation sigma factor of Bacillus subtilis. *PLoS Genet* **7**, e1002220.

Serrano, M., Zilhao, R., Ricca, E., Ozin, A. J., Moran, C. P., Jr. & Henriques, A. O. (1999). A Bacillus subtilis secreted protein with a role in endospore coat assembly and function. *J Bacteriol* **181**, 3632-43.

Soutourina, O. A., Monot, M., Boudry, P., Saujet, L., Pichon, C., Sismeiro, O., Semenova, E., Severinov, K., Le Bouguenec, C., Coppee, J. Y., Dupuy, B. & Martin-Verstraete, I. (2013). Genome-wide identification of regulatory RNAs in the human pathogen Clostridium difficile. *PLoS Genet* **9**, e1003493.

Sylvestre, P., Couture-Tosi, E. & Mock, M. (2002). A collagen-like surface glycoprotein is a structural component of the Bacillus anthracis exosporium. *Mol Microbiol* **45**, 169-78.

Walker, J. R., Gnanam, A. J., Blinkova, A. L., Hermandson, M. J., Karymov, M. A., Lyubchenko, Y. L., Graves, P. R., Haystead, T. A. & Linse, K. D. (2007). Clostridium taeniosporum spore ribbon-like appendage structure, composition and genes. *Mol Microbiol* **63**, 629-43.

Xue, Q., Gu, C., Rivera, J., Hook, M., Chen, X., Pozzi, A. & Xu, Y. (2011). Entry of Bacillus anthracis spores into epithelial cells is mediated by the spore surface protein BclA, integrin alpha2beta1 and complement component C1q. *Cell Microbiol* **13**, 620-34.

Yolton, D. P., Huettel, R. N., Simpson, D. K. & Rode, L. J. (1972). Isolation and partial chemical characterization of the spore appendages of Clostridium taeniosporum. *J Bacteriol* **109**, 881-5.

Zhang, J., Fitz-James, P. C. & Aronson, A. I. (1993). Cloning and characterization of a cluster of genes encoding polypeptides present in the insoluble fraction of the spore coat of Bacillus subtilis. *J Bacteriol* **175**, 3757-66.

Chapter 6

Discussion

One of the features that contribute to the pathogenicity of *C. difficile* is its ability to produce spores. Sporulation, which occurs promptly *in vivo*, allows the shedding of high amounts of spores, facilitating transmission, and the accumulation of spores inside the host, ensuring persistence and recurrence. Thus, it is not surprising that some of the most successful strategies that have been developed to either prevent or cure CDI are linked to the inhibition of spore formation and germination (Babakhani *et al.*, 2012; Howerton *et al.*, 2013). Still, at the time we started this work, little was known about the regulation of spore formation in this organism and the details of its functional architecture and composition. Here, we have elucidated the regulatory network that controls spore formation in *C. difficile*. Furthermore, we have identified a total of 225 genes whose expression is controlled by the main sporulation regulators σ^F, σ^E, σ^G and σ^K. Among the genes controlled by σ^K one, *CD1581* (*sp17*), was found to encode an abundant component of the spore surface with a clear role in host colonization.

The sporulation regulatory network of *C. difficile*

Using the fluorescent reporter for single cell analysis developed for *C. difficile* (Chapter 2) we have determined, for the first time in a Clostridial species, the compartments and main periods of activity of the sigma factors that govern spore formation (Chapter 3). Despite the high degree of similarity found between *C. difficile* and the model organism *B. subtilis*, we noticed that the activities of the sigma factors in the first are less segregated in time, revealing that in this organism sporulation is not as tightly controlled as it is in *B. subtilis*. In *B. subtilis*, the forespore and mother cell lines of gene expression are organized in a series of interlocked type I coherent and incoherent FFLs. Type I incoherent FFLs are thought to generate pulses of gene expression whereas type I coherent FFLs act as persistence detectors

and cause delays in gene expression; their action fine tunes gene transcription, minimise noise, and push morphogenesis forward (Eichenberger *et al.*, 2004; Wang *et al.*, 2006). Since most of the transcriptional regulators that are involved in these FFLs (Figure 1.6), are absent from the *C. difficile* genome, the corresponding FFLs are also absent, perhaps leading to reduced fidelity of the morphogenetic process. Furthermore, some of the cell-cell signalling pathways that operate to control the activation of the sigma factors of *B. subtilis* are absent in *C. difficile* (Chapter 3 and 4). This conclusion was also reached by the authors of an independent work that addressed the regulation of the sporulation pathway in *C. difficile* (Fimlaid *et al.*, 2013). Thus, and contrary to *B. subtilis*, σ^F is not strictly required for σ^E activity, the σ^G activation is independent of σ^E and σ^K activation does not rely on σ^G.

In *C. difficile*, partial activation of σ^E, and consequently of σ^K, occurs independently of σ^F (Chapters 3 and 4). We have further confirmed that this is due to the σ^F-independent production of SpoIIR, the signalling protein that activates the protease required for Pro-σ^E cleavage, and thus for σ^E activation (Chapter 4). Accordingly, transcription of *spoIIR*, which was initially thought to be completely dependent on σ^F, is still detected in a *sigF* mutant after asymmetric division, in the forespore compartment. Only one other sigma factor able of driving forespore-specific gene expression takes part during sporulation, and this is σ^G. Thus, either SpoIIR is partially controlled by σ^G, or other as yet unidentified factor with the ability to drive forespore-specific gene expression is at play in *C. difficile*.

Activity of σ^G in *C. difficile* is σ^F-dependent, and is detected after asymmetric division, in the forespore compartment (Chapter 3). Although the transcription of *sigG* is mainly driven from a σ^F-dependent promoter, in *B. subtilis sigG* is also co-transcribed with the *spoIIG* operon (Sun *et al.*, 1991). We assume this is also the case in *C. difficile*, although we failed to detect any transcription or activity of σ^G in *sigF* mutant cells. However, we cannot

discard the possibility that some σ^G activity originated in a σ^F-independent manner is still present at low but sufficient levels to drive *spoIIR* transcription (Figure 4.5). Because the consensus sequences recognized by the sporulation sigma factors are very similar in *B. subtilis* and *C. difficile* (Chapter 4), σ^F and σ^G may also have overlapping promoter specificities in the last, and both may be able to drive *spoIIR* expression. If this is the case, *spoIIR* transcription and the subsequent activation of σ^E should be totally abolished in a double *sigF sigG* mutant. Recent developments in the genetic tools available for *C. difficile* will allow the test of this idea in the near future (Ng *et al.*, 2013). As an alternative, mutating the consensus sequence for σ^F/σ^G binding in the *spoIIR* promoter should also provide valuable information.

Results from Chapters 3 and 4 also revealed that the activity of σ^G is σ^E-independent. However, the genes coding homologues of the proteins that form a mother cell-to-forespore channel required for σ^G activity in *B. subtilis* are present in the *C. difficile* genome, and seem to be regulated in a similar manner (Chapter 4). Nevertheless, the ability of SpoIIQ and SpoIIIAH, the two proteins that are crucial for the assembly of the *B. subtilis* channel (Camp and Losick, 2009; Doan *et al.*, 2009) to interact and to form a channel hasn´t been tested yet in *C. difficile*. Still, if present, such structure may be responsible for the increased activity of σ^G observed upon engulfment completion (Chapter 3). Importantly, σ^G activity and engulfment completion are uncoupled in *C. difficile.* This feature may contribute to elucidate the function(s) of the SpoIIIAH-SpoIIQ channel, which is not yet fully understood, even for *B. subtilis*.

Another striking difference in the regulation of *C. difficile* and *B. subtilis* sporulation has to due with the σ^G-independent activation of σ^K, resulting in coat deposition around the forespore of a *sigG* mutant (Chapter 3). In fact, the homologues of the genes involved in the σ^G to σ^K signalling pathway operating in *B. subtilis* are absent from the *C. difficile* genome, with

the exception of SpoIVB. While such pathway working at the post-translational level is absent, σ^K activation in *C. difficile* is still controlled at the transcription level and at the level of *sigK* recombination (Chapters 3 and 4). In *B. subtilis*, transcription of *sigK* is driven by the joint action of σ^E and SpoIIID (Kunkel *et al.*, 1989; Halberg and Kroos, 1994). SpoIIID is also required for the excision of the *skin*^{Bs} element that interrupts *sigK*, (Kunkel *et al.*, 1990), and to shut down many σ^E-controlled targets (Kroos *et al.*, 1989). In Chapter 4 we have shown that the *C. difficile* SpoIIID directs the repression of σ^E-controlled genes and contributes to the activation of *sigK* transcription, but no conclusions could be drawn about its effect on *sigK* recombination. To address this issue, we have constructed a fusion of the *SNAP*^{Cd} to the 5´ end of the interrupted *sigK* gene, that is only expressed when excision of the *skin*^{Cd} element occurs (Figure 6.1A). Preliminary results suggest that SpoIIID is indeed required for *sigK* recombination given that no fluorescence signal from a translational SigK-SNAP fusion could be detected in a *spoIIID* mutant (Figure 6.1B). To discard that this was a specific effect of SpoIIID on *sigK* recombination, rather than on *sigK* transcription, we have confirmed that transcription of *sigK* is not severely impaired by disruption of *spoIIID* (Figure 6.2B, bottom panel). Furthermore, amplification of a recombined *sigK* gene from the *C. difficile* genome, readily observed for wild type cells, is not observed in *spoIIID* (Figure 6.1C, red arrow).

In *B. subtilis*, SpoIIID is required for the transcription of *spoIVCA*, coding for a site-specific recombinase that mediates *skin*^{Bs} excision and *sigK* recombination (Popham and Stragier, 1992). The SpoIVCA recombinase is present in the *skin*^{Bs} element, and among the *C. difficile* homologues of SpoIVCA one, CD1231, was found to be also located in the s*kin*^{Cd} element. This recombinase is, however, expressed in all cells of the population, sporulating and non-sporulating, without any preferential period or compartment (Figure 6.1D). Indeed, inactivation of SpoIIID or σ^E does not impact on *CD1231* expression (Chapter 4). Thus, two possibilities for the

control of *sigK* recombination in *C. difficile* by SpoIIID are plausible: either the *skin*^{Cd} excision is mediated by another recombinase present in the *C. difficile* genome and whose transcription is controlled by SpoIIID; or CD1231 is the recombinase that acts on *sigK*, but its activity is controlled at the post-transcriptional level by an yet unknown factor that depends on SpoIIID. However, if the first hypothesis is correct, our transcriptomic analysis of the

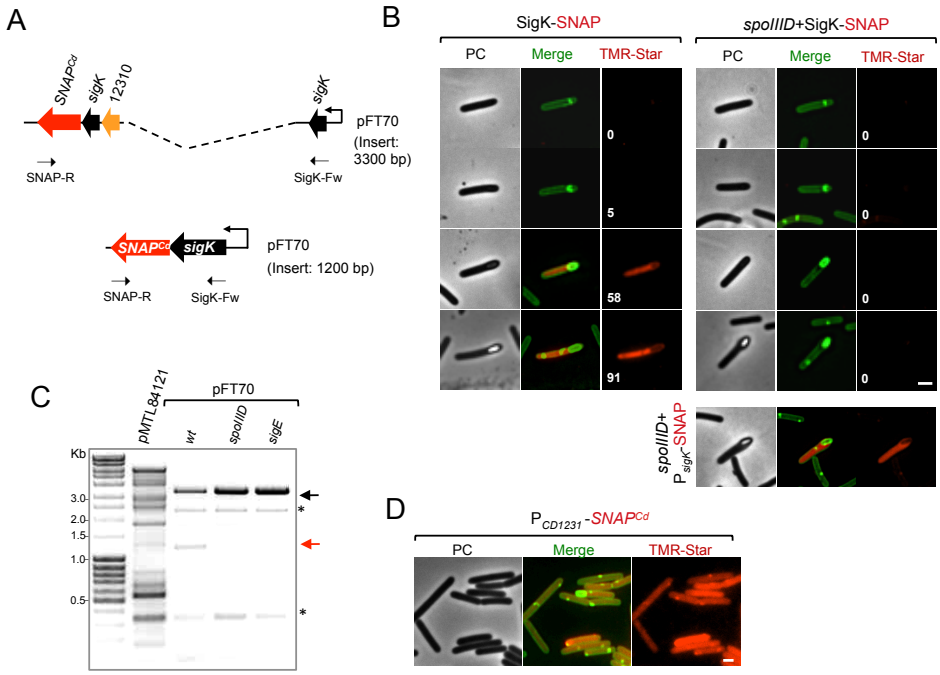


Figure 6.1. SpoIIID-mediated excision of the *skin*^{Cd} element. (A) Schematic representation of the plasmid pFT70, containing a SigK-SNAP fusion. The *sigK* gene is interrupted by a short version of the *skin*^{Cd} element (also present on pFT38, described on Chapter 3). When recombination occurs, the SigK-SNAP protein is produced. The size of the intact and recombined inserts is indicated. Primers used for the PCR are also indicated. **(B)** Fluorescence microscopy analysis of sporulating cells carrying SigK-SNAP (depicted in A) or P_{sigK}-SNAP (see Chapter 3) fusions in a *wt* or *spoIIID* mutant background. Cells grown in SM were collected at 24h and stained with the membrane dye MTG and the SNAP substrate TMR-Star. Cells belonging to four main stages of sporulation, ordered from the top to the bottom, are shown. **(C)** Analysis of *skin* excision by PCR. PCR was performed with the primer pair indicated in (A) on genomic DNA extracted from wt, *spoIIID* or *sigE* mutant sporulating cells, collected at 24 hours of growth in SM. The red arrow points to the recombined *sigK*, while the black arrow points to the interrupted gene. Non-specific bands are marked with an asterisk. Plasmid DNA from the pMTL84121 vector was used as a negative control. The size of the molecular marker (in Kb) is indicated on the left. **(D)** Fluorescence microscopy analysis of wt *C. difficile* cells carrying a P_{CD1231}-SNAP fusion. Cells grown in SM were collected at 24h and stained with the membrane dye MTG and the SNAP substrate TMR-Star. Scale bar in (B) and (D), 1µm.

spoIIID mutant failed to detect this other recombinase. Furthermore, favouring the second hypothesis, we have noticed that the expression of a gene (*CD1234*) present in the s*kin*^{Cd} element is down-regulated by a 1000 fold in a *spoIIID* mutant. We speculate that CD1234 acts by controlling translation or activity of the recombinase at the right time and in the right cell compartment. In *B. subtilis*, synthesis of SpoIVCA requires the action of the tmRNA (transfer-messenger RNA) SsrA, which rescues stalled ribosomes under stressful conditions (Abe *et al.*, 2008). A similar system may operate to control CD1231 synthesis in *C. difficile,* and perhaps CD1234 contributes to this system. The characterization of mutants for *CD1231* and *CD1234*, in progress, is an important research goal. In Chapter 3 we have shown that the presence of an intact *sigK* strongly interferes with sporulation, blocking the process prior to the asymmetric division. Supporting this hypothesis, two naturally occurring s*kin*$^{Cd-}$ strains sporulate poorly, and the expression of an intact *sigK* gene from a replicative plasmid in a *skin*$^{Cd+}$ strain also culminates in impaired sporulation (Haraldsen and Sonenshein, 2003). Surprisingly, we observed that among a panel of epidemic strains isolated in hospitals across Europe, several strains that lack the s*kin*^{Cd} element sporulate normally. In addition, in a parallel study addressing the function of σ^K in sporulation in the *C. difficile* strain JIR8094, the sporulation defect resulting from a *sigK*-disrupted gene could be rescued by a multicopy plasmid containing a copy of the intact *sigK* gene expressed under the control of its native promoter (Fimlaid *et al.*, 2013). These results are intriguing and suggest that other unidentified determinants present in the genomes of these strains are involved in the control of spore formation and *sigK* activation. Interestingly, the aforementioned panel of European epidemic strains contains strains with s*kin*^{Cd} elements of smaller sizes. Despite the high degree of variation among the size of the s*kin*^{Cd} element, the *CD1231* and *CD1234* genes are present in all s*kin*^{Cd}, emphasizing the important role that these genes may play in *sigK* recombination.

The σ^F, σ^E, σ^G and σ^K regulons of *C. difficile*

Our genome wide analysis of gene transcription during spore formation allowed the identification of the genes controlled by the sigma factors of sporulation. Figure 6.2 summarizes the chromosome location of the genes assigned to each regulon, highlighting the genes that were identified in either or both the studies of Saujet *et al*. (2013) and Fimlaid *et al*. (2013). Some discrepancies between the two studies do exist, probably due to the different conditions used to induce sporulation and time points analysed. Nevertheless, with the exception of the σ^F regulon, there is a good overlap among the genes identified in both studies (Figure 6.2). In both studies the σ^F regulon is the smallest and the σ^E regulon is the largest (Figure 6.2), as in *B. subtilis*. While in *B. subtilis* the σ^F regulon members are predominantly located in the left hand side of the genome, and a set is clustered near the replication origin, this asymmetry is not clearly observed for the members of the *C. difficile* σ^F regulon (Figure 6.2A). Strikingly, the *spoIIA* operon, whose location near the chromosome terminus is important for the compartmentalization of σ^F activity in *B. subtilis* (Dworkin and Losick, 2001), is positioned close to the one third portion of the *C. difficile* chromosome that is presumably trapped inside the forespore compartment (Figure 6.2A). Transposition of the *B. subtilis spoIIA* operon to a position equivalent to the one observed for *C. difficile* results in a 2-fold reduction in the sporulation efficiency. We presently do not know whether transient genetic asymmetry also contributes to compartmentalization of σ^F activity in *C. difficile* or exactly which portion of the chromosome is present in the forespore compartment after asymmetric division. In addition, it is not known whether the position of the *C. difficile spoIIA* operon affects the activation of σ^F. One way of testing this would be to transpose the *spoIIA* operon close to *oriC* or away from the boundary of the arc of the chromosome presumed to be trapped in the forespore. Despite the

discrepancies in the chromosomal distribution of the σ^F-controlled genes mentioned above, *spoIIR*, which is required for the activation of σ^E, is located near the origin of replication as in *B. subtilis*. This location enables rapid *spoIIR* transcription and subsequent σ^E activation, something that is required to prevent the formation of a second asymmetric septum in *B. subtilis* (Khvorova *et al.*, 2000; Zupancic *et al.*, 2001), and given the phenotypes of the *sigF* and *sigE* mutants here reported, also in *C. difficile*.

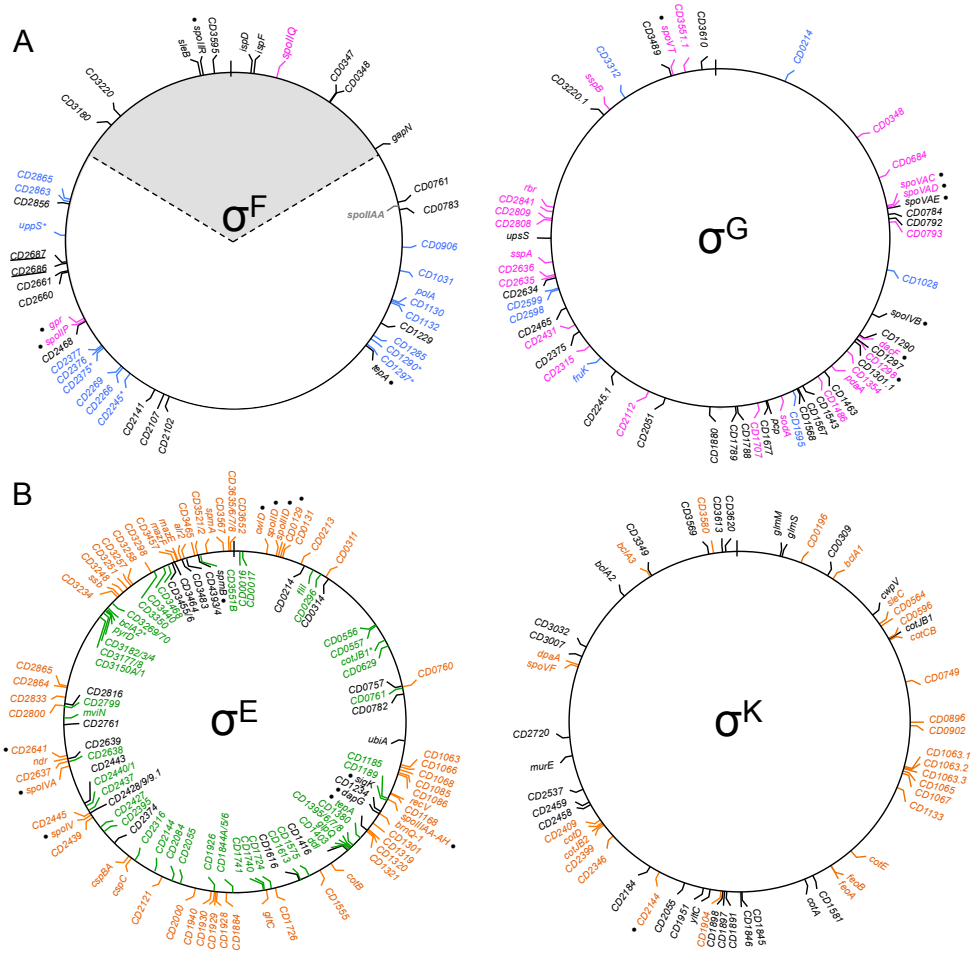


Figure 6.2. Map of the σ^F, σ^E, σ^G and σ^K regulons. (A) Genes found to be controlled by σ^F and σ^G in both transcriptional studies performed here and by Fimlaid *et al.*, are depicted in mangenta. Genes identified only in our study are shown in black, while genes identified only by Fimlaid *et al.* are shown in blue. Marked with an asterisk are genes found to be regulated by σ^G in our study. Underlined are genes found to be regulated by σ^G in Fimlaid *et al.* **(B)** Genes found to be controlled by σ^E and σ^K in both ours and Fimlaid *et al.* studies

are depicted in orange. Genes identified only in our study are shown in black, while genes identified only by Fimlaid *et al.* are shown in green. Marked with an asterisk are genes found to be regulated by σ^K in our study. The chromosome origin is depicted by a line crossing the main circle. The black dots in (A) and (B) denote the genes that are part of the genomic signature for sporulation proposed by Abecasis *et al.*, 2013.

Most of the genes assigned to the σ^F, σ^E, σ^G and σ^K regulons code for proteins whose function is unknown. Nevertheless, the information offered by our work concerning when (early or late) and where (forespore of mother cell) are these expressed provides some insights about their putative functions. These can be related to host colonization, germination, transmission or persistence in the host, all crucial processes for *C. difficile* pathogenesis.

A link between the *C. difficile* spore surface and host colonization

The surface layers of spores from different pathogens have been implicated in adherence and interaction with the host in a manner that is favourable for their survival. Here we have identified a major component of the *C. difficile* spore surface, termed Sp17, which is required for efficient colonization in a mice model for CDI (Chapter 5). The absence of Sp17 from *C. difficile* spores results in the impaired assembly of the spore coat and, we presume, exosporium layers. Although the presence of this last has been reported for the strain 630 here used, we have found this structure extremely fragile and difficult to observe by TEM. Thus, it would be interesting to also address impact of *sp17* inactivation in a strain with a more robust exosporium layer.

Following ingestion, the *C. difficile* spores have to germinate to establish a population in the colon. Thus, spores must adhere to the colonic mucosa in order to prevent the premature shed of incompletely germinated spores from the host. The outer layers of *C. difficile* have been implicated in the adhesion of the spore to colon epithelial cells (Panessa-Warren *et al.*, 1997; Panessa-Warren *et al.*, 2007; Paredes-Sabja and Sarker, 2012; Joshi *et*

al., 2012), and spores that lack BclA1, an homologue of the *B. anthracis* exosporium protein BclA, are impaired in colonization in a mouse model (Phetcharaburanin *et al.*, 2014). We propose that Sp17, as a component of these layers, also plays a role in spore adhesion, and that the decreased adherence of spores to the spore surface is the main cause for the reduced colonization. A putative contribution of Sp17 for spore adherence must go through the formation of an appendage-like structure that most likely is required for the interaction with a receptor located on the colonic epithelial cells (Chapter 5, Figure 5.11). This appendage is only observed for 22% of the spores in a population, and we speculate that the fraction of spores producing this appendage will adhere more efficiently and may ensure persistence in the host. Conversely, the remaining 78% of the population may enter growth, toxin production, and sporulation, and be more related to pathogen transmission. The basis for the formation of these two spore morphotypes is unknown, but we can speculate that the absence of the FFLs that govern gene expression later in the mother cell may result in a more random expression of the σ^K targets, such as *sp17*, and thus, in different spore morphotypes. Preliminary observations suggest that the dimensions of the appendage, as well as the fraction of spores that assemble it, differs among epidemic strains. If confirmed, these observations could help explain the characteristics of different epidemic strains. Separation of these two populations followed by adhesion and *in vivo* colonization assays will further elucidate the role played by these structures. Despite the many reports about the interaction of *C. difficile* with epithelial cells, a receptor has not been identified yet. Cross-linking assays of spores to epithelial cells should also be performed to identify the receptor, and an interaction of this putative receptor with Sp17 should be the tested to further clarify its role in adherence and colonization.

Finally, Sp17 is unique to *C. difficile,* and is present in the genome of all strains sequenced so far. Thus, and given the large amount of spores that

are shed by the some of epidemic strains, detection of an abundant spore component that is unique to *C. difficile* constitutes a good alternative to most of the diagnosis methods currently available. Furthermore, Sp17 is the most abundant component of the *C. difficile* spore surface, so it would be interesting to test whether this protein is also immunodominant. If so, Sp17 can be a strong candidate for a vaccine against *C. difficile*, something that has not yet been achieved.

Concluding Remarks

Here we have analysed the sporulation pathway of *C. difficile*, an important human pathogen. At the level of regulation, this pathway seems closer to the model organism *B. subtilis* than the other *Clostridia* studied so far. A large set of genes that are expressed during sporulation were identified, and among them are strong candidates for genes coding for new germination receptors and surface proteins involved in spore dissemination and host colonization. Sp17 is an example of the last, which we have here identified and characterized, establishing for the first time a direct link between the spore surface and colonization. While these are of paramount importance towards the understanding of the role played by spores in *C. difficile* pathogenesis, the fluorescent reporter for single cell analysis in *C. difficile* here developed represents an invaluable tool to dissect many other pathogenesis-related issues in this organism.

REFERENCES

Abe, T., Sakaki, K., Fujihara, A., Ujiie, H., Ushida, C., Himeno, H., Sato, T. & Muto, A. (2008). tmRNA-dependent trans-translation is required for sporulation in Bacillus subtilis. *Mol Microbiol* **69**, 1491-8.

Abecasis, A., Serrano, M., Alves, L., Quintais, L., Pereira-Leal, J. B. & Henriques, A. O. (2013). A genomic signature and the identification of new endosporulaion genes. *J Bacteriol* **195**, 2101-2115.

Babakhani, F., Bouillaut, L., Gomez, A., Sears, P., Nguyen, L. & Sonenshein, A. L. (2012). Fidaxomicin inhibits spore production in Clostridium difficile. *Clin Infect Dis* **55 Suppl 2**, S162-9.

Camp, A. H. & Losick, R. (2009). A feeding tube model for activation of a cell-specific transcription factor during sporulation in Bacillus subtilis. *Genes Dev* **23**, 1014-24.

Doan, T., Morlot, C., Meisner, J., Serrano, M., Henriques, A. O., Moran, C. P., Jr. & Rudner, D. Z. (2009). Novel secretion apparatus maintains spore integrity and developmental gene expression in Bacillus subtilis. *PLoS Genet* **5**, e1000566.

Dworkin, J. & Losick, R. (2001). Differential gene expression governed by chromosomal spatial asymmetry. *Cell* **107**, 339-46.

Eichenberger, P., Fujita, M., Jensen, S. T., Conlon, E. M., Rudner, D. Z., Wang, S. T., Ferguson, C., Haga, K., Sato, T., Liu, J. S. & Losick, R. (2004). The program of gene transcription for a single differentiating cell type during sporulation in Bacillus subtilis. *PLoS Biol* **2**, e328.

Fimlaid, K. A., Bond, J. P., Schutz, K. C., Putnam, E. E., Leung, J. M., Lawley, T. D. & Shen, A. (2013). Global analysis of the sporulation pathway of Clostridium difficile. *PLoS Genet* **9**, e1003660.

Halberg, R. & Kroos, L. (1994). Sporulation regulatory protein SpoIIID from Bacillus subtilis activates and represses transcription by both mother-cell-specific forms of RNA polymerase. *J Mol Biol* **243**, 425-36.

Haraldsen, J. D. & Sonenshein, A. L. (2003). Efficient sporulation in Clostridium difficile requires disruption of the sigmaK gene. *Mol Microbiol* **48**, 811-21.**Howerton, A., Patra, M. & Abel-Santos, E.** (2013). A new strategy for the prevention of Clostridium difficile infection. *J Infect Dis* **207**, 1498-504.

Joshi, L. T., Phillips, D. S., Williams, C. F., Alyousef, A. & Baillie, L. (2012). Contribution of spores to the ability of Clostridium difficile to adhere to surfaces. *Appl Environ Microbiol* **78**, 7671-9.

Khvorova, A., Chary, V. K., Hilbert, D. W. & Piggot, P. J. (2000). The chromosomal location of the Bacillus subtilis sporulation gene spoIIR is important for its function. *J Bacteriol* **182**, 4425-9.

Kroos, L., Kunkel, B. & Losick, R. (1989). Switch protein alters specificity of RNA polymerase containing a compartment-specific sigma factor. *Science* **243**, 526-9.

Kunkel, B., Kroos, L., Poth, H., Youngman, P. & Losick, R. (1989). Temporal and spatial control of the mother-cell regulatory gene spoIIID of Bacillus subtilis. *Genes Dev* **3**, 1735-44.

Kunkel, B., Losick, R. & Stragier, P. (1990). The Bacillus subtilis gene for the development transcription factor sigma K is generated by excision of a dispensable DNA element containing a sporulation recombinase gene. *Genes Dev* **4**, 525-35.

Ng, Y. K., Ehsaan, M., Philip, S., Collery, M. M., Janoir, C., Collignon, A., Cartman, S. T. & Minton, N. P. (2013). Expanding the repertoire of gene tools for precise manipulation of the Clostridium difficile genome: allelic exchange using pyrE alleles. *PLoS One* **8**, e56051.

Panessa-Warren, B. J., Tortora, G. T. & Warren, J. B. (1997). Exosporial membrane plasticity of Clostridium sporogenes and Clostridium difficile. *Tissue Cell* **29**, 449-61.

Panessa-Warren, B. J., Tortora, G. T. & Warren, J. B. (2007). High resolution FESEM and TEM reveal bacterial spore attachment. *Microsc Microanal* **13**, 251-66.

Paredes-Sabja, D. & Sarker, M. R. (2012). Adherence of Clostridium difficile spores to Caco-2 cells in culture. *J Med Microbiol* **61**, 1208-18.

Phetcharaburanin, J., Hong, H. A., Colenutt, C., Bianconi, I., Sempere, L., Permpoonpattana, P., Smith, K., Dembek, M., Tan, S., Brisson, M. C., Brisson, A. R., Fairweather, N. & Cutting, S. M. (2014). The Spore-Associated Protein BclA1 Affects the Susceptibility of Animals to Colonization and Infection by Clostridium difficile. *Mol Microbiol*.

Popham, D. L. & Stragier, P. (1992). Binding of the Bacillus subtilis spoIVCA product to the recombination sites of the element interrupting the sigma K-encoding gene. *Proc Natl Acad Sci U S A* **89**, 5991-5.

Sun, D. X., Cabrera-Martinez, R. M. & Setlow, P. (1991). Control of transcription of the Bacillus subtilis spoIIIG gene, which codes for the forespore-specific transcription factor sigma G. *J Bacteriol* **173**, 2977-84.

Wang, S. T., Setlow, B., Conlon, E. M., Lyon, J. L., Imamura, D., Sato, T., Setlow, P., Losick, R. & Eichenberger, P. (2006). The forespore line of gene expression in Bacillus subtilis. *J Mol Biol* **358**, 16-37.

Zupancic, M. L., Tran, H. & Hofmeister, A. E. (2001). Chromosomal organization governs the timing of cell type-specific gene expression required for spore formation in Bacillus subtilis. *Mol Microbiol* **39**, 1471-81.

Appendices

APPENDICES SUPPORTING TEXT

Table A1 lists the plasmids used or constructed during the course of the work here presented.

Table A2 lists the oligonucleotide primers used for PCR, qRT-PCR, intron retargeting or sequencing.

Tables A3 through A7 contain relevant information to support the findings described on Chapter 4.

Table A3 lists the genes under the direct control of the forespore sigma factors σ^F and σ^G, among the genes positively controlled by these sigma factors. This pool of genes was determined by combining the transcriptome data and promoter identifications. Genes are listed according to the functional groups of proteins encoded.

Table A4 is similar to Table A3, except that it lists the genes under the direct control of the mother cell sigma factors σ^E and σ^K.

Table A5 shows the validation of microarray data by qRT-PCR on selected genes belonging to each of the *sig* regulons. qRT-PCR experiments were performed on two different RNA preparations for each mutant. The results presented correspond to the mean of at least two independent experiments.

Table A6 contains qRT-PCR data that shows complementation of the mutants inactivated for the sigma factors of sporulation. Experiments were performed on two different RNA preparations for each mutant and each complemented strain. Cells were harvested after 14h of growth for the strain 630Δ*erm*, the *sigE* and *sigF* mutants and the *sigF* mutant containing pMTL84121-*sigF* and the *sigE* mutant containing pMTL84121-*sigE* ; after 20h of growth for the strain 630Δ*erm*, the *sigG* mutant and the *sigG* mutant containing pMTL84121-*sigG*; and after 24 h of growth for the strain 630Δ*erm*, the *sigK* mutant and the *sigK* mutant containing pMTL84121-

sigK$^{short\ skin}$. The results presented correspond to the mean of at least two independent experiments.

Finally, Table A7 lists genes that are controlled by SpoIIID in transcriptome. RNA was extracted from strain 630Δ*erm* strain and the *spoIIID* mutant after 15 h of growth in SM medium.

REFERENCES

Eichenberger, P., Fujita, M., Jensen, S. T., Conlon, E. M., Rudner, D. Z., Wang, S. T., Fergusson, C., Haga, K., Sato. T., Liu, J. S. and Losick, R. (2004). The program of gene transcription for a single differentiating cell type during sporulation in *Bacillus subtilis*. *PLoS Biol* **2**, e328.

Fagan, R. P. and Fairweather, N. F. (2011). *Clostridium difficile* has two parallel and essential Sec secretion systems. *J Biol Chem* **286**, 27483-27493.

Heap, J. T., Pennington, O. J., Cartman, S. T., Carter, G. P. and Minton, N. P. (2007). The ClosTron: a universal gene knock-out system for the genus Clostridium. *J Microbiol Methods* **70**, 452-464.

Heap, J. T., Pennington, O. J., Cartman, S. T. and Minton, N. P. (2009). A modular system for Clostridium shuttle plasmids. *J Microbiol Methods* **78**, 79-85.

Lawley, T. D., Croucher, N. J., Yu, L., Clare, S., Sebaihia, M., Goulding, D., Pickard, D. J., Parkhill, J., Choudhary, J. and Dougan, G. (2009). Proteomic and genomic characterization of highly infectious *Clostridium difficile* 630 spores. *J Bacteriol* **191**, 5377-5386.

Manganelli, R., Provvedi, R., Berneri, C., Oggioni, M., Pozzi, G. (1998). Insertion vectors for construction of recombinant conjugative transposons in *Bacillus subtilis* and *Enterococcus faecalis*. *FEMS microb lett*. **168**, 259–268.

Table A1 – Plasmids used in this study

Plasmid	Relevant features	Origin
pET33b	His-tag fusion protein expression vector (Km^R)	Novagen
pET28a	His-tag fusion protein expression vector (Km^R)	"
pMTL007	ClosTron plasmid (Cm^R/Tm^R)	Heap *et al.,* 2007
pMTL84121	*Clostridium* modular plasmid (Cm^R/Tm^R)	Heap *et al.*, 2009
pSMB47	Tn*916* integrational vector (Cm^R/Tm^R)	Manganelli *et al.*, 1998
pRPF185	pMTL960 plasmid carrying P_{tet}-*gusA* (Cm^R/Tm^R)	Fagan and Fairwearther, 2011
pFT07	pMTL007::Cdi-*sigE*-453s (Cm^R/Tm^R)	This work
pFT08	pMTL007::Cdi-*sigF*-459s (Cm^R/Tm^R)	"
pFT09	pMTL007::Cdi-*sigG*-546s (Cm^R/Tm^R)	"
pFT13	pMTL007::Cdi-*sigK*-102s (Cm^R/Tm^R)	"
pFT32	pMTL84121-*sigF* (Cm^R/Tm^R)	"
pFT34	pET28a-*sigF*-his_6	"
pFT35	pET28a-*sigE*-his_6	"
pFT38	pMTL84121-$sigK^{skin+}$ (Cm^R/Tm^R)	"
pFT39	pMTL84121-*sigE* (Cm^R/Tm^R)	"
pFT40	pMTL84121-*sigG* (Cm^R/Tm^R)	"
pFT41	*sp17* in pET33b (Km^R)	
pFT42	pMTL84121-$sigK^{skin-}$ (Cm^R/Tm^R)	"
pFT43	pMTL84121-*sp17* (Cm^R/Tm^R)	"
pFT46	P_{tet}-$SNAP^{Cd}$ (Cm^R/Tm^R)	"
pFT47	pMTL84121-$SNAP^{Cd}$ (Cm^R/Tm^R)	"
pFT48	pFT47 containing the *sigF* promoter region	"
pFT49	pFT47 containing the *sigE* promoter region	"
pFT50	pFT47 containing the *sigG* promoter region	"
pFT51	pFT47 containing the *sigK* promoter region	"
pFT52	pFT47 containing the *sp17* promoter region	"
pFT53	pFT47 containing the *gpr* promoter region	"
pFT54	pFT47 containing the *spoIIIAA* promoter region	"
pFT55	pFT47 containing the *sspA* promoter region	"
pFT57	Sp17-SNAP in pMTL84121	"
pFT58	pMTL84121-linker-$SNAP^{Cd}$ (Cm^R/Tm^R)	"
pFT63	pFT58-*cotB*	"
pFT64	pFT58-*cotE*	"
pFT69	pFT47 containing *cotE* promoter region	"
pFT71	pSMB47- *sp17*	"
pFT73	P_{tet}-$Clip_I^{Cd}$	"
pMS462	pFT47-P*spoIIR*-$SNAP^{Cd}$	"
pMS459	pMTL007::Cdi-*CD3564*-38a	"
pMS480	pFT58-*spoIIQ*	"
pMS481	pFT58-*spoIIIAH*	"
pDIA6123	pMTL007::Cdi-*CD0216*-39s	"
pDIA6124	pMTL007::Cdi-*CD3499*-157a	"
pDIA6132	pMTL84121-*spoVT*	"
pDIA6133	pMTL84121-*spoIIID*	"
pDIA6135	pMTL84121-*spoIIR*	"

Table A2 – Oligonucleotide primers used in this study

Primer[a]	Sequence (5´ to 3´)[b]	Description
Primers used for cloning in vectors that allow protein overexpression and purification		
5´-sigF-NcoI	GAGCTCCATGGGAGGTTTATTAATGGAAGTAACTG	*sigF*-pet28a Fw
3' -sigF-XhoI	TGCCTCGAGCGATATATATTCTTTTAACTTAG	*sigF*-pet28a Rev
5´-sigE-NcoI	GAGCTCCATGGGCTTTATTACTATGTTAGGTATTA	*sigE*-pet28a Fw
3' -sigE-XhoI	TGCCTCGAGTACAAATTTTTTCATTTCTTTTTGC	*sigE*-pet28a Rev
sp17-pET33b Fw	GATATGTTCCATGGAAAATAAAAAATGTTATTC	5´-*sp17* NcoI
sp17-pET33b Rev	CTAATTCTCGAGTTTTCTACAGCAGTTACA	3´-*sp17* XhoI-no STOP
Primers to retarget the type II intron		
EBS	CGAAATTAGAAACTTGCGTTCAGTAAAC	PCR ClosTron
CT001	AAAAAAGCTTATAATTATCCTTAGGAAACGAACTTGTGCGCCCAGATAGGGTG	IBS-*sigF*-459s
CT002	CAGATTGTACAAATGTGGTGATAACAGATAAGTCGAACTTTTAACTTACCTTTCTTTGT	EBS1d-*sigF*-459s
CT003	TGAACGCAAGTTTCTAATTTCGATTTTTCCTCGATAGAGGAAAGTGTCT	EBS2-*sigF*-459s
CT004	AAAAAAGCTTATAATTATCCTTAGATTTCGTAATGGTGCGCCCAGATAGGGTG	IBS-*sigE*-453s
CT005	CAGATTGTACAAATGTGGTGATAACAGATAAGTCGTAATGGCTAACTTACCTTTCTTTGT	EBS1d-*sigE*-453s
CT006	TGAACGCAAGTTTCTAATTTCGATTAAATCTCGATAGAGGAAAGTGTCT	EBS2-*sigE*-453s
CT007	AAAAAAGCTTATAATTATCCTTAGATGCCATATTCGTGCGCCCAGATAGGGTG	IBS-*sigG*-546s
CT008	CAGATTGTACAAATGTGGTGATAACAGATAAGTCATATTCGTTAACTTACCTTTCTTTGT	EBS1d-*sigG*-546s
CT009	TGAACGCAAGTTTCTAATTTCGATTGCATCTCGATAGAGGAAAGTGTCT	EBS2-*sigG*-546s
CT010	AAAAAAGCTTATAATTATCCTTAGCAAACGATACAGTGCGCCCAGATAGGGTG	IBS-*sigK*-102s
CT011	CAGATTGTACAAATGTGGTGATAACAGATAAGTCGATACATTTAACTTACCTTTCTTTGT	EBS1d-*sigK*-102s
CT012	TGAACGCAAGTTTCTAATTTCGATTTTTGCTCGATAGAGGAAAGTGTCT	EBS2-*sigF*-102s
CT025	AAAAAAGCTTATAATTATCCTTAGATATCAGACTTGTGCGCCCAGATAGGGTG	IBS-*spoIIR*-39a
CT026	CAGATTGTACAAATGTGGTGATAACAGATAAGTCAGACTTAATAACTTACCTTTCTTTGT	EBS1d-*spoIIR*-39a
CT027	TGAACGCAAGTTTCTAATTTCGATTATATCTCGATAGAGGAAAGTGTCT	EBS2-*spoIIR*-39a
LS171	AAAAAAGCTTATAATTATCCTTAATAGTCGTAGCAGTGCGCCCAGATAGGGTG	IBS-*spoIID*-39s
LS172	CAGATTGTACAAATGTGGTGATAACAGATAAGTCGTAGCAAATAACTTACCTTTCTTTGT	EBS1d-*spoIID*-39s
LS173	TGAACGCAAGTTTCTAATTTCGGTTACTATCCGATAGAG GAAAGTGTCT	EBS2-*spoIID*-39s
LS174	AAAAAAGCTTATAATTATCCTTACATTTCATTCACGTGCGCCCAGATAGGGTG	IBS-*spoVT*-157a
LS175	CAGATTGTACAAATGTGGTGATAACAGATAAGTCATTCACCTTAACTTACCTTTCTTTGT	EBS1d-*spoVT*-157a
LS176	TGAACGCAAGTTTCTAATTTCGGTTAAATGTCGATAGAGGAAAGTGTCT	EBS2-*spoVT*-157a
pMTL007-F	TTAAGGAGGTGTATTTCATATGACCATGATTACG	Sequencing pMTL007
pMTL007-R	AGGGTATCCCCAGTTAGTGTTAAGTCTTGG	Sequencing pMTL007
Primers to confirm proper type I intron splicing and type II intron insertion		
ErmRAM-F	ACGCGTTATATTGATAAAAATAATAATAGTGGG	PCR *erm* intron
ErmRAM-R	ACGCGTGCGACTCATAGAATTATTTCCTCCCG	PCR *erm* intron

[a] Underlined primers were also used for transcriptional fusions to *SNAP*[CD]
[b] Introduced restriction sites are underlined.

Table A2 – Oligonucleotide primers used in this study (cont.)

Primer	Sequence (5' to 3')	Description
CDsigF_616 Rev	GACGCTCTCTCTTATCCAG	*sigF* 3'-intron
CDsigE_331 Fw	GCAACTTATGCTTCGAGATG	*sigE* 5'-intron
CDsigE_625 Rev	CCAAGCATTCCAGCAACTTC	*sigE* 3'-intron
CDsigG_99 Fw	GGTGATGAAGAAGCCAGAC	*sigG* 5'-intron
CDsigG_700 Rev	CATCAGCAACTTCTATCTGAG	*sigG* 3'-intron
CDsigK_35 Fw	CCCCTGAAGAAGAGATTGAG	*sigK* 5'-intron
CDsigK_238 Rev	GCATATGTTGCTAATCGAGTT	*sigK* 3'-intron
IMV649	AAATTAAGATTAGATTGGGACTA	5' CD3564-*spoIIR*
LS184	ATCTCATATAGAGGAAAGG	5' CD0126-*spoIIID*
LS185	CATTTTTAACTTCTTTGGC	3' CD0126-*spoIIID*
LS186	TTCACAGCAAAAGATGGAG	5' CD3499-*spoVT*
LS187	ACACTCCATATCCTAGTAC	3' CD3499-*spoVT*
OBD522	ATCTGTAGGAGAACCTATGGGAAC	Southern blot probe
OBD523	CACGTAATAAATATCTGGACGTAAAA	Southern blot probe
Primers used in complementation studies		
M13R	CAGGAAACAGCTATGAC	Sequencing pMTL84121
M13F	GTTTTCCCAGTCACGAC	Sequencing pMTL84121
PCDsigF Fw	GCTGCCGCCGCCGAGCTTTATTCAGA	5'-promoter *sigF*-NotI
PCDsigF Rev	GGCAACAGTTACTTCCATCAAAAATCCCTCCTT	3'-promoter *sigF*
CDsigF Fw	AAGGAGGGATTTTTGATGGAAGTAACTGTTGCC	5'-*sigF*
CDsigF_EcoRI Rev	GCTGAATTCTCACGATATATATTC	3'-*sigF*-EcoRI
PCDsigE Fw	GCTGCGGCCGCGACCATTGACAGTTGGAC	5'-promoter *sigE*-NotI
PCDsigE Rev	TTCTTTCAATCGTAACATGTATCCCCCTTAAAACC	3'-promoter *sigE*
CDsigE Fw	GGTTTTAAGGGGGGATACATGTTACGATTGAAAGAA	5'-*sigE*
CDsigE_EcoRI Rev	GCTGAATTCTTATACAAATTTTTCAT	3'-*sigE*-EcoRI
PCDsigG Fw	GCTGCGGCCGCCGAGATGTATAGAAAATG	5'-promoter *sigG*-NotI
CDsigG_EcoRI Rev	GCTGAATTCTTATACATATTTTCTCAT	3'-*sigG*-EcoRI
PCDsigK5' Fw	GCTGCGGCCGCTGACTGATACTTTTG	5'-promoter *sigK*-NotI
CDsigK5' Rev	CACTTTTCTTAAATCACTAGCTATAGAAATTGATGA	3'-promoter *sigK*
CDsigK3' Fw	TCATCAATTTCTATAGCTAGTGATTTAAGAAAAGTG	5'-*sigK*
CDsigK3'_EcoRI Rev	GCTGAATTCTTAACTTTCTTGAACAAG	3'-*sigK*-EcoRI
IMV641	CGGGATCCTTAATATTTTCTGTTGGTTC	3'BamHI-*spoIIR*
IMV642	CCGCTCGAGATAAGGCTATGATTAGT	5'XhoI-*spoIIR*
IMV644	CCGCTCGAGGTAGAAAAAGATGAAAAAG	5'XhoI-*spoVT*
IMV648	CGGGATCCTACGATTACACCTGCTA	3'BamHI-*spoVT*
IMV647	CCGCTCGAGCAATGAGTCTAATTAAGT	5'XhoI-*spoIIID*
IMV649	CGGGATCCAAATCTCGCTTTAATCAA	3'BamHI-*spoIIID*
Psp17 Fw	GCTGTCGACGCTAGAATAGACAATGAG	5'-promoter *sp17* SalI
sp17 Rev	CAATTACTCGAGTTATTTTCTACAGCAGTT	3'-*sp17* XhoI

Table A2 – Oligonucleotide primers used in this study (cont.)

Primer	Sequence (5´ to 3´)	Description
SNAPtag_SacI Fw	GCTCGAGAGCTCGGAGGAACTACTATGG	5´-*SNAPCd*-SacI
SNAPtag_BamHI Rev	GCTGTTGGATCCAAGCTTTCCTTACCC	3´-*SNAPCd*-BamHI
SNAP-linker BamHIFw	GGTGGTGGTGGATCCGCAGCTGCTGATAAAGATTGTGAAATGAAG	5´- *SNAPCd*-linker BamHI
CDspoIIQ40D_EcoRI	AAAAGAATTCATGCCGGAACTGTAG	5´-promoter *spoIIQ*-EcoRI
CDspoIIQR_BamHI	AGCTGCGGATCCACCACCACCAAGCTTAATTAGACTCATTGGG	3´-*spoIIQ*-linker BamHI
CDspoIIIAA-AH	CCCTTATAATTAAACTTCATCTCTTGTCCTCCTCCTTTG	3´-promoter *spoIIIAA*
CDspoIIIAAD	ATGAAGTTTAATTATAAGGGAAGAGG	5´-*spoIIIAH*
CDspoIIIAHR_BamHI	AGCTGCGGATCCACCACCACCAAGCTTATTACTATTATTTATTTGTAAG	3´-*cotB*-linker BamHI
PCDsigF-XhoI Rev	GGAATAATTCTCGAGCAAAAAATCCCTCCTT	3´-promoter *sigF*-XhoI
PCDsigE-XhoI Rev	CAATATACTCGAGGTATCCCCCCTTAAAACC	3´-promoter *sigE*-XhoI
PCDsigG-XhoI Rev	ATTAACCTCGAGATATTTTCAGTCCTCTCT	3´-promoter *sigG*-XhoI
PCDsigK-XhoI Rev	AAGAGCCTCGAGTAGGTCACCCCCTTTTCA	3´-promoter *sigK*-XhoI
Pgpr-EcoRI Fw	AAGAGCTATGAATTCGGGTATTTATCACTT	5´-promoter *gpr*-EcoRI
Pgpr-XhoI Rev	CATATAAACCTCGAGTATAAATTAAATGCT	3´-promoter *gpr*-XhoI
PspoIIIAA-EcoRI Fw	TAGATGGTGGAATTCCTAGGGCTTACCAAA	5´-promoter *spoIIIAA*-EcoRI
PspoIIIAA-XhoI Rev	GTTTATTCATCTCTTGCTCGAGTCCTTTG	3´-promoter *spoIIIAA*-XhoI
PsspA-EcoRI Fw	AGATGAGGAGAATTCGGATAAAAGAGTTCA	5´-promoter *sspA*-EcoRI
PsspA-XhoI Rev	CTTCCTTCTCTCGAGTTTATTTTGTGTTGC	3´-promoter *sspA*-XhoI
PcotB-EcoRI Fw	TACCTAGAATTCTAAGTGCAGCTATTAG	5´-promoter *cotB*-EcoRI
CotB Linker Rev	GCTGCGGATCCACCACCACCAAGCATGTTTTTATAACTCTCC	3´-*cotB*-linker BamHI
PcotE-EcoRI Fw	CATAACTGAATTCATTAACACTAATGC	5´-promoter *cotE*-EcoRI
CotE Linker Rev	GCTGCGGATCCACCACCACCAAGGAATTGCCCATAAATACCTTC	3´-*cotE*-linker BamHI
PcotE-XhoI Rev	CATGTAAATCACCTCGAGATAAATTATCATTT	3´-promoter *cotE*-XhoI
PspoIIR-SNAP EcoRIFw	AAGCACTTTGAATTCCGATGTAGATGCGTT	5´ - PspoIIR-EcoRI Fw
Psp17 Rev	TGCAGCCTCGAGTTATATAACATATCTCTCCC	3-promoter *sp17* XhoI
sp17 Linker Rev	GGATCCACCACCACCAAGTTTTCTACAGCAGTTACA	3´-*sp17*-linker
Primers used in qRT-PCR experiments		
DNApoIIII-F	TCCATCTATTGCAGGGTGGT	5´ CD1305-*dnaF*
DNApoIIII-R	CCCAACTCTTCGCTAAGCAC	3´ CD1305
CD1581-F	TGTGGATGTTCAGATTCAGGAG	5´ CD1581-sp17
CD1581-R	TGCTAATGCCTCTTTCTTTTTC	3´ CD1581
IMV552	CCAGGAGCAAAAGAGGCTTTA	5´ CD2688-*sspA*
IMV553	TCCAGCCATTTGTTGTTCAG	3´ CD2688
LS100	CACCACCTCAATGTGGAAAA	5´ CD1192-*spoIIIAA*
LS101	GCTCCTGCTATCTCATTACGC	3´ CD1192
LS114	TGCATTGGCTTCAAGAAAAA	5´ CD2470-*gpr*
LS115	GCAATAACATCCACGCCTAAA	3´ CD2470
LS143	CAGGCCCAAATGGTAGAAAA	5´ CD1433-*cotE*
LS144	GAAGGCATTCCAGCATTCTC	3´ CD1433

Table A2 – Oligonucleotide primers used in this study (cont.)

Primer	Sequence (5´ to 3´)	Description
LS157	GCAGTAGGTGGAGTTTGGTTT	3' CD0125
LS102	GGGCCATAGTGGTAGCAAAA	5' CD0126
LS103	TGGCAAGGGATGGATTTATT	3' CD0126
LS206	GCGGAGGAGTACTTTCTGGA	5' CD0332-*bclA1*
LS207	GCTGATTGCCCATTTCGTAT	3' CD0332
LS147	CCAATGTCCTGACCAAATGA	5' CD0551-*sleC*
LS148	CTGCCCAGAAGAACCAACTT	3' CD0551
LS158	TTAAGGCTGCTGGACTTGGT	5' CD0598-*cotCB*
LS159	AGTTACCGAATCGCCAAAGA	3' CD0598
LS218	AATCCAGCTCCTGTCTTTGC	5' CD0684
LS219	CGTCTAATCCAGCAACGTCTC	3' CD0684
IMV550	TTGGAAGGGTAGTTATCCCAAA	5' CD3499-*spoVT*
IMV551	GCATATTCTTGCGAAAATTCA	3' CD3499
LS151	AGCATCACCAAATCCAATCC	5' CD1067
LS152	AACTTGGCTTTCCACTTCCA	3' CD1067
LS166	TTGATATGGCCCACAAAGAA	5'CD1133
LS167	CACTTTCTTGATATGCACACTTTTT	3' CD1133
LS100	CACCACCTCAATGTGGAAAA	5' CD1192-*spoIIIAA*
LS101	GCTCCTGCTATCTCATTACGC	3' CD1192
LS155	GAAAAACCCTTAACCCCTGA	5' CD1230-*sigK*
LS156	TCATCCTGATCTTCCGTTGA	3' CD1230
LS214	TTCAAAAAGCCATCCATCAA	5' CD1430
LS215	TCGCTAAATTTACCCATTGGA	3' CD1430
LS143	CAGGCCCAAATGGTAGAAAA	5' CD1433-*cotE*
LS144	GAAGGCATTCCAGCATTCTC	3' CD1433
LS114	TGCATTGGCTTCAAGAAAAA	5' CD2470-*gpr*
LS115	GCAATAACATCCACGCCTAAA	3' CD2470
LS220	TAATCGCTGAAGGGGGAAAGA	5' CD1486
LS221	CCAACATCTTCAGTTCCATGA	3' CD1486
LS141	TTGCTATGGAAGCTGACCACT	5' CD1511-*cotB*
LS142	TTGTTCTGCAACTGATGTGAGA	3' CD1511
LS145	GCTGCATCTACTCCATTAGCAA	5' CD1613-*cotA*
LS146	GCAATCATCACAATCGCAGT	3' CD1613
LS232	GGATTTTTGGGAATAATTTCAAG	5' CD1940
LS233	TTTTGTCGGATGTTTCTGGA	3' CD1940
LS226	GGGCAAGAGAGGAGTTTGAA	5' CD2000
LS227	TCCATTTCCTTGTTCCGTTC	3' CD2000
LS212	AAAGTAGATGGAATGCAAGCAA	5' CD2375
LS213	CATTTTACAGCTTCACAATGAATG	3' CD2375
LS160	AAAGCTGCAGGTCTTGGTTC	5' CD2401-*cotD*

Table A2 – Oligonucleotide primers used in this study (cont.)

Primer	Sequence (5´ to 3´)	Description
LS106	CCCCAAAGTGGTTCAGGTAA	5' CD2629-*spoIVA*
LS107	TGCCCTAAAGCTCCTTCAAC	3' CD2629
IMV552	CCAGGAGCAAAAGAGGCTTTA	5' CD2688-*sspA*
IMV553	TCCAGCCATTTGTTGTTCAG	3' CD2688
LS208	AGCAGCAAACAATGCACAAT	5' CD3230-*bclA2*
LS209	TCCCTCCAACTGCAAGATTT	3' CD3230
IMV554	CAGTAGTTCCAGAAGCAAAAGC	5' CD3249-*sspB*
IMV555	TTCCAGCCATTTGTCTTTCA	3' CD3249
LS216	AGGTGCTAGTTCTGCAACATCA	5' CD0773-*spoVAC*
LS217	ACCAGCTCCAGCTCTTTTTC	3' CD0773
LS112	GTTCTTCCAGCTGGTGAATATAAA	5' CD3564-*spoIIR*
LS113	CAACAAAACATAGTGGAGGGAAC	3' CD3564
LS230	AACTGGGGTGCAATTGGTAT	5' CD3551.1
LS231	TCCAGCTAAACCTACAATGAAGT	3' CD3551.1
LS164	CGTTGACGAAATAGCACCTG	5' CD3580
LS165	TGGGGAAATACAATAGAAGCGTA	3' CD3580
LS200	GGTCATGCCATAAAAGATTCG	5' CD1213-*spoIVB*
LS201	CCCATTAAATTGCTGTCATTAAAA	3' CD1213
LS245	CAAACAATACCCATCAAGACCA	5' CD2636
LS246	GTCAAGGCTGCAAACTCCTT	3' CD2636
LS243	TTTGGGGTTTTAGGAGCTGT	5' CD0792
LS244	GGCTTTTTATCGGATTTTGC	3' CD0792
LS239	TGGATGATGTTTCAAGACAAAA	5' CD1290
LS240	AAACACCTTTATTAGACGCATTTTC	3' CD1290
LS257	TTCCCCAAGATGAATTGAGAA	5' CD0311
LS258	TGTTAGGCAAGTTGTTTGGAGA	3' CD0311
IMV617	TGGCAAAAGAGTTGGGTATG	5' CD1616
IMV618	CTTCAATTGGCATTGGCTTT	3' CD1616
LS247	AGGGCTTATGTGTGGAAATG	5' CD1746
LS248	TAATCCAACCCCAACTGCTC	3' CD1746
LS259	GGTGGTTCTCCAGGTGAAAT	5' CD3521
LS260	CCAGCTTCAGGTTCAAGACC	3' CD3521
LS261	TGGGATATTTAAAAGAGGATAAACAA	5' CD0596
LS262	TGCTGTTTTTGGATTGTATGTG	3' CD0596
LS241	GTATGGGCTGAAATGCCTGT	5' CD0580-*gapN*
LS242	CTTCTGCCCTTGCAGATTTT	3' CD0580
LS255	AAACAAGGGCATTTTCAAGAAG	5' CD2856
LS256	TGCAAATATATTATTGGGGTCAA	3' CD2856
LS249	GCTTTTAGCTCAAACTGGTGGT	5' CD0783- *spoIVB'*
LS250	CCTACAGCTGGGTCGTTTTC	3' CD0783
LS354	CCTGGTGGATATGTCGCTTT	3' CD1229

Table A3. The forespore line of expression with the identification of the σ^F and σ^G regulons

Gene	Name	Function	Expression ratio[a]	Promoter[b]	Detection in spore[c]	Operon
Members of the σ^F regulon			**sigF/630Δ*erm***			
Sporulation						
CD0125	*spoIIQ*	Stage II sporulation protein Q	0.11	σ^F	-	
CD2470	*gpr*	Spore endopeptidase	0.22	σ^F	+	*CD2470-CD2468*
CD2469	*spoIIP*	Stage II sporulation protein P	0.14		-	
CD2468		Conserved hypothetical protein	0.28		-	
CD3564	*spoIIR*	Pro-sigma(E) endopeptidase (stage II)	0.43	σ^F	-	*CD3564-CD3563*
CD3563	*sleB*	Spore-cortex-lytic protein	0.16		-	
CD0783	*spoIVB'*	SpoIVB, S55 peptidase family	0.10		-	
Enveloppes						
CD2141		Serine-type D-Ala-D-Ala carboxypeptidase	0.30		-	
CD1229		Putative peptidoglycan glycosyltransferase	0.42	σ^F	-	
CD2686		Putative membrane protein	0.33	σ^F	-	
CD2856		Putative membrane protein	0.46	σ^F	-	
CD2107		Xanthine/uracil/thiamine permease family	0.37		-	
CD2102		Putative Na(+)/H(+) antiporter	0.49	σ^A	-	
Miscellaneous						
CD0580	*gapN*	Glyceraldehyde-3-phosphate dehydrogenase	0.36	σ^F	+	
CD3595		Aminopeptidase	0.55	σ^F	+	
CD2661		Putative peptidase, M16 family	0.47		+	*CD2661-CD2660*
CD2660		Putative peptidase, M16 family	0.48		+	
CD0047	*ispD*	2-C-methyl-D-erythritol 4-phosphate	0.48	σ^A	-	*CD0047-CD0048*
CD0048	*ispF*	2-C-methyl-D-erythritol 2,4-cyclodiphosphate	0.50		-	
CD3220		Putative methyltransferase	0.39	σ^A	-	
CD0761		Putative ATP-dependent RNA helicase	0.46	σ^F and σ^A	+	
CD1323	*tepA*	Protein export-enhancing factor	0.38		+	
Proteins of unknown function						
CD0347		Conserved hypothetical protein	0.35	σ^F	-	*CD0347-CD0048*
CD0348		Fragment of conserved hypothetical protein	0.35		-	
CD2687		Conserved hypothetical protein	0.36		-	

Table A3. The forespore line of expression with the identification of the σ^F and σ^G regulons (cont.)

Gene	Name	Function	Expression ratio[a]		Promoter[b]	Detection in spore[c]	Operon
Members of the σ^G regulon			**sigF/630Δ*erm***	**sigG/630Δ*erm***			
Sporulation							
CD0773	*spoVAC*	DPA uptake protein, SpoVAC	0.24	0.26	σ^3	-	*CD0773-CD0775*
CD0774	*spoVAD*	DPA uptake protein, SpoVAD	0.27	0.30		+	
CD0775	*spoVAE*	DPA uptake protein, SpoVAE	0.32	0.28		-	
CD2688	*sspA*	Small, acid-soluble spore protein alpha	0.00	0.02	σ^3	+	
CD3249	*sspB*	Small, acid-soluble spore protein alpha	0.01	0.07	σ^3	+	
CD1290		Putative small acid-soluble spore protein SASP	0.29	0.52	σ^3	-	
CD3220.1		Small acid-soluble spore protein	0.28	0.66	σ^3	+	
CD3499	*spoVT*	Transcriptional regulator, SpoVT	0.07	0.22	σ^F/σ^G?	+	
CD1213	*spoIVB*	Stage IVB sporulation protein B, S55 peptidase		0.6	σ^3	-	
Envelopes							
CD1430	*pdaA*	Putative d-lactam-biosynthetic deacteylase	0.06	0.21	σ^G	-	
CD1291	*dacF*	D-alanyl-D-alanine carboxypeptidase	0.11	0.31	σ^G	+	
CD0784		Putative N-acetylmuramoyl-L-alanine amidase	0.23	0.41		-	
CD2762	*uppS*	Putative undecaprenyl pyrophosphate synthetase	0.20	0.62	σ^G	-	
CD0793		Putative membrane protein, DUF81 family	0.26	0.33		-	*CD0793-CD0792*
CD0792		Putative membrane protein, DUF81 family	0.39	0.43		-	
CD2636		Putative membrane protein	0.20	0.34	σ^G?	-	*CD2636-CD2634*
CD2635		Putative membrane protein	0.33	0.50		+	
CD2634		Conserved hypothetical protein	0.27	0.65		+	
CD2315		Putative exported protein	0.31	0.38		-	
CD3551.1		Putative membrane protein	0.11	0.18	σ^G	-	
CD2051		Putative membrane protein		0.52		-	
CD1789		Putative membrane protein, DUF421 family		0.61	σ^G	+	*CD1789-CD1788*
CD1788		Conserved hypothetical protein		0.43		-	
CD2465		Putative amino acid/polyamine transporter	0.36	0.52	σ^G	+	
Stress							
CD1567	*cotG*	putative manganese catalase	0.25	0.27	σ^G	+	
CD2845	*rbr*	Rubrerythrin	0.14	0.27	σ^G	+	
CD1631	*sodA*	Superoxide dismutase (Mn)	0.06	0.32		+	

Table A3. The forespore line of expression with the identification of the σ^F and σ^G regulons (cont.)

Gene	Name	Function	Expression ratio[a]		Promoter[b]	Detection in spore[c]	Operon
Members of the σ^G regulon			**sigF/630Δ*erm***	**sigG/630Δ*erm***			
Miscelaneous							
CD1486		Putative ribosome recycling factor	0.10	0.17	σ^{G}?	+	
CD0684		Putative ATP-dependent peptidase, M41 family	0.29	0.46	σ^{G}	+	
CD1543		putative FMN-dependent NADH-azoreductase	0.16	0.43	σ^{G}	+	
CD2431		Putative nitrite/sulphite reductase	0.23	0.54		+	
CD1707		Putative C4-dicarboxylate anaerobic carrier	0.29	0.53	σ^{G}	+	
CD3489		Putative oligoendopeptidase F, peptidase M3B …		0.62	σ^{G}	-	
CD1676	*pcp*	Pyrrolidone-carboxylate peptidase		0.66		+	*CD1677-CD1676*
CD1677		Putative membrane protein	0.48	0.51		-	
CD2841		Putatie amidohydrolase		0.57	σ^{G}	+	
Proteins of unknown function							
CD0543		Conserved hypothetical protein	0.38	0.57	σ^{G}	-	
CD1297		"	0.31	0.66		-	
CD1298		"	0.19	0.55		-	
CD1301.1		"	0.28	0.45		-	
CD1354		"	0.22	0.48	σ^{G}	-	
CD1463		"	0.06	0.22	σ^{G}	+	
CD1880		"	0.07	0.25	σ^{G}	+	
CD2112		"	0.03	0.26	σ^{G}	+	
CD2245.1		"	0.10	0.30	σ^{G}	+	
CD2375		"	0.14	0.24	σ^{G}	+	
CD2808		"	0.05	0.64	$\sigma^{?}$	-	
CD2809		"	0.05	0.45	$\sigma^{?}$?	+	
CD1568		"		0.60	$\sigma^{?}$	-	
CD3610		"		0.66	$\sigma^{?}$	+	

[a] A gene is considered differentially expressed when the P value is <0.05 using the statistical analysis described in Materials and Methods.
[b] We searched for promoters recognized by the different sigma factors active in forespore upstream of the TSS mapped by RNA-seq. Promoters with the -10 and -35 elements less conserved are indicated with a question mark.
[c] Proteins that are detected associated with the spore by a proteomic approach (Lawley *et al.*, 2009).

Table A4. The mother cell line of expression with the identification of the σ^E and σ^K regulons

Gene	Name	Function	Expression ratio[a]		Promoter[b]	Detection in spore[c]	Operon
			sigE/630Δ*erm*	*spoIIID*/630Δ*erm*			
Members of the σ^E regulon							
Sporulation							
CD0124	*spoIID*	peptidoglycan hydrolase SpoIID	0.24		σ^E	-	
CD1192	*spoIIIAA*	ATP-binding stage III sporulation protein	0.05	3.9[d]	σ^E	-	*spoIIIAA-spoIIIAH*
CD1193	*spoIIIAB*	Stage III sporulation protein AB	0.05	3.8		-	
CD1194	*spoIIIAC*	Stage III sporulation protein AC	0.03	4.9		-	
CD1195	*spoIIIAD*	Stage III sporulation protein AD	0.07	3.4		-	
CD1196	*spoIIIAE*	Stage III sporulation protein AE	0.09	2.4		-	
CD1197	*spoIIIAF*	Stage III sporulation protein AF	0.08	3.6		-	
CD1198	*spoIIIAG*	Stage III sporulation protein AG	0.02	7.8[d]	σ^E	-	
CD1199	*spoIIIAH*	Stage III sporulation ratchet engulfment protein	0.05	3.3		-	
CD0126	*spoIIID*	Transcriptional regulator SpoIIID	0.04		σ^E	-	
CD2629	*spoIVA*	SpoIVA, assembly of cortex and coat layers	0.02	3.5	σ^E	+	
CD3657	*sipL*	Functional homologue of SpoVID	0.05	3.0	σ^E	+	
CD2443		Conserved hypothetical protein YqfC-like	0.05		σ^E	+	*CD2443-CD2442*
CD2442	*spoIV*	Stage IV sporulation protein YqfD-like	0.14			-	
CD1230	*sigK*	Fragment of RNA polymerase sigma-K factor	0.13		σ^E and σ^K	-	
CD1234		Putative phage protein, skin element	0.22		σ^E	-	
CD0106	*cwlD*	N-acetylmuramoyl-L-alanine amidase	0.14		σ^E	-	
CD2247	*cspBA*	Subtilisin-like serine germination related protease	0.11		σ^E	+	*cspAB-cspC*
CD2246	*cspC*	Subtilisin-like serine germination related protease	0.16			+	
CD3542	*spmA*	Spore maturation protein A	0.16		σ^E	-	*spmA-spmB*
CD3541	*spmB*	Spore maturation protein B	0.11	2.0		-	
CD1511	*cotB*	Spore outer coat layer protein CotB	0.05	3.5	σ^E	+	
CD0213		Putative spore coat protein	0.42			-	*CD0214-CD0213*
CD0214		Conserved hypothetical protein	0.23			+	
CD1320		Putative peptidase, M16 family	0.25			-	*CD1320-CD1322*
CD1321		Putative sporulation protein	0.23			-	
CD1322	*dapG*	Aspartokinase 1 (Aspartokinase I)	0.32			-	
CD2641		Putative sporulation protein	0.13		σ^E	-	*CD2641-CD2639*
CD2640	*nrdR*	Transcriptional regulator, repressor NrdR family	0.23			-	
CD0782		Putative sporulation protein YunB	0.32			-	
CD1168		YbJ-like protein/ spore cortex formation	0.18			-	
CD3455		Carboxy-terminal protease, homolog of CtpB	0.25		σ^E	-	

Table A4. The mother cell line of expression with the identification of the σ^E and σ^K regulons (cont.)

Gene	Name	Function	Expression ratio[a]		Promoter[b]	Detection in spore[c]	Operon
Members of the σ^E regulon			***sigE*/630Δ*erm***	***spoIIID*/630Δ*erm***			
CD3494		Putative spore protein	0.20		σ^E	-	*CD3494-CD3493*
CD3493		Putative membrane protein	0.18			-	
Envelopes							
CD3464		Conserved hypothetical protein	0.04		σ^E	-	*CD3464-CD3463*
CD3463	*alr2*	Alanine racemase 2	0.05			+	
CD2761		Putative N-acetylmuramoyl-L-alanine amidase	0.40			-	
CD2833		Putative calcium-transporting ATPase	0.22	2.3[d]	σ^E	-	
CD0760		Putative Ca2+/Na+ antiporter	0.20		σ^E	-	
CD3483		Putative zinc/iron permease	0.34		σ^E	-	
CD2445		Transmembrane signaling protein, TspO family	0.07	2.4	σ^E	-	
CD0131		Putative membrane protein	0.18			-	
CD0314		"	0.15		σ^E	-	
CD1301		"	0.33			-	
CD1416		"	0.20	2.2	σ^E	-	
CD1928		"	0.14	3.0[d]		-	*CD1928-CD1929*
CD1929		"	0.27			-	
CD1940		"	0.08		σ^E	-	
CD2800		"	0.11	2.6	σ^E	-	
CD3636		"	0.42			-	*CD3636-CD3635*
CD3635		Conserved hypothetical protein	0.29			-	
Metabolism							
CD3638		Conserved hypothetical protein	0.21		σ^E	-	*CD3638-CD3637*
CD3637		Putative NADPH-dependent FMN reductase	0.27			-	
CD2429.1		4Fe-4S ferredoxin	0.27		σ^E	-	*CD2429.1-CD2428*
CD2429		Flavodoxin/ferredoxin oxidoreductase a subunit	0.31			-	
CD2428		Flavodoxin/ferredoxin oxidoreductase b subunit	0.38			-	
CD3251		Putative dehydrogenase	0.25		σ^E	-	
CD3258		Iron hydrogenase	0.08	2.3[d]	σ^E	-	

Table A4. The mother cell line of expression with the identification of the σ^E and σ^K regulons (cont.)

Gene	Name	Function	Expression ratio[a]		Promoter[b]	Detection in spore[c]	Operon
			sigE/630Δ*erm*	*spoIID*/630Δ*erm*			
Members of the σ^E regulon							
CD2000	*isp*	Intracellular serine protease	0.18		σ^E	-	
CD3652		Putative peptidase, M1 family	0.12			+	
CD3521		Putative peptidase T, M20B family	0.33		σ^E	+	
CD1085		Putative membrane protein	0.26		σ^A	-	*CD1085-CD1086*
CD1086		Putative peptidase, M20D family	0.26			-	
CD1555		Putative amino acid permease	0.22			-	
CD1746	*gltC*	Sodium/glutamate symporter	0.20		σ^E	-	
CD1259	*brnQ-1*	Branched chain amino acid transporter	0.39			-	
CD2439		Diacylglycerol kinase/undecaprenol kinase	0.45			+	
CD1068		Polysaccharide biosynthesis/sporulation protein	0.23			-	
CD3257		Polysaccharide deacetylase	0.11	2.1		-	
CD3248		"	0.09	2.3[d]	σ^E	-	
CD1319		"	0.06	3.2[d]	σ^E	+	
CD0982	*ubiA*	Putative UbiA prenyltransferase	0.39			-	
Miscellaneous							
CD3235	*ssb*	Single-stranded DNA-binding protein	0.03	3.6[d]	σ^E	-	
CD1167	*recV*	Tyrosine DNA recombinase, XerC/XerD family	0.26		σ^E	-	
CD2864		Putative hydrolase	0.08	4.0	σ^E	-	
CD3298		Putative ATP/GTP-binding protein	0.19		σ^E	-	
CD3462	*mazE*	Putative antitoxin EndoAI	0.05			-	*CD3641-CD3642*
CD3461	*mazF*	Endoribonuclease toxin	0.11			-	
CD2865		Putative bacterioferritin	0.19	2.2[d]	σ^E	+	
CD0757		Putative diguanylate kinase signaling protein	0.45			-	
CD1616		Putative diguanylate kinase signaling protein	0.37			-	
CD2637		Two-component sensor histidine kinase	0.30			-	
Proteins of unknown function							
CD0129		Conserved hypothetical protein, DUF1256 family	0.15		σ^E	-	
CD2121		Conserved hypothetical protein	0.23			-	
CD0311		"	0.02	4.1	σ^E	-	
CD2374		"	0.23			-	

Table A4. The mother cell line of expression with the identification of the σ^E and σ^K regulons (cont.)

Gene	Name	Function	Expression ratio[a]		Promoter[b]	Detection in spore[c]	Operon
Members of the σ^E regulon			***sigE*/630Δ*erm***	***spoIIID*/630Δ*erm***			
CD1884		Conserved hypothetical protein	0.11		σ^E	-	
CD1726		"	0.05	3.9[d]	σ^E	-	
CD3457		"	0.05	2.6[d]	σ^E	-	
CD3465		"	0.40			-	*CD3466-CD3465*
CD1066		"	0.16			-	
CD3522		"	0.09		σ^E	+	
CD2816		"	0.11		σ^E	-	
CD3234		"	0.09			-	
CD1930		"	0.27			+	
CD1063		"	0.11	3.0	σ^E	-	
Members of the σ^K regulon			***sigE*/630Δ*erm***	***sigK*/630Δ*erm***			
Sporulation							
CD1613	*cotA*	Spore outer coat layer protein CotA	0.10	0.05	σ^K	+	
CD0596		Conserved hypothetical protein		0.02	σ^K?[e]	+	*CD0596-cotCB*
CD0597	*cotJB1*	Spore coat peptide assembly protein		0.02		+	
CD0598	*cotCB*	Spore-coat protein CotCB manganese catalase	0.18	0.01		+	
CD2399		Conserved hypothetical protein	0.14	0.01	σ^K?[e]	+	*CD2399-CD2401*
CD2400	*cotJB2*	Spore coat peptide assembly protein CotJB 2	0.10	0.01		+	
CD2401	*cotD*	Spore coat protein CotD manganese catalase	0.11	0.01		+	
CD1433	*cotE*	Spore coat protein CotE peroxiredoxin/chitinase	0.14	0.03	σ^K	+	
CD0551	*sleC*	Spore cortex-lytic enzyme pre-pro-form	0.15	0.14	σ^K	+	
CD0332	*bclA1*	Putative exosporium glycoprotein	0.21	0.08	σ^K?[e]	+	
CD3230	*bclA2*	Putative exosporium glycoprotein	0.12	0.04	σ^K?[e]	-	
CD3349	*bclA3*	Exosporium glycoprotein BclA3	0.01	0.02	σ^K?[e]	-	
CD2968	*dpaA*	Dipicolinate synthase subunit A	0.30	0.08	σ^K	-	*CD2968-CD2967*
CD2967	*dpaB*	Dipicolinate synthase subunit B	0.29	0.09		-	
CD3569		Sporulation-specific protease, YabG-like protein		0.22	σ^K	-	
Envelopes							
CD0119	*glmM*	Phosphoglucosamine mutase		0.49		-	
CD0120	*glmS*	Glucosamine-fructose-6P aminotransferase		0.52	σ^K/σ^A ?	-	

Table A4. The mother cell line of expression with the identification of the σ^E and σ^K regulons (cont.)

Gene	Name	Function	Expression ratio[a]		Promoter[b]	Detection in spore[c]	Operon
			sigE/630Δ*erm*	*sigK*/630Δ*erm*			
Members of the σ^K regulon							
CD2664	*murE*	UDP-N-acetylmuramyl-tripeptide synthetase	0.45	0.42	$\sigma^{?}/\sigma^{K}$?	-	
CD3007		Putative L,D-transpeptidases	0.37	0.48		+	
CD2184		Putative N-acetylmuramoyl-L-alanine amidase	0.21	0.13	σ^K	-	
CD1898		Putative N-acetylmuramoyl-L-alanine amidase		0.12	σ^K	+	*CD1898-CD1897*
CD1897		Conserved hypothetical protein		0.13		-	
CD0514	*cwpV*	Cell surface protein	0.63	0.20		+	
CD1518	*feoA*	Ferrous iron transport protein		0.05	σ^K	-	*CD1518-CD1517*
CD1517	*feoB*	Ferrous iron transport protein B		0.05		-	
CD1904		ABC-type transport system, permease	0.26	0.13		-	
CD1927	*yttC*	Putative ABC transporter component, ATP-binding	0.3	0.08	σ^K	-	
CD1891		Fragment of ABC-type transport system, SBP	0.26	0.08	σ^K	-	
CD2720		Putative transporter		0.16		-	
CD0902		Putative cation efflux protein	0.29	0.17	σ^K	-	
CD2144		Putative membrane protein	0.46	0.12	σ^K	-	
CD2346		Putative membrane protein		0.24		-	
CD2459		Putative glucokinase, ROK family		0.41	σ^K	-	*CD2459-CD2458*
CD2458		Transporter, Major Facilitator Superfamily		0.39		-	
Miscellaneous							
CD3032		Putative phosphate-dependent transferase	0.32	0.11		+	
CD0309		Putative hydrolase, HAD superfamily, subfamily IB		0.38		-	
CD3350		Putative glycosyl transferase, family 2		0.09		-	
CD1951		Putative Acyl-CoA N-acyltransferase		0.19		-	
CD0749		Putative DNA helicase, UvrD/REP type	0.32	0.05		-	
CD2537		Membrane-associated 5'-nucleotidase/phosphoesterase	0.15	0.22		-	
CD1845		Putative membrane protein Tn1549-like,	0.14	0.17		-	*CD1845-CD1846*
CD1846		Putative conjugative protein Tn1549-like	0.15	0.20		-	
CD0564		Putative ATP-dependent protease		0.39	σ^K	-	
Proteins of unknown function							
CD0196		Fragment of conserved hypothetical protein	0.18	0.18		-	
CD0896		Conserved hypothetical protein	0.05	0.02	σ^K	-	

Table A4. The mother cell line of expression with the identification of the σ^E and σ^K regulons (cont.)

Gene	Name	Function	Expression ratio[a]		Promoter[b]	Detection in spore[c]	Operon
Members of the σ^K regulon			***sigE*/630Δ*erm***	***sigK*/630Δ*erm***			
CD1063.1		Conserved hypothetical protein	0.17	0.21	σ^K	+	
CD1063.2		"	0.01	0.04	σ^K	-	*CD1063.2-CD1063.3*
CD1063.3		"	0.01	0.04		-	
CD1065		"	0.06	0.30	σ^K?[e]	-	
CD1067		"	0.00	0.03	σ^K	+	
CD1133		"	0.12	0.16	σ^K	+	
CD1581		"	0.01	0.01		+	
CD2055		"		0.33		-	
CD2409		"		0.11	σ^K	-	
CD3580		"	0.11	0.12	σ^K	+	
CD3613		"		0.04	σ^K	+	
CD3620		Conserved hypothetical protein, similar to YmaF	0.34	0.04		+	

[a] A gene is considered differentially expressed when the P value is <0.05 using the statistical analysis described in Materials and Methods.
[b] We searched for promoters recognized by the different sigma factors active in forespore upstream of the TSS mapped by RNA-seq. Promoters with the -10 and -35 elements less conserved are indicated with a question mark.
[c] Proteins that are detected associated with the spore by a proteomic approach (Lawley *et al.*, 2009).
[d] Using the consensus sequence recognized by SpoIIID in *B. subtilis* (Eichenberger *et al.*, 2004).
[e] σ^K? indicates the presence of sequences similar to σ^K consensus elements upstream of these genes but the TSS was not mapped.

Table A5. Validation of microarrays data by qRT-PCR on selected genes

Gene	Name	Transcriptome[a]	qRT-PCR[b]
SigF-controlled genes		***sigF*/630Δ*erm***	
CD0125	*spoIIQ*	0.11	0.08
CD0332	*bclA1*	0.43	0.23
CD0580	*gapN*	0.36	0.11
CD0783	*spoIVB*	0.10	0.12
CD0792		0.39	0.006
CD1067		0.07	0.22
CD1229		0.42	0.08
CD1290		0.29	0.006
CD1486		0.10	0.035
CD1613	*cotA*	0.13	0.04
CD1940		0.25	0.18
CD2470	*gpr*	0.22	0.001
CD2636		0.20	0.008
CD2856		0.46	0.24
CD3499	*spoVT*	0.07	0.004
CD3564	*spoIIR*	0.43	0.025
SigE-controlled genes		***sigE*/630Δ*erm***	
CD0126	*spoIIID*	0.04	0.001
CD0311		0.02	0.0002
CD0598	*cotCB*	0.18	0.09
CD1192	*spoIIIAA*	0.05	0.0003
CD1230	*sigK*	0.13	0.002
CD1511	*cotB*	0.05	0.009
CD1613	*cotA*	0.10	0.0035
CD1616		0.37	0.13
CD1746	*atlC*	0.20	0.0015
CD1940		0.08	0.002
CD2000	*lsp*	0.18	0.04
CD2629	*spoIVA*	0.02	0.004
CD3521		0.33	0.17
SigG-controlled genes		***sigG*/630Δ*erm***	
CD0684		0.46	0.11
CD0773	*spoVAC*	0.26	0.04
CD0792		0.43	0.015
CD1213	*spoIVB*	0.60	0.12
CD1290		0.52	0.02
CD1430	*pdaA*	0.21	0.03
CD1486		0.17	0.02
CD2375		0.24	0.06
CD2636		0.34	0.03
CD2688	*sspA*	0.02	0.02
CD3249	*sspB*	0.07	0.08
CD3499	*spoVT*	0.22	0.08
SigK-controlled genes		***sigK*/630Δ*erm***	
CD0332	*bclA1*	0.08	0.04
CD0551	*sleC*	0.14	0.011
CD0598	*cotCB*	0.01	0.003
CD1067		0.03	0.1
CD1133		0.16	0.15
CD1433	*cotE*	0.03	0.01
CD1613	*cotA*	0.05	0.002
CD2401	*cotD*	0.01	0.006
CD3230	*bclA2*	0.04	0.01
CD3580		0.12	0.07
SpoIIID-controlled genes		***spoIIID*/630Δ*erm***	
CD1067		0.03	0.0003
CD1192	*spoIIIAA*	**3.85**	30
CD1511	*cotB*	**3.54**	5.6
CD2629	*spoIVA*	**3.54**	11
CD2375		**5**	3
CD2864		**4**	9
CD3349	*bclA3*	0.19	0.07
CD3490	*spoIIE*	**3.14**	5.5
CD3499	*spoVT*	**2.26**	3.7

[a] A gene is considered to be differentially expressed when P value is <0.05. Genes significantly downregulated are in grey, while genes significantly upregulated are in bold.
[b] Results are from at least 2 independent experiments

Table A6. Complementation of the mutants inactivated for the sigma factors of sporulation

***sigF* mutant complementation**		***sigF*/630Δ*erm***	**Complemented *sigF*/630Δ*erm***
CD2469	*spoIIP*	0.001	12
CD3564	*spoIIR*	0.03	9
CD3499	*spoVT*	0.004	19
***sigE* mutant complementation**		***sigE*/630Δ*erm***	**Complemented *sigE*/630Δ*erm***
CD1198	*spoIIIAG*	0.002	2
CD0126	*spoIIID*	0.001	0.7
CD2629	*spoIVA*	0.004	2.5
CD1230	*sigK*	0.002	0.4
***sigG* mutant complementation**		***sigG*/630Δ*erm***	**Complemented *sigG*/630Δ*erm***
CD2688	*sspA*	0.002	0.7
CD3249	*sspB*	0.001	1.2
CD3499	*spoVT*	0.08	0.5
CD3551.1		0.006	1.3
***sigK* mutant complementation**		***sigK*/630Δ*erm***	**Complemented *sigK*/630Δ*erm***
CD2401	*cotD*	0.005	0.15
CD0551	*sleC*	0.011	0.25
CD1433	*cotE*	0.01	0.15
CD1613	*cotA*	0.0005	0.05

Table A7. List of genes controlled by SpoIIID in the transcriptome

Gene ID	Name	Function	Expression ratio[a]
sporulation			
CD1214	*spo0A*	Stage 0 sporulation protein A	**0.47**
CD2644	*spoIIGA*	Sigma-E factor processing peptidase	**2.36**
CD2643	*sigE*	RNA polymerase sigma-E factor	**1.85**
CD2642	*sigG*	RNA polymerase sigma-G factor protein	**2.72**
CD3490	*spoIIE*	Phosphoprotein phosphatase	**3.14**
CD3563	*sleB*	Spore-cortex-lytic protein	**1.94**
CD0125	*spoIIQ*	Stage II sporulation protein Q	**3.02**
CD1192	*spoIIIAA*	Stage III sporulation protein AA	**3.85**
CD1193	*spoIIIAB*	Stage III sporulation protein AB	**3.81**
CD1194	*spoIIIAC*	Stage III sporulation protein AC	**4.91**
CD1195	*spoIIIAD*	Stage III sporulation protein AD	**3.44**
CD1196	*spoIIIAE*	Stage III sporulation protein AE	**2.42**
CD1197	*spoIIIAF*	Stage III sporulation protein AF	**3.63**
CD1198	*spoIIIAG*	Stage III sporulation protein AG	**7.75**
CD1199	*spoIIIAH*	Stage III sporulation protein AH	**3.31**
CD2629	*spoIVA*	Stage IV sporulation protein A	**3.54**
CD3567	*sipL*	SpoIVA-Interacting protein, coat localization	**3.06**
CD3541	*spmB*	Spore maturation protein B	**2.04**
CD1511	*cotB*	Spore outer coat layer protein CotB	**3.54**
CD2688	*sspA*	Small, acid-soluble spore protein alpha	**13.45**
CD3249	*sspB*	Small, acid-soluble spore protein alpha	**11.62**
CD3499	*spoVT*	Stage V sporulation protein T	**2.26**
CD3349	*bclA3*	Exosporium glycoprotein BclA3	**0.19**
stress			
CD1631	*sodA*	Spore coat protein, superoxide dismutase	**2.87**
CD2865		Putative bacterioferritin	**2.20**
envelopes			
CD0147		Putative transporter	**20.56**
CD0146		Putative membrane protein	**2.24**
CD1928		Putative membrane proteinTspO	**3.05**
CD2445		Transmembrane signaling protein, TspOT	**2.37**
CD2737		Putative nitrilase/cyanide hydratase	**2.25**
CD2738		Putative cytosine permease	**2.14**
CD2833		Putative calcium-transporting ATPase	**2.25**
CD2800		Putative membrane protein	**2.60**
CD3073		Putative membrane protein	**2.19**
CD2664	*murE*	UDP-N-acetylmuramyl-tripeptide synthetase	**1.93**
CD0090	*prlA*	Preprotein translocase SecY subunit	**0.44**
metabolism			
CD0740		Putative PLP-dependent aminotransferase	**2.17**
CD0777		Putative membrane protein	**1.96**
CD0778		Conserved hypothetical protein	**2.15**
CD0779		Putative amidohydrolase	**2.13**
CD0780		Conserved hypothetical protein, DUF1177	**2.46**
CD0994		Putative serine-pyruvate aminotransferase	**1.91**
CD0995	*serA*	Putative D-3-phosphoglycerate dehydrogenase	**2.95**
CD0996		Conserved hypothetical protein	**2.22**
CD1319		Putative polysaccharide deacetylase	**3.24**
CD1697	*ribH*	6,7-dimethyl-8-ribityllumazine synthase	**15.04**
CD1698	*ribBA*	Riboflavin biosynthesis protein ribBA	**5.68**
CD1699	*ribE*	Riboflavin synthase alpha subunit	**3.84**
CD1700	*ribD*	Riboflavin biosynthesis protein ribD	**3.88**
CD1749		Putative 2-hydroxyacyl-CoA dehydratase	**1.94**
CD1750		Putative CoA enzyme activase	**2.20**
CD1767	*gapB*	Glyceraldehyde-3-phosphate dehydrogenase	**2.70**
CD2387		a-hydroxy acid dehydrogenase,	**2.02**
CD2819		Putative aminho acid racemase	**2.25**
CD2864		Putative hydrolase	**4.00**

Gene ID	Name	Function	Expression ratio[a]
CD3236		Putative membrane protein	**1.89**
CD3238		Putative component of proline reductase	**1.96**
CD3241	*pdrB*	Proline reductase (selenocysteine)	**2.07**
CD3243		Conserved hypothetical protein	**2.19**
CD3244	*pdrA*	D-proline reductase proprotein pdrA	**2.09**
CD3248		Polysaccharide deacetylase	**2.27**
CD3257		Putative polysaccharide deacetylase	**2.13**
CD3258		Iron hydrogenase	**2.26**
CD3489		Putative oligoendopeptidase F	**2.29**
CD2323		Putative sugar-phosphate dehydrogenase	**0.53**
CD2324		Putative sugar-phosphate dehydrogenase	**0.44**
CD2325	*gatC*	PTS system, specific IIB component lactose	**0.47**
CD2326		PTS system, specific IIB component lactose	**0.47**
CD2327	*gatA*	PTS system, galactitol-specific IIB component	**0.48**
CD2600	*cstA*	Carbon starvation protein. CstA	**0.42**
regulators			
CD2214	*sinR*	Transcriptional regulator, HTH-type	**8.23**
CD2215		Transcriptional regulator, HTH-type	**8.22**
miscellaneous			
CD0356	*xis*	Excisionase Tn916-like, CTn1-Orf2	**2.45**
CD3235	*ssb*	Single-stranded DNA-binding protein	**3.63**
CD2517.1		Putative phage protein	**0.25**
Proteins of unknown function			
CD0311		Conserved hypothetical protein	**4.12**
CD1063		"	**2.95**
CD1463		"	**2.41**
CD1726		"	**3.87**
CD1941		"	**2.38**
CD2046		"	**2.25**
CD2098		"	**2.09**
CD2112		"	**2.76**
CD2245.1		"	**3.35**
CD2366		"	**2.76**
CD2369		"	**2.70**
CD2375		"	**5.09**
CD2409		"	**2.28**
CD2752		"	**2.30**
CD2808		"	**3.04**
CD2809		"	**2.52**
CD3457		"	**2.62**
CD3610		"	**2.13**
CD3618		"	**2.31**
CD1067		"	**0.03**
CD2216		"	**0.39**

[a] A gene is considered to be differentially expressed when P value is <0.05. Genes significantly downregulated are in grey, while genes significantly upregulated are in bold.

ITQB-UNL | Av. da República, 2780-157 Oeiras, Portugal
Tel (+351) 214 469 100 | Fax (+351) 214 411 277

www.itqb.unl.pt