Lourenço Maria Corrêa Monteiro Cayolla Bonneville
BSc in Biology

CHARACTERIZATION OF BIOFILM FORMATION BY *CLOSTRIDIOIDES DIFFICILE*

MASTER IN MEDICAL MICROBIOLOGY
NOVA University Lisbon
November, 2021

Lourenço Maria Corrêa Monteiro Cayolla Bonneville
BSc in Biology

CHARACTERIZATION OF BIOFILM FORMATION BY *CLOSTRIDIOIDES DIFFICILE*

MASTER IN MEDICAL MICROBIOLOGY
NOVA University Lisbon
November, 2021

CHARACTERIZATION OF BIOFILM FORMATION BY *CLOSTRIDIOIDES DIFFICILE*

LOURENÇO MARIA CORRÊA MONTEIRO CAYOLLA BONNEVILLE

Bachelor of Science in Biology

Adviser: Dr. Mónica Paula Fernandes Serrano Miranda,
Researcher, ITQB-NOVA

Co-adviser: Dr. Adriano José Alves Oliveira Henriques,
Associate Professor, ITQB-NOVA

Examination Committee:

Chair: Dr. José Paulo Sampaio,
Associate Professor, DCV FCT-NOVA

Rapporteur: Dr. Leonilde Moreira,
Associate Professor, IST-UL

Adviser: Dr. Mónica Paula Fernandes Serrano Miranda,
Researcher, ITQB-NOVA

MASTER IN MEDICAL MICROBIOLOGY

NOVA University Lisbon
November, 2021

Characterization of biofilm formation by *Clostridioides difficile*

Acknowledgments

Firstly, I would like to thank the Scientific Committee of the MSc in Medical Microbiology of UNL and FCT-NOVA for the planning and organization of this master. To the Instituto de Tecnologia Química e Biológica António Xavier of Universidade Nova de Lisboa for receiving me as a master student throughout this year who imposed challenges for new opportunities

A special acknowledgement to Professor of the year Jaime Mota, for all the organisation and management of the Ninth Edition of MSc Medical Microbiology, especially in a pandemic year.

I would like to thank my supervisor the “Supé Chefe” Dr. Mónica Serrano and to my co-supervisor the “el gran chefe” Prof. Dr. Adriano O. Henriques for receiving a student in pandemic times, but as well for guiding me, teaching me, and always give a positive word of encouragement when needed. For all the chocolates and “mumus” as well.

I would also like to thank Dr. Tiago Cordeiro for the structural model of Veg. To Bruno Salgueiro for assisting with Thermofluor assays.

A big and special thank you to my two “fairy godmothers”, Cristina Timóteo and Teresa Silva, for all the love, laughs, and for providing me with all the training to be able to conduct this research.

To all my lab colleagues, Bruno Gonçalves and Sara Ramalhete and Aristides Mendes for the strains and plasmids, as well as all the wisdom and unwind moments. To Mónica Louro, for accompanying me in the first work weeks. A special thank you for Diogo Martins, for all the advice, for all the lab “parenting” and for dishing when needed. To Zoé Vaz da Silva, for all the nagging, advice, small talk and for the recipe of granola, which I will cherish forever. For the nicest words of encouragement and a great collecting system, to Khira Amara. To Inês Morais, Mariana Valente and Carmen Olivença, for all the empathy, sympathy, and resilience when worked seemed overwhelming. To the new master students, Isabel Roseiro and Constança, that your future may bring everything you propose to.

To all my friends, that endlessly heard me talk about bacteria and microbiology and scientific gibberish. With a special shoutout to my colleagues, Raquel, Laura and Filipe for all the study and “sueca” nights.

Moreover, to my family, my parents Sofia “Pi Vinagre” Bonneville and to Miguel Bonneville, to my brother, Manuel Bonneville, and to my dogs, Conde and Coca, that have given me the possibility, opportunity, and conditions to proceed with my education, as well as guidance and support in critical moments.

This work was supported by the FCT (“Fundação para a Ciência e a Tecnologia") through the award PTDC/BIA-MIC/29293/2017.

Finally, to all that made this work possible, even in a pandemic scenario, that worked so that all our work could continue.

Abstract

Clostridioides difficile is an urgent threat level enteric pathogen, that can produce endospores, toxins and biofilms. This work focused on understanding the regulatory circuits controlling gene expression during biofilm formation in *C. difficile*. We characterized the Veg protein, a highly conserved protein among Gram-positive bacteria, and that in *B. subtilis* was shown to stimulate biofilm formation by inducing biofilm-specific gene expression. Veg, is a 10 kDa, Sm-like protein. Sm proteins form multimers that bind to RNA. We show that the overproduction of Veg also stimulated biofilm formation in a *C. difficile* lab strain. *In vitro,* the oligomeric state of Veg is pH-dependent and shifts the protein from monomer to dimer, and higher order protein complexes. We have no evidence for binding to RNA but we show that Veg binds to DNA, specifically to a TA rich region in the *cdeM* promotor. *cdeM* is a late sporulation gene involved in assembly of the spore surface layers and whose expression is dependent on *yabG,* the gene adjacent to *veg,* coding for a cysteine protease. We propose a model of repression of *cdeM* by Veg, that could be lifted by proteolysis of Veg by YabG.

Keywords

Clostridioides difficile; biofilm; Veg; Sm-like protein; protein-DNA interaction.

Resumo

Clostridioides difficile é um patogéneo entérico humano, de grande importância médica e económica, com capacidade de produzir endósporos, toxinas e biofilmes. O presente trabalho focou-se na compreensão dos circuitos regulatórios que controlam a expressão genética durante a formação de biofilme em *C. difficile*. Em particular, procedemos à caracterização de Veg, uma proteína conservada em bactérias Gram-positivas, que, em *B. subtilis*, induz a expressão de genes específicos para a formação de biofilme e estimula a produção destas estruturas. Veg é uma proteína de 10 kDa com um domínio Sm, característico de proteínas que formam oligómeros e ligam a RNA. Neste trabalho demonstramos que a sobreprodução de Veg estimula a produção de biofilme numa estirpe de laboratório de *C. difficile*. *In vitro*, Veg existe maioritariamente como um monómero a pH igual ou inferir ao pI, e como um dímero e formas de elevado peso molecular acima do pI. Apesar de não termos evidência de uma interação com RNA, demonstramos que Veg liga a DNA, especificamente a regiões ricas em AT no promotor *cdeM*. O gene *cdeM* é transcrito em fases tardias da esporulação, na célula mãe, e é essencial para a montagem da superfície do esporo. A expressão de *cdeM* é dependente de *yabG.* Por sua vez, *yabG*, imediatamente a montante de *veg,* codifica para uma protéase de cisteína. Propomos um modelo de repressão de *cdeM* por Veg, que pode ser levantada através de proteólise de Veg por YabG.

Palavras-Chave

Clostridioides difficile; biofilme; Veg; Proteína Sm-*like*; interação proteína-DNA.

Table of Contents

List of figures

List of tables

Symbols and abbreviations

%	Percentage
°C	Celsius degrees
Abs	Absorbance
ACE	Allele coupled exchange
APS	Ammoniumperoxidisulphate
BHI	Brain-Heart Infusion Broth
BHIS	Brain Heart Infusion Supplemented L-cystein hydrochloride
BHISG-DOC	Brain Heart Infusion Supplemented L-cystein hydrochloride, glucose and DOC
bp	base pair
CA	Cholate
CDC	Centre for Disease Control
CDI	*C. difficile* Infection
CDMM	*C. difficile* Minimal Medium
CFU	Colony forming units
CV	Crystal Violet
DB	Dextran Blue
DNA	Desoxyrribonucleic acid
DNAase	DNA endonuclease I
DOC	Deoxycholate
DTT	Dithiothetriol
EDTA	Ethylenediamine tetraacetic acid
FoA	5-Fluoorotic Acid
FMT	Fecal Microbiota Transplantation
EMSA	Electrophoretic Mobility-Shift Assay
EtOH	Ethanol
g	grams
g	acceleration of gravity
GA	Glutaraldehyde
h	hours
kDa	kilodaltons
kg.cm-2	kilagrams per square cm
L	Litre
LB	Luria-Bertani Broth
M	Molar
mA	milliampere
mg	Milligrams
min	minutes
mm	millimeter
mM	millimolar
ms	millisecond
MW	Molecular Weight
MWCO	Molecular Weight Cut-Off
nm	nanometers
nM	nanomolar

o/d	over-day
o/n	over-night
OD	Optical Density
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate Buffered Saline
PBS-T	PBS with 0.1 % Tween-20
PCR	Polychain Reaction
pI	Isoelectric Point
RBS	Ribosome Binding Site
RNA	Ribonucleic Acid
RUG	Resorufin-beta-D-glucuronic acid methyl ester
RT	Ribotype
RNase	Ribonuclease
s	second
SDS	Sodium dodecyl sulfate
SEC	Size-Exclusion Chromatography
TCA	Taurocholate
TAE	Tris-Acetate-EDTA buffer
TEMED	Tetramethylethylenediamine
Tris	Tris-base
UV	Ultraviolet
µL	Microliter
µM	Micromolar
µg	Microgram
V	Volt
VC	Volumes of Column
v/v	Volume per volume
WT	Wildtype
λ	Wavelength
σ	Sigma

The abbreviation listed is according to the recommendation published by the Journal of Biological Chemistry (JBC). All the other abbreviations are defined in the text.

1. Introduction

1.1 *Clostridioides difficile,* an urgent threat level pathogen

1.1.1 *Clostridioides difficile*, a brief introduction

Clostridioides difficile (formerly known as *Clostridium difficile*) is an obligatory anaerobic endospore-forming gram-positive bacillus, that has the ability to produce toxins and biofilm (Deakin et al., 2012a; Theriot et al., 2014). This bacterium is an enteric pathogen that can colonize the human gut and causes disease with symptoms ranging from mild diarrhoea to abdominal pain, fever, and severe diarrhoea. Complicated *C. difficile* infection (CDI) is characterised by inflammatory lesions, pseudomembranous colitis, toxic megacolon, bowel perforation, sepsis, shock and even death (Rupnik, 2007; Rupnik et al., 2009).

C. difficile is presently one of the major causative agents of nosocomial diseases and is categorized by the Centre for Disease Control (CDC) as an urgent threat due to the numerous hospitalised cases and deaths every year as well as to the economic cost associated. The CDC estimated that in the year 2018 around 295 thousand new cases in hospitalised patients appeared and almost 13 thousand deaths, comprehending costs of a thousand million US dollars attributable to healthcare (2018 Annual Report for the Emerging Infections Program for *Clostridioides difficile* Infection). While CDI healthcare-associated cases are decreasing, the spread of community-associated cases are increasing.

The burden of healthcare-associated CDIs in acute care hospitals in the European Economic Area was estimated at 123,997 cases annually. In the United States, *C. difficile* is the most common cause of healthcare–associated infections, accounting for approximately 15% of them. According to data from 2012, *C. difficile* caused approximately half a million infections and 29,000 deaths in the US. The pooled incidence rate of CDI in Asia was calculated by meta-analysis at 5.3/10,000 patient days (95% CI 4.0–6.7) (de Roo et al., 2020; McDonald et al., 2018).

One of the most important risk factors is the antibiotic therapy during a long-term hospitalisation, this represents Hospital-Acquired CDI (HA-CDI), where the average age is 72 years old, but HA-CDI is not the only form of CDI known. Community Acquired CDI (CA-CDI) is now a common form of CDI. CA-CDI patients are usually younger than the HA-CDI patients, with the average age of 50. It is still not clear what increases the risk factors of CA-CDI, but antibiotic therapies combined with the increase of asymptomatic hosts, as well as the increased risk of food and water contamination, may be contributing (Gupta & Khanna, 2014).

CDI is usually associated with extensive antibiotic therapy as treatment for underlying conditions (figure 1.1). Antibiotics can disrupt the intestinal colon, or gut, microbiome, which creates a niche opportunity for *C. difficile* to develop and generate active CDI. Normally, the intestinal microbiota mediates colonization resistance against *C. difficile,* but, an antibiotic

treatment disrupts the host microbiota, resulting in *C. difficile* growth, colonization of the intestine and toxin production**.** Toxin activity causes epithelial damage and resultant disease symptoms, shedding endospores (hereinafter spores for simplicity), which promote transmission of the organism.

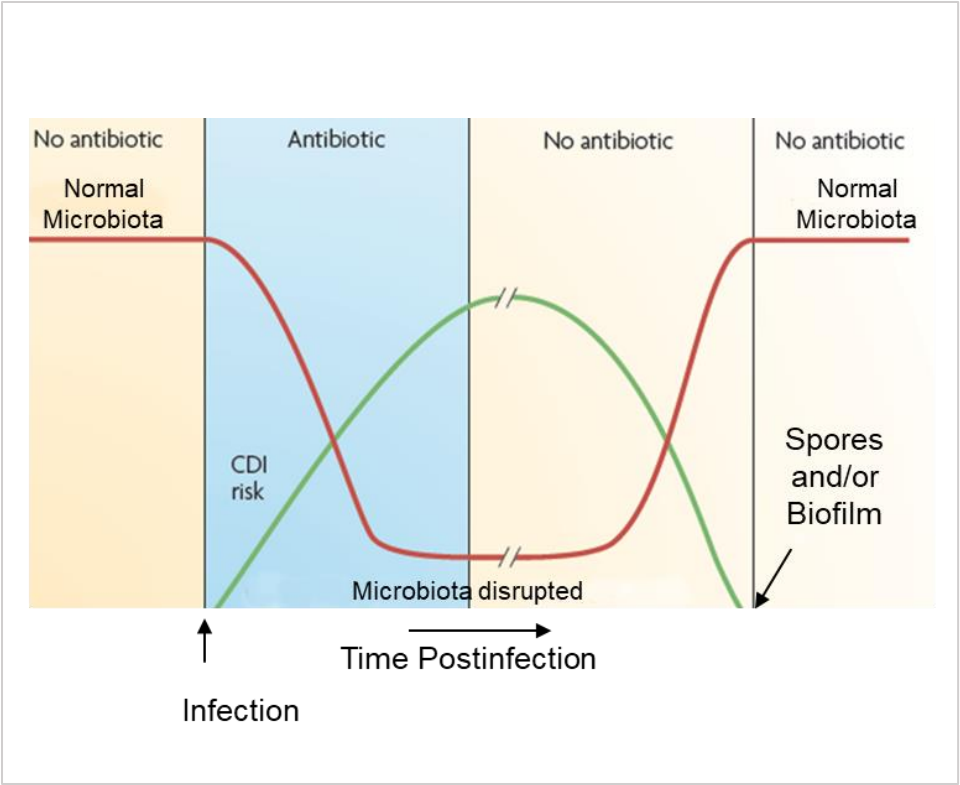

Figure 1.1 - The effect of antibiotics on the normal gut microbiota and the risk of *C. difficile* infection (CDI). A normal microbiota is present with no antibiotic exposure. Upon antibiotic therapy, there is a disruption of the normal microbiota, which leads to dysbiosis, which in turn increases the risk of CDI. After the cessation of the antibiotic therapy, the microbiota will eventually recover to a normal state. This recovery is slow as the microbiota remains disturbed for a variable period of time, represented by the break in the graph. After this period, the risk of CDI decreases. Recurrence occurs within two weeks of the first successful treatment. Recurrence is due to persistence of *C. difficile* in the gut or the host immediate environment, but it is not clear if the organism persists as spores, biofilms, or spores within biofilms. Figure adapted from Rupnik and co-authors (Gamier & Cole, 1988; Rupnik et al., 2009).

After a successful treatment, there is still the risk of recurrence, or a relapse of CDI within 2-8 weeks after the first episode. It is estimate that recurrence occurs in 15-35 % of CDI patients. Recurrence is thought to be due to spores that remain in the organism; spores may have an intracellular niche (Sarker & Paredes-Sabja, 2012) or may remain in the gut associated with biofilms that may have formed. This combination may be responsible for the persistence of *C. difficile* in the gut, where it can, if the conditions for CDI meet again, produce a recurrent infection (figure 1.1, (Rupnik et al., 2009)).

1.1.2 Gut colonization, role of the gut microbiome and infection cycle

The human intestine is an example of an extremely complex ecosystem that can harbour multiple species, or microbiota, which plays an important role in protecting the host from enteric pathogens, in modulating the immune response as well as in nutrition and pathological processes (Blaut et al., 2002; Cummings, 2009). This ecosystem contains a large number of species that include beneficial bacteria along with commensal bacteria and pathobionts. Multiple enteric pathogens, such as *Salmonella enterica*, *Yersinia*, *Escherichia coli* and *C. difficile* can modify the dynamics of this complex ecosystem to establish infection. These gut pathogens have highly adhesive and invasive properties. Some of the bacterial virulence factors include adhesins and pili, which can promote gut colonization and infection (Calabi et al., 2002; Tasteyre et al., 2001; Wright et al., 2008). Infection starts with colonization of the host which usually occurs through ingestion of the bacteria or spores which may then settle in the infection niche and start infection in favourable conditions.

For CDI to occur, first spores have to be ingested. These spores must then endure several challenges until they reach the large intestine, the preferred niche for colonization. When favourable conditions are met, the spores will germinate, resulting in growth and infection (K. H. Wilson & Perini, 1988). *C. difficile* may then form microcolonies, which are biofilm-like structures, on the epithelium, which facilitate persistence in the gut. CDI is then considered active if there is production of the *C. difficile* toxins which cause damage to the epithelial tissues, potentially leading to medical conditions (see 1.1.1).

One of the factors that is thought to have an influence in the germination of spores is the gut microbiota. The microbial diversity of an adult gut microbiota usually includes bacteria of the *Bacteroidetes* and *Firmicutes* phyla (Eckburg et al., 2005). The ratio between the abundance of these two phyla has been widely discussed and it has been described to be involved in nutrients absorption and food fermentation, stimulation of the host immune system, and barrier effects against pathogens(Bäckhed et al., 2012; Macpherson, 2006; Mowat, 2003). Within this complex ecosystem, *Clostridium scidens*, produces an enzyme, 7α-dehydroxylase, that converts primary bile salts, such as cholate (CA) and chenodeoxycholate (CDCA), into secondary bile salts, like deoxycholate (DOC) or lithocholate (LCA). The presence of DOC and other secondary bile salts represent a stress to vegetative cells of *C. difficile*, since these affect the integrity of bacterial cell membrane, damage the DNA and cause protein denaturation(Begley et al., 2005). The primary bile salts, either unconjugated or conjugated, CA and Taurocholate (TCA), respectively, induce *C. difficile* spore germination and promote growth, unlike the secondary bile salt DOC which induce germination of the spores but inhibits vegetative growth (Sorg & Sonenshein, 2008). In a state of dysbiosis, the ratio of bile salts may shift to an accumulation of the primary bile salts , CA and TCA, and favour germination of the spores and vegetative growth, and, subsequently, create an active CDI (Theriot et al., 2014; Y. Zhao et al., 2013).

Dysbiosis of the gut microbiome is a phenomenon associated with antibiotic therapy for extended periods of time with broad-spectrum drugs. Gut bacteria, that are sensible to the antibiotic therapy, are compromised leading to the creation of opportunities for pathogenic colonization and infection since *C. difficile* is resistant to multiple antibiotics mainly due to the presence of mobile elements in its genome. Besides dysbiosis, the massive use of antibiotics and long antibiotic therapies can also lead to the emergence of antibiotic resistant strains. A striking example was the intense use of fluoroquinolones in Northern American hospitals during the years 2000's, which produced a selective pressure for strains of *C. difficile*, to become resistant to these antibiotics and their subsequence dissemination worldwide (Brown et al., 2013; Isidro et al., 2017; Linder et al., 2005; Sebaihia et al., 2006; Slimings & Riley, 2014) (figure 1.2).

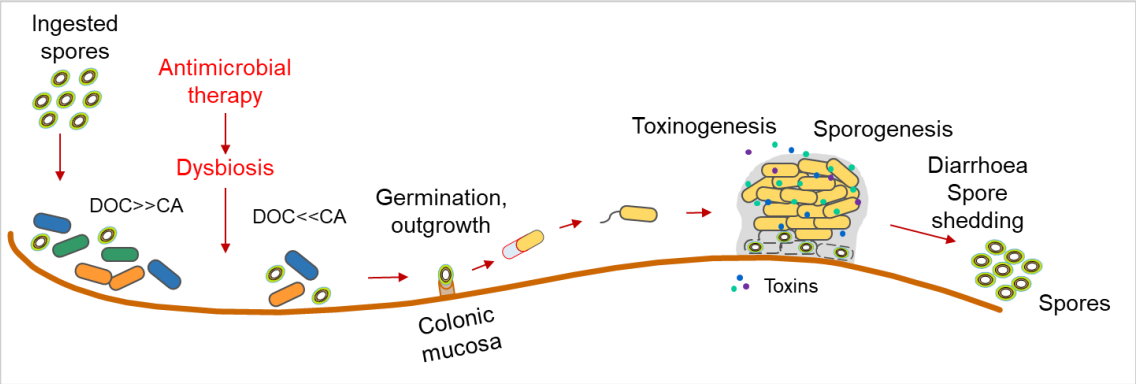

Figure 1.2 – Gut colonization, role of the gut microbiome and infection cycle of *C. difficile*. Infection starts with colonization, which in a scenario of dysbiosis, occurs following the ingestion of the bacterial spores. Spores may germinate in the gut since there is presence of primary bile salts (such as cholate, CA), but the presence of secondary bile salts (such as deoxycholate, DOC) inhibits vegetative growth. Upon antibiotic therapy, usually with broad-spectrum drugs, dysbiosis of the gut may occur, compromising the group of bacteria that convert primary bile salts into secondary bile salts, causing an accumulation of the precursors (CA) in the gut. CA ratio increases in the colon, allowing spores to complete germination. Spore germination may further depend on the recognition of a specific receptor in the colonic mucosa. Following spore germination, the vegetative cells will grow. Most vegetative cells will then form peritrichous flagella, which allow motility and are an important factor in promoting adherence to the gut mucosa. Once the cells adhere to the mucosa, micro-colonies or biofilms can be formed, which contribute to the virulence of infection. In the biofilm cells are incased in a self-produced matrix composed of protein, exopolysaccharides (EPS) and eDNA and they produce toxins as well as spores. In the last phase of biofilm formation, dispersion, there is release of spores from the micro-colony. Spores are then released into the environment and can colonize a next host, completing the cycle. Figure adapted from Isidro and colleagues (Isidro et al., 2017).

Presently, treatment of CDI usually consist in antibiotic therapy, bacteriotherapy and immunological therapies. Antibiotics remain the therapy by default, with vancomycin and metronidazole being the primary option. These two antibiotics are also of broad-spectrum, which means they will also affect the already disturbed microbiota. A novel drug, fidaxomicin, is being adopted in new therapies, as it has a narrower-spectrum and also inhibits spore formation (Babakhani et al., 2012). Bacteriotherapies include the usage of pre and probiotics, ingestion of spores of non-pathogenic strains and faecal microbiota transplant (FMT). Of the three, only FMT

is recognized to be a prospective alternative to antibiotic therapy, as it has 90% of success rates in the recovery of the microbiota from a dysbiotic state (Kociolek & Gerding, 2016). There are some questions regarding FMT, not only related to the acceptance by the patient, but as well as what is considered a "healthy" microbiota and which characteristics to look for in the microbiota of an FMT donor. The third therapy relies in immunologically targeting the *C. difficile* toxins, TcdA and TcdB, preventing binding of the toxins to the epithelial cell receptor (Isidro et al., 2017).

Another challenge in case of *C. difficile* that must not be overlooked are the oxygen-resistant spores. The spores represent the transmission vehicle between hosts and to the environment, where due to their resistance proprieties (see 1.1.3.2), they can survive for long periods of time. Since spores are also resistant to antibiotic therapy, these can persist in the gut after CDI treatment. The persistence of the spores may lead to recurrence of CDI in case of new dysbiosis due to another antibiotic therapy. Another form of microbial development that may facilitate the persistence of spores in the gut is biofilms. Biofilms increase the contact of *C. difficile* cells and the epithelium, while conferring additional protection against aggressors like antibiotics. In these structures, spores are formed, and may persist in the gut (Semenyuk et al., 2015). It is not clear if it is solely spores or biofilms, or a combination of both, that are responsible for the persistence and recurrence of CDI. According to the Centre for Disease Control (CDC), recurrence is defined as a relapse of CDI within 2-8 weeks after successful treatment of the first episode (McDonald et al., 2018). Recurrence is expected to occur in 15-35% of the patients (Marsh et al., 2012). If a case of CDI occurs 8 weeks after the first successful treatment, the CDC defines the case as reinfection (McDonald et al., 2018).

1.1.3 Virulence factors

C. difficile virulence factors allow it to infect and cause disease, and include among others, toxin production, sporulation, cell-surface associated proteins, flagella and biofilm formation (Awad et al., 2014). Here we will focus on toxin production, spores, and biofilm formation.

1.1.3.1 Toxin production

Some *C. difficile* strains have the ability to produce toxins. These toxins are what confer the ability to cause disease symptoms and are considered to be the main virulence factor (Burns et al., 2010; Carter et al., 2012; Deakin et al., 2012a; Rupnik et al., 2009; Sarker & Paredes-Sabja, 2012). The spectrum of diseases caused by *C. difficile* is thought to depend on the level of toxin production, supporting the hypothesis that regulation of toxin production is a key factor of *C. difficile* pathogenicity (Åkerlund et al., 2008). Toxin production is associated with a stationary phase of growth (Hundsberger et al., 1997). Many other environmental factors may influence toxin regulation and production, such as simple sugars, like glucose, or certain aminoacids, like

cysteine or proline, which can inhibit toxin production (Dupuy & Sonenshein, 1998; Karlsson et al., 2000).

There are two main toxins, TcdA, and TcdB, that are thought to be the main cause of extensive epithelial damage and pathology(Carter et al., 2010). Some *C. difficile* strains produce a third toxin, CDT, a binary toxin, whose role is not quite understood although recent data suggest an impact of CDT in host gut inflammation and immune response during infection (Cowardin et al., 2016). The TcdA and TcdB toxins are glucosyltransferases that have a cytotoxic effect on the epithelial cells of the host where they target the Rho family GTPases, Rho, Rac and Cdc42 (Oezguen et al., 2012). The TcdA and TcdB toxins are encoded by the *tcdA* and *tcdB* genes located in a Pathogenicity Locus, or PaLoc, which carries three other genes, *tcdR*, *tcdE* and *tcdC*, coding for proteins that have a role in toxin regulation (Hammond & Johnson, 1995). TcdR is an RNA polymerase sigma factor that binds to the promoter regions upstream of the *tcdA* and *tcdB* genes (Mani & Dupuy, 2001a). TcdE is a holin-like protein, and it is thought to be involved in the export of the toxins (Govind & Dupuy, 2012; Olling et al., 2012). TcdC was considered to be an anti-sigma repressor that impairs the interaction between TcdR and the RNA polymerase hence negatively regulating toxin production (Matamouros et al., 2007). However, the function and relevance of TcdC remains unclear, since several studies failed to show an association between a *tcdC* genotype and toxin production (Hunt & Ballard, 2013; Janoir, 2016).

Other regulators located outside the PaLoc are involved in toxin production, (Dineen et al., 2007) such as, Spo0A and σ^H, regulators that control post-exponential events, (Saujet et al., 2011; Underwood et al., 2009), σ^D, the sigma factor that controls flagellar synthesis and *tcdR* expression (Fimlaid & Shen, 2015; Mani & Dupuy, 2001b), RstA, a transcriptional regulator (Daou et al., 2019; Dineen et al., 2007) and CodY and CcpA, that respond to nutritional cues (Antunes et al., 2011; Dineen et al., 2007). Direct binding of RstA, CodY and CcpA to *tcdR*, *tcdA* and *tcdB* genes was previously shown (Antunes et al., 2012a; Fimlaid & Shen, 2015).

1.1.3.2 Sporulation

C. difficile spores are considered to be a virulence factor (Awad et al., 2014). As a strict-anaerobe, the bacteria rely on spores, which are oxygen resistant, to disseminate in aerobic environments (Paredes-Sabja et al., 2014a; Sarker & Paredes-Sabja, 2012). This dormant cell type confers additional protection against aggressors not only to oxygen, but as well to antibiotics, lysozyme, stomach acids, etc., and it is what allows the persistence of the bacteria in the environment and in the host.

The developmental process of sporulation is largely controlled by a cascade of cell-type specific RNA polymerase sigma factors. During this process, there is cell to cell signalling pathways that ensures that the mechanisms of gene expression in the mother cell and in the forespore are triggered in a coordinated manner (Zhu et al., 2018). At the end of the sporulation process, the mother cell lysis and the spore is released into the environment (figure 1.3).

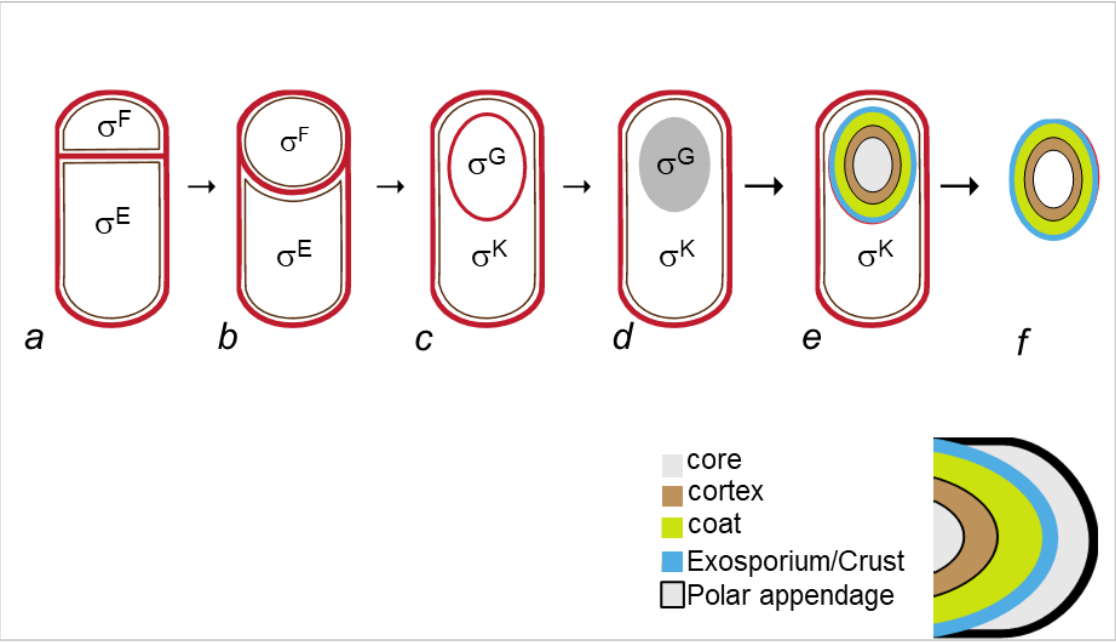

Figure 1.3 – Morphological stages during sporulation and gene expression in *C. difficile.* (Upper panel) The process begins with the asymmetric division (a), followed with the mother cells membrane surrounding the forespore, engulfing it (b). At the end of engulfment, a process coordinated by the sigma factors σ^E and σ^F, the forespore is a free protoplast inside of the mother cell cytoplasm (c). The sigma factors, σ^G and σ^K become active after engulfment, initiating the formation of the coat and assembly of the exosporium (d). The synthesis of the coat and late phase proteins (e), is coordinated by the late phase mother cell-specific σ^K. Lysis of the mother cell allows the release of the mature spore to the environment. A mature spore remains in this dormant state until conditions that trigger germination are met. The cell where the σ^F, σ^E, σ^G and σ^K sigma factors are active is indicated; **(Lower panel)** Legend of the spore layers. (Pereira et al., 2013).

The developmental process of sporulation starts with the activation, by phosphorylation, of a response regulator, Spo0A, which is conserved among sporeformers (Abecasis et al., 2013; Deakin et al., 2012a; Ramos-Silva et al., 2019). Spo0A is an ambivalent transcription factor as it acts both as a repressor and an activator, depending on its intracellular concentration and position, relative to the promoter, of its binding sites in the genes it controls (Chastanet et al., 2010; Ribis et al., 2018). Spo0A is transcribed in the vegetative cell under the control of two sigma factors, σ^A and σ^H (Fimlaid et al., 2015; Fimlaid & Shen, 2015). In *C. difficile,* Spo0A is activated directly by a group of sensor kinases that respond to external signals by auto-phosphorylating and subsequently transfer the phosphoryl group to Spo0A. Spo0A has been shown to control approximately 300 genes in *C. difficile*. Many of these genes were associated to biofilm formation, motility, toxin production and the genes that code for the first two sporulation-specific sigma factors, σ^F and σ^E. Spo0A further controls the expression of the genes required for asymmetric division of the cell entering sporulation. Following asymmetric division, the sigma factors become active in a cell type-specific manner: σ^E becomes active in the mother cell and σ^F in the forespore, which creates a difference in gene expression caused by the spatial division. After activation of σ^F in the forespore, it activates σ^E in the mother cell, which then activates the process of engulfment of the forespore, a phagocytic-like process where the septal membranes migrate

towards the proximal cell pole eventually fusing and releasing the forespore as a free protoplast inside the mother cell cytoplasm. At later times in development, following engulfment completion, σ^F and σ^E are replaced by σ^G and σ^K, respectively, which control the final stages of spore formation (Abecasis et al., 2013; Fimlaid et al., 2015; Fimlaid & Shen, 2015; Paredes-Sabja et al., 2014a; Ribis et al., 2018; Saujet et al., 2011; Wörner et al., 2006). Upon engulfment, the spore is involved in two layers of peptidoglycan, an inner thinner layer called the primordial cell wall which has a composition similar to the vegetative cell wall, and an outer second thicker layer with chemically distinctive peptidoglycan called the spore cortex. The cortex is essential in spore morphogenesis since it is responsible for the maintenance of the dehydrated state of the spore core, hence dormancy and heat resistance (Henriques & Moran, 2007; Sarker & Paredes-Sabja, 2012). The spore cortex is surrounded by the coat, which is composed of more than 80 different proteins orderly assembled in a multi-layered structure around the spore. Its assembly is coordinated with the end of cortex assembly (Driks & Eichenberger, 2016a). The coat confers protection of the cortex against the peptidoglycan hydrolytic activity of enzymes such as lysozyme, and it is also required for resistance to UV radiation and chemicals, such as peroxides (Driks & Eichenberger, 2016b; Henriques & Moran, 2007; Rodriguez-Palacios & LeJeune, 2011). *C. difficile* spores have an external, more electrodense, layer, termed exosporium, closely apposed to the underlying coat (Abecasis et al., 2013; Phetcharaburanin et al., 2014; Pizarro-Guajardo et al., 2016; Rabi et al., 2017). The exosporium is permeable to germinants but excludes lytic enzymes and antibodies, and it is the first line of contact of the spore with host cells, the immune system, and the environment (Rodriguez-Palacios & LeJeune, 2011).

1.2 Biofilm formation

1.2.1 Biofilm are ubiquitous form of bacterial development

Many, if not all, bacterial species have the ability to form biofilms, which may in fact be their normal mode of growth in natural settings (Flemming & Wuertz, 2019). Bacterial biofilms are defined as sessile microbial communities formed by cells adhered to a surface or substrate as well as to each other (Branda et al., 2005). Biofilms are usually contained in a self-produced matrix composed by a mix of exopolysaccharides (EPS), proteins and extracellular DNA (eDNA) (Arnaouteli et al., 2021b). This matrix facilitates the biofilm attachment to abiotic or biotic surfaces and is one of the hallmarks of biofilm formation (Flemming & Wuertz, 2019). These surfaces or environments that promote biofilm formation often include niches that can be colonized by other species of bacteria (Domínguez-Manzano et al., 2012; Pandit et al., 2020). Biofilms would naturally occur as multi-species assemblies (Kolenbrander et al., 2010; Tan et al., 2017) where inter-species interactions can occur. In this state of microbial development, the bacteria exhibit different patterns of growth rate and gene expression on the course of biofilm formation (Donlan, 2002). The ability to form these structures is considered important for virulence (Stoodley et al., 2002) since they confer additional protection against environmental aggressors such as

antibiotics and the host immune system (Costerton et al., 1995). This form of bacterial development, although ubiquitous, is still poorly understood, making a species-to-species approach necessary to better comprehend the underlying mechanisms.

Biofilms have great impact in different sectors of society. Biofilms of pathogenic bacteria can have an impact on food processing environments, such as meat broilers, cheese factories, or Ready-to-Eat food manufacturers where these structures can form on stainless steel surfaces and become persistent, but also in environments such as healthcare settings, with biofilms associated with prothesis, such as catheters or ventilators, which can lead to further systemic infections (Alvarez-Ordóñez et al., 2019; Carpentier & Cerf, 1993; Gominet et al., 2017; Huq et al., 2008; Vishwakarma, 2020). These situations can lead to economic loss and epidemiologic outbreaks. Comprehending biofilm formation is of great importance as it can lead to targets for new therapies/treatments that allow the eradication of pathogenic species to be more effective(Bonneville et al., 2021; Vishwakarma, 2020).

1.2.2 Biofilm formation, regulatory pathways, and cellular heterogeneity

Biofilm formation is an intricate process regulated by a transcriptional network that is triggered by environmental cues. In *B. subtilis*, a major model for biofilm formation by sporeformers, Spo0A, ComA and DegU are three essential master regulators that control the different stages of biofilm formation (Mielich-Süss & Lopez, 2015a). When cells become sessile, DegU, ComA or Spo0A are activated by phosphorylation (~P). Cells which accumulate ComA~P differentiate into two different cell types, surfactin producers and competent cells (Chen et al., 2020). Surfactin is a paracrine signal that induces the differentiation of matrix producers. Motile cells have low levels of DegU~P, which are sufficient for transcription of flagellar genes. In the biofilm, cells with high levels DegU~P are specialized in the secretion of exoproteases that degrade large polymers into smaller peptides that will be used by the population as nutrients.

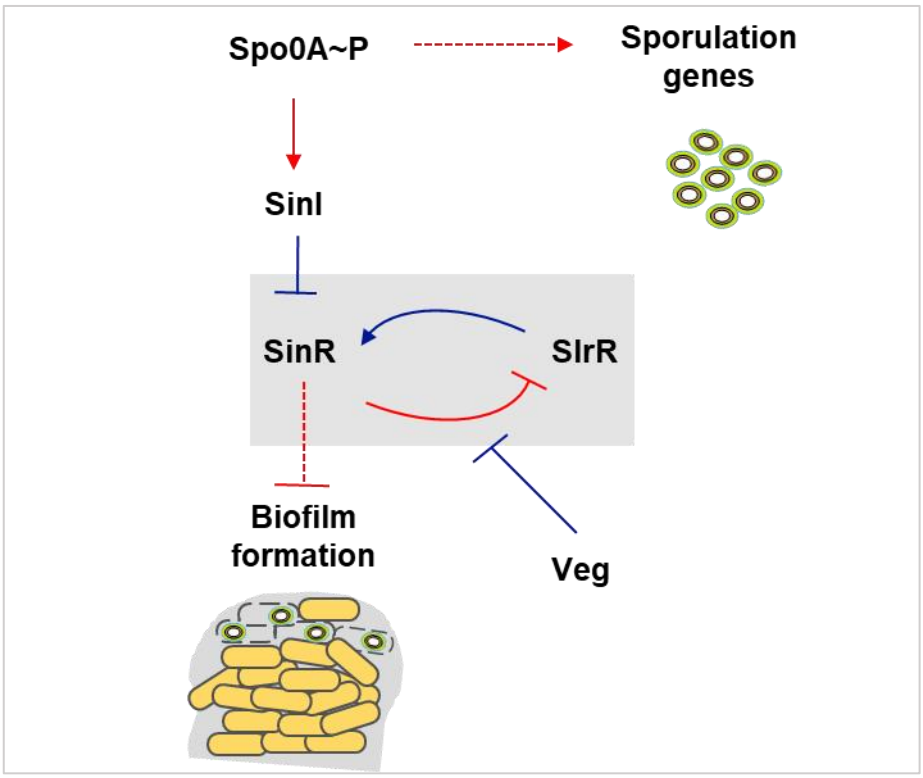

Figure 1.4 – Regulatory pathways involved in biofilm formation in *B. subtilis*. Arrows represent positive regulation or activation, and blunt arrows represent negative regulation or repression. In red, transcriptional regulation, in blue, protein-protein regulation. SinI is an anti-repressor which blocks SinR through the formation of a heterodimeric SinI-SinR complex. SinR activity is regulated by Slr, and the transcription of *slrR* is repressed by SinR, creating a double negative feedback loop. Veg acts as an additional regulatory repressor that contributes to the control of matrix production. Veg is thought to repress SinR activity independently of the SinI and SlrR pathways.

High levels of DegU~P also induces expression of *bslA* (Kobayashi & Iwano, 2012), coding for a matrix protein localized on the surface of the biofilm and required for its integrity (Arnaouteli et al., 2021a; Chen et al., 2020; Mielich-Süss & Lopez, 2015b); high levels of DegU~P in turn, repress transcription of the flagellar genes. Both Spo0A and Spo0A~P are present in the cell at the same time, and it is the relative level of these two forms that differentiates gene expression (Chastanet et al., 2010). While high levels of Spo0A~P induce sporulation-specific gene expression, intermediate levels of the regulator induce biofilm matrix production (Jiang et al., 2000). Under the control of Spo0A, and involved in biofilm formation, are a set of regulatory proteins (figure 1.4). These are the DNA-binding proteins SinI, SinR and AbrB (Shafikhani et al., 2002). AbrB and SinR repress the expression of the matrix encoding operons while SinI is an antirepressor that acts by blocking SinR through the formation of an heterodimeric complex SinI-SinR, thereby preventing SinR from binding to DNA (Bai et al., 1993; Chai et al., 2008; Kearns et al., 2005; Mandic-Mulec et al., 1992; Newman et al., 2013). The SinI-SinR heterodimer complex acts as a repressor for genes involved in motility (*hag* gene, flagellin) and involved in cellular separation (*lytABC* and *lytF*, autolysins) (Chai et al., 2008; Kearns et al., 2005). Another protein that may induce biofilm formation by inhibiting SinR, is Veg (figure 1.5). Veg is thought to repress SinR

activity independently of the SinR-SlrR pathways. Veg is an 86 amino acid, highly conserved, protein encoded by the *veg* gene, which is transcribed at very high levels during both exponential growth and sporulation (le Grice et al., 1986; Lei et al., 2013a; Ollington et al., 1981). Previous studies also reported delayed spore germination in *veg* mutants (Fukushima et al., 2003).

1.2.3 Biofilm formation in *Clostridioides difficile*

As the majority of the bacteria, *C. difficile* has the ability to form biofilms. After being firstly reported (Donelli et al., 2012), biofilms of *C. difficile* with clinical origin were produced on abiotic surfaces and quantified through crystal violet staining (Dapa et al., 2013; Dawson et al., 2012). It was quickly realized that *C. difficile* biofilms are multi-layered structures, encased in a matrix composed of bacterial proteins, extracellular DNA (eDNA), and exopolysaccharides; however, it is not clear if there is an attributable main matrix component, as the composition and morphology of the biofilm is strain dependent (Dapa et al., 2013; Dawson et al., 2012, 2021; Semenyuk et al., 2015). *C. difficile* biofilm communities are likely critical in recurrent CDI. Although there is some understanding of *C. difficile* biofilm formation and regulation *in vitro*, several questions regarding the relevance of such biofilm communities in bacterial persistence, and the bacterial and host factors regulating their formation *in vivo* remain unanswered (Frostid et al., 2021).

Biofilm formation usually undergoes four stages: adherence, proliferation, maturation, and dispersion. These four stages are characterized with different regulatory processes that allow the development of the biofilm. After germination in the gut, the vegetative cell is thought to be free until environmental triggers induce biofilm formation. Sessile bacteria display a different pattern of gene expression than their planktonic counteracts. It is not clear whether the interaction of the vegetative cell and the gut epithelium acts as a switch for biofilm formation; nonetheless, some stimuli that may act as signals to start biofilm formation have been identified (Dubois et al., 2019a; Girinathan et al., 2018a; Mani & Dupuy, 2001a; Meza-Torres et al., 2021; Tremblay et al., 2021). One of them has been mentioned previously, DOC, a secondary bile salt, that in sub-lethal concentrations has the ability to induce biofilm formation *in vitro* (Dubois et al., 2019a; Girinathan et al., 2018b). A second signal that appears to promote biofilm formation, is the presence of extracellular pyruvate. Extracellular pyruvate in the presence of DOC shows to increase the biomass of the biofilm but is dependent of pyruvate uptake by the membrane protein CstA (Tremblay et al., 2021). Cyclic-di-GMP (c-di-GMP) plays a role in the switch from motile to sessile biofilm state through the repression of the flagellar synthesis and induction of pilli and adhesins (Purcell et al., 2016). Several adhesins have been identified in *C. difficile*. The surface layer protein SlpA belongs to the S-layer class of proteins that are secreted and present in many bacterial species (Calabi et al., 2002). Additional to the role in cell structure, S-layer proteins have also been implicated in immune invasion, blocking the complemented-mediated lysis, and adhesion (Sára & Sleytr, 2000). Another family of proteins involved in adhesion is the Cell wall proteins (Cwp) family, such as the Cwp84 and Cwp66 (Wright et al., 2008). Other adhesins that

have been identified in *C. difficile* include the flagellin FliC and the flagellar cap protein FliD (Tasteyre et al., 2001), the heat-shock protein GroEL (Hennequin et al., 2001) and fibronectin-binding protein Fbp68 (Hennequin et al., 2003).

After adhesion, the now sessile cells may proliferate, attached to each other, and encased in the self-produced matrix, which, as described above, does not have an attributable key component in *C. difficile*. It is thought that, with biofilm maturation, the eDNA that is released with cell lysis is an important component of the matrix (Tremblay et al., 2021). Spores and toxins A and B are present in the mature biofilm (Semenyuk et al., 2015), suggesting that the biofilm may take a role in *C. difficile* virulence. *C. difficile* biofilm formation is characterized by a metabolic shift from glycolysis and the pentose phosphate pathway to the Stickland fermentation pathways and the Wood-Ljungdhal pathway, which are less efficient at producing energy (Neumann-Schaal et al., 2019).

When mature, the biofilm starts to disperse. In dispersion, it is thought to occur release of parts of the biofilm, as well as spores. The release of spores into the environment is what allows the colonization of new hosts. Even after successful treatment, recurrence may occur as explained in section 1.1.1. What is not understood if it is attributable to spores that persist in the gut, to the biofilm microcolonies that are encased in the protecting matrix, or both.

Several transcriptional regulation pathways are in place during *C. difficile* biofilm formation, involving the transcriptional regulators CcpA, CodY and Spo0A (Antunes et al., 2012b; Daou et al., 2019). CcpA and CodY are global regulators of the metabolism and support the idea that biofilm formation is dependent on a metabolic shift in planktonic cells ((Antunes et al., 2012b; Daou et al., 2019; Dubois et al., 2019b); see 1.2.3). Furthermore, the transition phase sigma factor σ^H and the master regulator of the sporulation Spo0A are also important for biofilm formation independent of the sporulation process ((Dawson et al., 2012; Deakin et al., 2012a); see 1.1.3.2). Additionally, σ^H and Spo0A are also associated with the metabolism of *C. difficile* (Dawson et al., 2012; Deakin et al., 2012b). *sinR* is also present in *C. difficile* genome and the members of its regulon were previously identified (Girinathan et al., 2018a). The *sin* locus is essential to establish a successful infection as it can regulate the transcription of key factors in the sporulation, toxin production and motility pathways. In contrast, some conserved proteins such as ComA, DegU, AbrB and Veg, coded for by the *comA* (*CD1269*), *degU* (*CD1269*), *abrB* (*CD1859* and *CD3120*) and *veg* genes are present in the *C. difficile* genome, but no functional characterization of these genes has been reported; yet they are strong candidates for genes with roles in biofilm formation. The Veg protein in particular, is thought to act as a repressor of SinR, based in the *B. subtilis* model (Lei et al., 2013a; figure 1.4). Although it shares a high degree of sequence similarity to the Veg protein of *B. subtilis* Veg, some differences are present that may be pivotal for a different regulatory role or *modus operandi*.

1.2.4 Veg protein in *Clostridioides difficile*

In *B. subtilis*, the overproduction of Veg stimulated biofilm formation via inducing transcription of the genes that encode the matrix (Lei et al., 2013a). Deletion of *veg* strongly impaired the biofilm formation and reduced the transcription of the same genes (Lei et al., 2013a). In *sinR* mutants the overproduction or deletion of *veg* made no significant difference in biofilm formation, whereas in *sinI* mutant, the overexpression of *veg* restored the impairment of biofilm formation, suggesting that Veg negatively regulates SinR activity independently of any known antirepressors (Lei et al., 2013a). *veg* is constitutively transcribed, at high levels, in both the exponential and stationary phase. Veg is highly unstable in the exponential phase, which makes it undetectable through Western Blot. Increasing Veg accumulation to a detectable level leads to induction of biofilm matrix genes. Since the protein is unstable, this suggests that rather than the transcriptional induction of the expression of *veg*, stabilization of Veg protein in the stationary phase, leads to the activation of genes that promote biofilm development. It is still unclear what signal stabilizes Veg protein to stimulate biofilm formation. (le Grice et al., 1986; Lei et al., 2013a; Ollington et al., 1981).

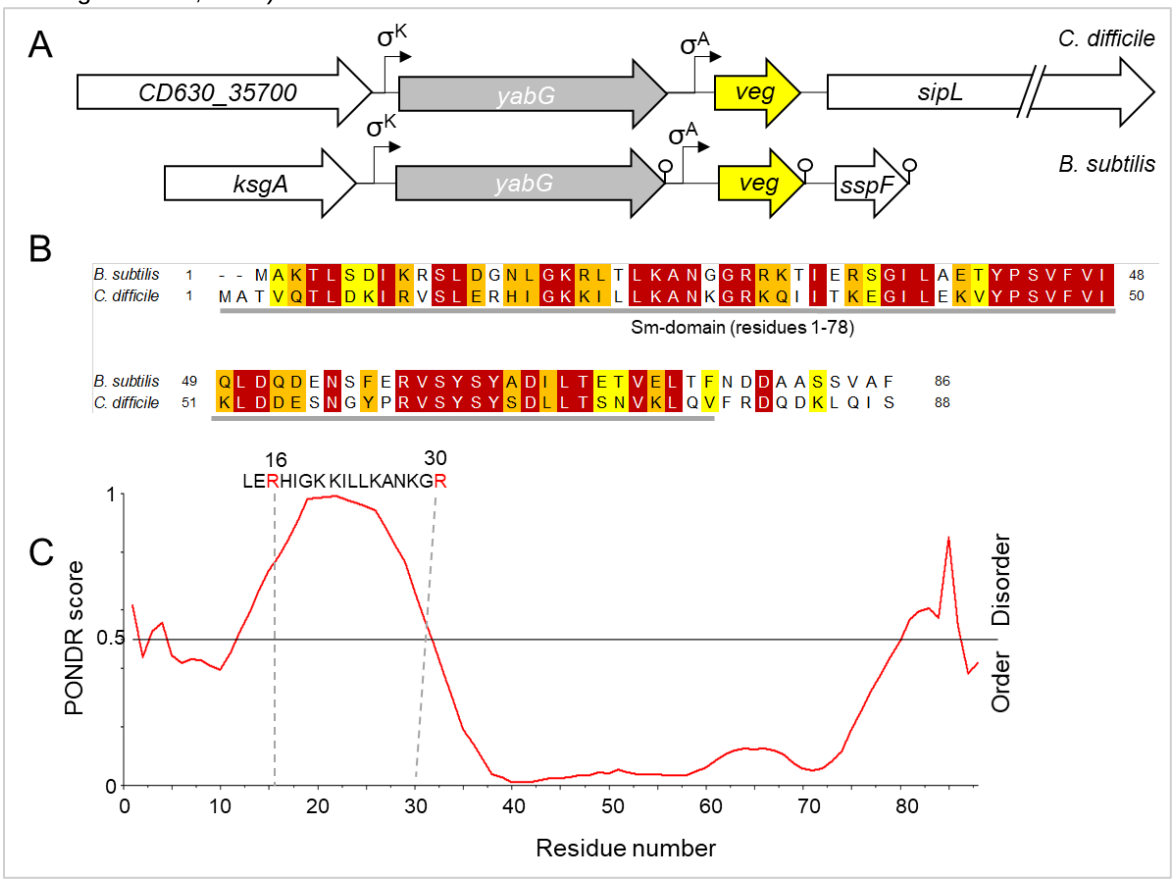

Figure 1.5 – Veg is conserved among Gram-positives. **(A)** Genomic map of the region surrounding the *veg* gene in the *C. difficile* genome (upper panel) and in *B. subtilis* (lower panel). The position of *veg* can be seen relative to its neighboring genes; the positioning of *veg* downstream *of yabG* which encodes for a cysteine protease is conserved across sporeformers. **(B)** Clustal W (https://www.genome.jp/tools-bin/clustalw) alignment of the Veg proteins from *C. difficile* and *B. subtilis.* The red background highlight identical residues. Residues are shadowed in orange or yellow according to the degree of conservation. **(C)** Plot represents the output of the PONDR server used to analyze the Veg sequence for order/disorder. Regions above a score of 0.5 are predicted to be disordered (Li et al., n.d.) Two arginine residues that flank a short-disordered region in Veg are highlighted in red. Adapted from (Ramalhete, 2021).

C. difficile Veg is a small basic protein, with 88 residues, with significant similarity to *B. subtilis* Veg (figure 1.5). To date no functional characterization of the *veg* gene has been reported in this gut pathogen. The *veg* gene is downstream of *yabG*, a gene that is part of a genomic signature for sporulation. *yabG* codes for a cysteine protease, that in *C. difficile* has a role in the assembly of the spore surface layers and in regulation of gene expression late during sporulation ((Kevorkian et al., 2016; Shrestha et al., 2019); figure 1.5). As mentioned above, *B. subtilis* Veg is unstable in some growth conditions. It was previously suggested that a protease, yet to be identified, would be involved in proteolysis of Veg (Ramalhete, 2021). The synteny of *veg* and *yabG* genes (figure 1.5) prompt us to hypothesize that the specific protease contributing to Veg instability may be YabG (Fukushima et al., 2003).

Veg contains an Sm-like domain. Sm-like domains are present in nucleic acid binding proteins involved in various roles regarding the translation, procession, either protecting or promoting degradation of mRNA (Bandyra et al., 2013; Mura et al., 2013). This family of proteins include regulator proteins like the Hfq protein, and usually behave in a ring-shape homo-oligomer of 6 or 7 units, that would bind the nucleic acid (Achsel et al., 1999, 2001a) (figure 1.6).

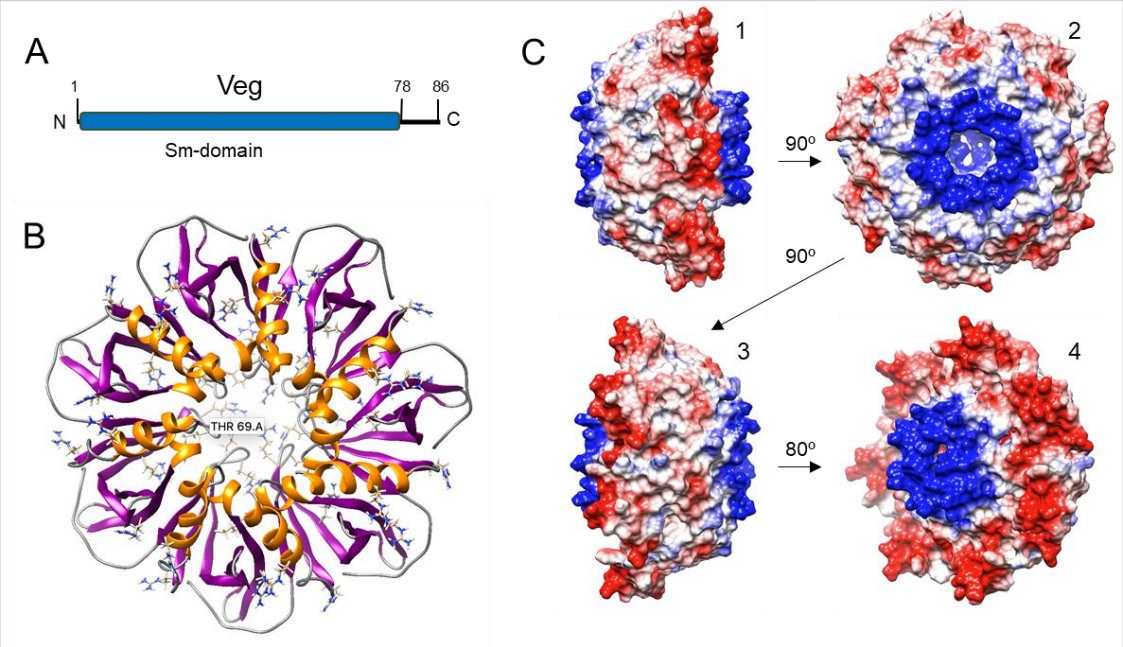

Figure 1.6 – The structure of *C. difficile* Veg. **(A)** Structural organization of Veg. Veg contains an Sm-domain, an ancient RNA-binding motif with oligo(U) specificity. **(B)** Homology model generated for Veg using as template a protein with unknown function from *S. pneumoniae* TIGR4 (PDB code: 3FB9). This structural model show that the Veg protein may form an heptamer with a central channel, an arrangement that is characteristic of Sm-like proteins. **(C)** Different views of the electrostatic surface map of the ring formed by the Veg heptamer. Red denotes negatives and blue positive charges.

When a Prediction of Natural Disordered Regions of Veg protein, or a PONDR model, (figure 1.5) is performed, Veg shows a major disordered region that is delimited by two Arginine residues, which are positively charged. This may suggest that this disordered region is capable to interact with negatively-charged nucleic-acid strands (Kambach et al., 1999; Urlaub et al.,

2001). Ordered regions suggest a form of secondary fold, such as an alpha-helix or a beta-sheet (figure 1.5).

Previously, a homology model of the Veg was produced to assess the overall conformation, fold, and position of the conserved amino acids. The structure of the protein with unknown function from *Streptococcus pneumoniae TIGR4* (protein data bank accession code: 3FB9) was used to build the model. The homology model of Veg shows the characteristic heptameric fold of the Sm-like domain (Achsel et al., 2001b; Bandyra et al., 2013; Mura et al., 2013; Urlaub et al., 2001) (figure 1.6). The model shows the presence of several surface-exposed positively charged arginine residues (figure 1.6). A surface electrostatic potential map representation of Veg shows that some parts of the surface of the oligomer are positively charged as is the inside of the ring (figure 1.6). These are features that could facilitate the interaction with the negative charged single stranded nucleic acids (Boudry et al., 2014, 2021; Sauer et al., n.d.). The oligomeric form of Veg still needs clarification outside of the *in silico* model. (figure 1.6).

1.3 Aim of this work

The characterization of regulatory proteins in model organisms is an essential element to understand cellular mechanisms. In other bacteria, such as *C. difficile*, essential processes may need to be re-examined since they may differ significantly from the model organisms. This poses a challenge by itself, since some techniques may be ineffective in other species, which represents a limitation to the insight in some of the processes. It is still unclear the role of the Veg protein in *C. difficile*, as few studies have tried to characterize this protein and its role of development. Understanding these processes is essential to better choose drug-target for therapy.

In order to investigate the biological significance of Veg, a structural characterization, to assess the stability, conformation, and oligomeric state of this protein, as well as a functional characterization of the protein in its role in physiological developmental processes, such as biofilm formation and sporulation were performed.

2. Materials and Methods

2.1 Microbiological techniques

2.1.1 Bacterial strains and growth conditions

The *C. difficile* and *E. coli* strains used in this study are listed in table 1 and 2, respectively. *C. difficile* strains were stored at - 80 °C in Brain Heart Infusion broth (BHI; table 5) (Biokar Diagnostics, Beauvais, France) containing 20 % (v/v) glycerol. *C. difficile* was routinely grown in BHI at 37 °C under anaerobic conditions (5 % H_2, 10 % CO_2, 85 % N_2). When necessary, cefoxitin (25 µg/mL) and/or thiamphenicol (15 µg/mL) was added to the cultures.

E. coli strain DH5α was used as a host for molecular cloning and for plasmid propagation. *E. coli* strain HB101 (RP4) was used for conjugation of plasmids into *C. difficile*. *E. coli* strain BL21 (DE3) was used for protein overproduction. Luria Bertani medium (LB; table 5) was routinely used for growth and maintenance of *E. coli* strains. When necessary, bacteriological agar was added to a final concentration of 1.6 % and when appropriate, chloramphenicol (20 µg/mL) and/or ampicillin (100 µg/mL) was added to the culture medium.

2.1.2 Biofilm formation assays

The method used to form biofilms was adapted from the work previously developed by Dubois et al., 2019b. Briefly, a 24-well tissue culture-treated microplate (Greiner bio-one, Austria) with 1 mL of BHI medium supplemented with L-cystein hydrochloride, D-glucose, and sodium deoxycholate (BHISG-DOC; table 5) per well, was inoculated with 10 µL of a BHI overnight culture. Two negative controls were added, with no inocula, to ensure sterility. The microplates were incubated at 37 ºC in anaerobic conditions for 24 h and 48 h.

2.1.3 Biofilm quantification

. After incubation for 24 or 48 hours, the media was removed by inversion of the microplate. The wells were washed twice with Phosphate-Buffered Saline solution (PBS; table 6) and air dried for ten minutes between each wash. The biofilms were then stained with Cristal Violet (CV; 0.2 % v/v in milliQH_2O) for a minimum of 30 minutes. The CV was removed by inversion and the wells were air dried for 15 min. The wells were washed twice with PBS to remove all the CV that was not bound to biofilm. Dye bound biofilm mass was solubilized by adding 1 mL of a 50:50 ethanol-acetone solution and incubated for a maximum of four minutes. The absorbance of the CV solution, corresponding to the biofilm biomass, was then measured at a 600 nm with a spectrophotometer (Ultraspec2100 pro). When needed (OD > 1.0), the solubilized CV was diluted 1:100 in the same 50:50 ethanol-acetone solution and measured again. Data obtained was checked for agreement to the normal distribution and for homogeneity of variance (Anderson-Darling test and Levene's test, respectively) by using the MiniTab17 software (Minitab, Inc., Pennsylvania, United States). When normality and homogeneity of variance were confirmed, one-way ANOVA with Tukey's test was performed to calculate statistical differences between average values. When the data did not comply with ANOVA assumptions, the non-parametric Kruskal–

Wallis median test was used. The software used was Statistica version 7.0 (Statsoft Inc., Tulsa, United States). For all tests, a probability greater than 95 % ($p < 0.05$) was considered as significant.

2.1.4 Growth rate of *Clostridioides difficile* strains

Growth rates are calculated with based on the slope of the curve correspondent to the exponential phase of growth from the strains. Growth curves were performed by measuring the OD at 600 nm (OD_{600}) hourly for 14 hours after inoculation. For inoculation, a pre-inocula was performed, so that all strains start at an equal OD_{600} of 0.05. Antibiotics were supplied to the media of growth when necessary. Generation time was determined according to the following equation: generation time = ln(2)/growth rate. Data obtained was checked for agreement and statistical difference use GraphPad Prism (GraphPad Software, San Diego, USA). A Dunn's multiple comparison test was performed to calculate statistical differences between each strain and it's control. For all tests, a probability greater than 95 % ($p < 0.05$) was considered as significant.

2.1.5 Sporulation Efficiency of *Clostridioides difficile* strains

Sporulation assays were performed as described previously (Putnam et al., 2013; Serrano et al., 2016). 150 µL of a BHI overnight culture was inoculated in a plate of 70:30 medium (table 5). After 24 h, serial dilutions were performed (up to 10^{-5}) and three spots (20 µL) of each of the dilution performed were spotted onto BHI plates supplemented with 0.1 % of sodium taurocholate (TA; Roth, Karlsruhe, Germany) to induce spore germination. After platting, the dilutions were removed from the anaerobic chamber and incubated 30 min at 70 ºC. This step is performed to assure that all CFUs platted after the heat shock are originated from spores and not viable vegetative cells. After the heat shock, they were reintroduced in the chamber where the platting of the three spots per dilution was repeated. After 24 h of incubation the colonies were counted, and the assessment of sporulation efficiency was calculated according to the formula:

$$\text{Sporulation efficiency } (\%) = \frac{\text{CFU / mL of heat} - \text{treated cells}}{\text{CFU/mL of non} - \text{treated cells} + \text{CFU/mL of heat} - \text{treated cells}} \times 100$$

Data obtained was checked for agreement to the normal distribution and for homogeneity of variance (Anderson-Darling test and Levene's test, respectively) by using the MiniTab17 software (Minitab, Inc., Pennsylvania, United States). When normality and homogeneity of variance were confirmed, one-way ANOVA with Tukey's test was performed to calculate statistical differences between average values. When the data did not comply with ANOVA assumptions, the non-parametric Kruskal–Wallis median test was used. The software used was Statistica version 7.0 (Statsoft Inc., Tulsa, United States). For all tests, a probability greater than 95 % ($p < 0.05$) was considered as significant.

2.2 Genetic and molecular biology techniques

2.2.1 Molecular cloning

DNA fragments for cloning were generated by the polymerase chain reaction (PCR) using the high fidelity Phusion DNA polymerase (Thermo scientific, Rockford, USA). All oligonucleotide primers used in this work are listed in the table 3. PCR products were purified and concentrated using the DNA Clean and Concentrator™ – 5 kit (Zymo research, Irvine, USA.). All other general cloning methodologies were as previously described (Sambroock and Green, 2012). All DNA restriction and modification enzymes were obtained from Thermo Fisher Scientific and used according to the manufacturer's guidelines. All the plasmids used and constructed during this work are listed in table 4. All newly constructed plasmids were verified by DNA sequencing.

2.2.2 Colony PCR

Genomic DNA was extracted using the Chelex (Sigma-Aldrich) resin that binds cellular polar components while the RNA and DNA remain in water solution. A single colony was resuspended in 100 µL of 5 % Chelex resin. Then, cells were subjected to heat at 95 ºC for ten minutes followed by a centrifugation at 10000 x *g* for one minute. Finally, 50 µL of the supernatant were collected and one µL was directly used for PCR with target primers using the DreamTaq DNA polymerase (Thermo Scientific, Rockford, USA) according to the manufacturer guidelines.

2.2.3 Extraction of plasmid DNA

To identify the transformants carrying the desired construct, plasmid DNA was extracted and analysed by digestion with the appropriate restriction endonucleases. Single colonies from *E. coli* strain DH5α were incubated overnight in five mL of LB medium with the appropriate antibiotic at 37 °C with orbital shaking. Afterwards, two mL of this culture was centrifuged for five minutes at 13000 x *g* and the sediment was suspended in 360 µL of STET buffer (table 6), 24 µL of lysozyme (10 mg/mL) and 10 µL of RNAse (10 mg/mL). The mix was incubated at 37 °C for 30 minutes and then boiled at 100 °C for one minute. The tubes were then centrifuged at 13000 x *g* for ten minutes and the sediment was removed with a loop. Isopropanol was added to the remaining content to a final concentration of 70 % (v/v) and the mix was then centrifuged at 13000 x *g* for 45 minutes at 4 °C. Finally, the supernatant was carefully decanted, and the sediment was air dried. The sediment was then resuspended in 20 µL of MilliQ H_2O. To verify the presence of the insert, digestion with restriction enzymes was performed. The digested product was then verified through gel electrophoresis (see 2.2.4).

To obtain the ideal purity, quality, and concentration of the DNA for genetic transformation experiments, cloning and DNA sequencing, the isolation of the plasmid DNA from *E. coli* was performed using the "NZYMiniprep" extraction kit (NZYTech, Lisbon, Portugal) according to the

protocol provided by the manufacturer. “NZYMiniprep” extraction kit is based on the alkaline lysis of the cells and the adsorption of DNA to a silica matrix immobilized in a column.

2.2.4 DNA gel electrophoresis

DNA samples were subjected to gel electrophoretical analysis to verify the presence, size (bp) and integrity of specific DNA fragments Agarose gels (1% agarose) were prepared in TAE buffer 1x (table 6) with ethidium bromide (0.001 % (v/v)). Orange G loading buffer (table 6) was added to the DNA samples prior to loading into the agarose gel. The gel electrophoresis was conducted at a constant voltage of 100 V and DNA was visualized under UV light (205 nm). The size of the fragments was estimated by comparison with a commercial molecular weight marker, 1 Kbp plus DNA Ladder (Invitrogen, Carlsbad, California, USA.).

2.2.5 Competence development in *E. coli* and transformation

Competent *E. coli* cells were prepared as follows, LB medium (100 mL) was inoculated with 200 µL of an overnight culture and incubated at 37 °C with orbital shaking until an OD_{600} between 0.3-0.4 was reached. The culture was then incubated in ice for 15 minutes and centrifuged at 900 x *g* for 15 minutes at 4 °C. The resultant supernatant was then decanted, and the sediment suspended in 30 mL of ice-cold RF1 buffer (table 6). The resuspended sediment was then incubated in ice for 15 minutes and centrifuged at 900 x *g* for 15 minutes at 4 °C and the supernatant decanted. The resulting sediment was then resuspended in eight mL of ice-cold RF2 buffer (table 6).

For transformation, 10 µL of the ligation mixture or 1 µL of plasmid DNA were added to 200 µL of competent *E. coli* cells and incubated for 30 minutes on ice. A thermal shock was performed, at 42 °C for 90 seconds followed by two minutes in ice. One mL of LB medium was then added, and the culture incubated for two hours at 37 °C with orbital shaking. After the incubation period, the culture was then centrifuged for two minutes at 6000 x *g*. 900 µL of the resultant supernatant was discarded and the sediment was resuspended in the remaining volume. The *E. coli* cells were then plated in LA medium with appropriate selective antibiotic and incubated overnight at 37 °C.

2.2.6 Conjugation of plasmids into *Clostridioides difficile*

The conjugation process is a DNA transference method that requires direct cell-to-cell contact and a bridge-like connection between a donor and receiver cell. Using *E. coli* strain HB101 (RP4) as a donor, the plasmid DNA was transferred by conjugation into the *C. difficile* receiver strain as previously described (Bhattacharjee & Sorg, 2020). Briefly, the receiver *C. difficile* strain was streaked onto a BHI agar plate supplemented with cefoxitin (25 µg/mL) and incubated at 37 °C for 48 hours in anaerobic conditions (5 % H_2, 10 % CO_2, 85 % N_2). The next day, the *E. coli*

HB101 (RP4) containing the plasmid of interest, was streaked onto LB agar media plate supplemented with the appropriate antibiotics and incubated overnight at 37 °C. A inocula of each strain was performed until an OD_{600} of about 1 was reached. After, 1 mL of the *E. coli* (donor) strain was centrifuged, and the sediment resuspended in new LB media. The cells were then centrifuged a second time and the supernatant removed. The cellular sediment of the donor strain was then resuspended in 300 µL of the *C. difficile* receiver strain, and the resulting mixture was spot-plated onto non-selective BHI agar and incubated anaerobically at 37 °C overnight. The resulting spots were then retrieved with a loop and resuspended in one mL of BHI. The suspension was plated in triplicate in BHI supplemented with cefoxitin (25 µg/mL) and thiamphenicol (15 µg/mL) to counter-select for *E. coli* cells. The plates were then incubated anaerobically for 48 hours at 37 °C. Isolated colonies were streaked a second time in BHI plates with cefoxitin and thiamphenicol to unsure the selection of conjugated *C. difficile* cells. Conjugated strains were then kept at - 80 °C with glycerol for further use.

2.2.7 *Clostridioides difficile* mutant construction through Allele Couple Exchange (ACE)

In ACE mutagenesis, a specific plasmid originated from pMTL-YN3 (table 4) containing a truncated version of the target gene and its flanking regions (to allow recombination), is expected to integrate into the chromosome and replace the wild type gene by the truncated version of itself (Ng et al., 2013). Plasmid pSR86 (table 4) was introduced into *E. coli* HB101 (RP4) and then transferred to strain 630Δ*ermΔpyrE* by conjugation. Following two passages on BHI agar supplemented with 25 µg/mL cefoxitin and 15 µg/mL thiamphenicol, colonies that were noticeably larger (indicative of plasmid integration) were streaked onto *C. difficile* minimal medium (CDMM; Karasawa et al., 1995; table 5) supplemented with 5 µg/mL uracil and 2 mg/mL 5-fluoroorotic acid (FOA) to select for plasmid excision. The isolated FOA-resistant colonies were screened by colony PCR using primers Veg-vef-fw and Veg-vef-Rv (table 3). Double-crossover mutants, in which the mutant allele was successfully integrated yielded products smaller than those seen in WT revertants (931 bp instead of 1091 bp).

2.2 Biochemical techniques

2.3.1 Protein production by auto-induction in *E. coli*

E. coli BL21 (DE3) (table 2) transformed with the recombinant plasmid pBG1 (table 4) was grown in 100 mL of autoinduction medium (Studier, 2005, table 5) supplement with ampicillin (100 µg/mL) at 37 °C, with orbital shaking for 18 hours. After incubation, the cells were collected by centrifugation at 7000 x *g* for 30 min at 4 °C and the resultant sediment was resuspended in Start buffer (table 6) and stored at -20 °C until needed.

2.3.2 Protein purification by affinity chromatography

The cells resulting from the auto-induction growth were thawed, 0.1 µL of benzonase ® nuclease was added per ten mL of bacterial suspension and incubated for four minutes in ice before proceeding with cell disruption. The cell suspension was passed through a French press cell at 63.276 $kg.cm^{-2}$. After clearing cell debris by centrifugation at 18 000 x *g* for one hour at 4 °C, the soluble fraction was then kept in ice until it was used for protein purification through affinity chromatography. The resultant sediment was resuspended in ten mL of Start buffer. The fractions collected through the cell disruption process (total cell extract, sediment, soluble fraction) were analysed through 15 % SDS-PAGE (see 2.3.3).

The His_6-Tag fusion protein was purified over a 5 mL Nickel His-Trap column (GE Healthcare Bio-Sciences AB, Uppsala, Sweden) in a ÄKTA purifier 10 (GE Healthcare Bio-Sciences AB, Uppsala, Sweden) system. All the reagents used during the purification procedure were filtered before loaded into the system. All method programmes used in the ÄKTA system took in account the maximum internal system pressure. The maximum column pressure and the flow (mL/min) of injection into the column was set accordingly. All absorbance measurements in the ÄKTA system were performed with a wavelength of 280 nm. The column was first washed with five volumes of column (VC) of MilliQ water, followed by equilibration with five VC of Start buffer. The soluble fraction was then loaded into the column and the flow-through was collected. The column was washed with five VC of Start buffer, and the protein was eluted with a continuous gradient of imidazole from 20 mM to 500 mM of imidazole (Start buffer and Elution buffer, respectively; table 6) in an 80 mL range with a 0,8 mL/min flow. The elution fractions were collected using a Fractioner 900 with 2 mL per fraction. The fraction collected were then electrophoretic resolved into a 15 % SDS-PAGE (see 2.3.3).

The purified protein was dialysed overnight at 4 °C in SnakeSkin™ Dialysis tubing (3.5 KDa MWCO, 16 mm; Thermo scientific, Rockford, USA) for buffer exchange to CHES buffer (500 mM NaCl; table 6). The samples were then concentrated in a Vivaspin column (MWCO 3,000. GE Healthcare Bio-Sciences AB, Uppsala, Sweden) until a concentration of 450 µM (~ 4 mg/mL) was reached. Concentration measurements were made in a Nanodrop ND-2000-C (Thermo scientific, Rockford, USA) at a wavelength (λ) of 280 and 190 nm. The concentrated purified protein was stored in -80 °C until usage.

2.3.3 SDS-PAGE

Purified proteins and products of the cross-linking reactions were analysed by SDS-PAGE. Samples were loaded on 15 %, SDS-PAGE gels (resolving and stacking: table 6) after boiling for five min at 100 °C in loading buffer (table 6) . The Precision Plus Protein™ All Blue Ladder (BioRad, Hercules, USA) was used as a molecular weight marker. To stain, the gel was incubated for one hour in the coomassie solution (table 6) and then, in destaining solution (table 6) until the background was clear. Stained gels were imaged using an iBright system (Thermo scientific, Rockford, USA), using the universal tool with the Coomassie channel.

2.3.4 Western Blot

Proteins were electrophoretically transferred from SDS-PAGE gel to nitrocellulose membranes (Supported Nitrocellulose, 0.45 µm; BioRad, Hercules, USA) at 100 V for 90 minutes using transfer buffer (table 6). The membrane was then incubated in ten mL of 5 % milk in PBS-T (table 6) for one hour with agitation at room temperature. The blocking solution was then removed, and the antibody anti-his-tag (Merck, Darmstadt, Germany) was added at a dilution of 1:1000 in ten mL of PBS-T with 0.5 % milk. The membrane was incubated overnight with the antibody solution at 4 °C without agitation. The antibody solution was then discarded, and the membrane washed 3 times with PBS-T for ten min, each wash. A mouse peroxidase-conjugated secondary antibody (Sigma, Darmstadt, Germany) was added, at a 1:2000 dilution in 10 mL PBS-T with 0.5 % milk. The membrane was incubated for 30 min at room temperature with agitation. Finally, the membrane was washed 3 more times in PBS- T for ten min each wash. Proteins were detected using the "SuperSignal West Pico Chemiluminescent" reagents (Thermo Scientific, Rockford, USA) and imaged in the iBright system (Thermo Scientific, Rockford, USA) with the "Chemi Blot" tool. Images were adjusted, cropped, and quantified using the iBright system image analysis (Thermo Scientific, Rockford, USA).

2.3.5 Thermofluor assay for buffer screening

Thermofluor assays allows to determine the stability of the protein in solution through the melting temperature (T_m). Protein melting temperature determination was performed through monitorization of the protein unfold using a fluoroprobe SYPRO Orange dye (Molecular Probes), which upon binding to hydrophobic protein regions emits fluorescence that can be measured as a function of temperature. The thermal shift assay was performed on an iCycle iQ5 Real Time PCR Detection System (Bio-Rad), equipped with a charge-coupled device (CCD) camera and a CY3 filter with excitation and emission wavelengths of 548 and 595 nm, respectively. Optimization of the protein and dye concentration in the fixed assay reaction volume were performed in a primary assay to assess signal strength. Using a fixed protein concentration of 0.05 mg/mL, increasing dye concentrations were tested (from 1- to 10-fold, diluted from the initial 5000-fold stock in HEPES buffer (table 6). The 96-well plates were sealed with Optical Quality Sealing Tape (Bio-Rad) and centrifuged at 2500 x*g* for 2 min immediately before the assay to remove possible air bubbles. For the thermal denaturation the plates were heated from 10 to 100 °C with stepwise increments of 1 °C per minute and a ten second hold step for every point, followed by the fluorescence reading. The best signal-to-noise ratio was obtained using 7.5 µg of protein per well and ten- fold dye as final assay concentrations, and subsequently these conditions were used for the remaining of this work. Buffer formulation screening was prepared based on the Solubility kit from Jena Biosciences (100 mM buffer concentration and pH range 3–10) with the addition of increasing NaCl concentrations (0, 150 and 500 mM). The assay was prepared by adding 2.5 µL of protein–dye mixture solution previously prepared in HEPES buffer with 22.5 µL of the different

screening buffers. The reference experiment was prepared using the protein purification buffer (Start Buffer, table 6). This protocol was adapted from Santos et al., 2012.

2.3.6 Size Exclusion Chromatography

Size Exclusion Chromatography (SEC) assays allow to better comprehend the oligomeric equilibrium of a protein. Purified Veg was taken for buffer exchange by dialysis to CHES buffer with two concentrations of NaCl (500 mM NaCl and 100 mM NaCl; table 6), Tris-HCl (table 6) and CAPS (table 6). Each sample was load in a a SuperDex75 10/300 increase column (GE Healthcare Bio-Sciences AB, Uppsala, Sweden) previously equilibrated with two VC of the respective buffer. For the study of the oligomeric state of the Veg protein in the presence of nucleic acids, a 10 poly-U sequence or a 10 poly-T sequence (Metabion, Planegg, Germany; table3) was added to Veg in solution in a 7:1 proportion (Veg:nucleic acid) and incubated for ten minutes at room temperature before injection into the SuperDex75 10/300 increase column. The method used was adapted from the work previously described by (Kilic et al., 2006; G. M. Wilson et al., 1999).

The column was calibrated with the gel filtration molecular marker Gel Filtration Standard (BioRad, Hercules, USA; Thyroglobulin (bovine) 670 kDa; γ-globulin (bovine) 158 kDa; Ovalbumin (chicken) 44 kDa; Myoglobulin (horse) 17 kDa; Vitamin B12 1.35 kDa), which was prepared in the same buffer used for equilibration of the column. After each run, fractions containing protein and the nucleic acids were subjected to SDS-PAGE and immunoblot analysis.

2.3.7 Electric Mobility Shift Assay (EMSA)

Electric Mobility Shift Assay (EMSA) takes base in the electrophoretic mobility shift of the interaction complex between protein and nucleic acid. The DNA fragments were PCR amplified with primers labelled with CY3 fluorophore at the 5'-terminal (table 3). Different concentrations of Veg (500-1000 nM) were incubated with 3 ng/µL of each PCR product in EMSA buffer (table 6) for 30 minutes in the dark. After the incubation, samples were resolved by 6 % polyacrylamide gel (in TAE buffer) for two hours at 100 V, 4 °C. The gels were scanned in the iBright system (Thermo Scientific, Rockford, USA), with the "Fluorescent blots" tool. Images were adjusted, cropped, and quantified using the iBright image analysis package (Thermo Scientific, Rockford, USA).

2.3.8 Crosslinking assays with Glutaraldehyde

Crosslinking assays were performed as described previously by Fadouloglou et al., 2008. Glutaraldehyde (GA) was added to a mix of Veg (at 210 µM) and RNA in a 7:1 ratio (Veg:RNA), to a final concentration of 0.05 % (v/v). Time points were then retrieved from the reaction point just after addition (time 0) and at 15 s, 30 s, 1 min, 5 min, 10 min, 20 min and 30 min thereafter. Quenching buffer was added to the samples (table 6). After quenching, the samples were boiled

for five minutes at 100 ºC and loaded into an SDS-PAGE gel followed by Western Blot that were performed as previously described (2.3.3 and 2.3.4, respectively).

3. Results

3.1 Role of *veg* in *Clostridioides difficile*

A phenotypical analysis of the *veg* mutant could provide further insights into the function of Veg during growth, sporulation, and biofilm formation in *C. difficile*. Previous studies in *B. subtilis* have shown that Veg influences both biofilm formation and spore assembly (see 1.2.4; (Lei et al., 2013)).

3.1.1 Construction of a multicopy allele of *veg*

A classical approach to understand the function of a gene is to perform a deletion (knock-out mutant) and evaluate the resulting phenotype. After deletion of the gene of interest, a complementation strain would be constructed to verify if the wild-type (wt) phenotype would be restored. In the impossibility of the deletion of a gene, *i.e.,* an essential gene, other approaches, such as propagating the gene in a multicopy vector (knock-in mutant), may be taken.

Multiple attempts to construct a strain carrying a *veg* in frame deletion using allelic-coupled exchange (ACE) methodology (Ng et al., 2013) were performed (see 2.2.7). This system allows the construction of in-frame deletion mutants in any location in the genome using a heterologous *pyrE* gene, as a negative selection marker. More importantly, the mutant can be complemented concomitantly with restoration of the *pyrE* gene. 630Δ*erm* transconjugants obtained with pSR86 (table 4), were re-streaked two times onto BHIS plates with thiamphenicol and cefoxitin to select for the single crossover. Subsequently, single colonies were re-streaked onto minimal medium supplemented with FOA and uracil to select for cells in which the integrated plasmid had excised. Depending on which homology arm undergoes recombination, plasmid excision can result in either the desired double crossover mutant, or a wildtype cell. All the FOA resistant colonies yielded a 1091 bp DNA fragment, consistent with the presence of a wild-type copy of the gene. The difficulty in deleting the *veg* gene may be an indication that *veg* could be essential for growth.

In the impossibility to perform a knock-out of *veg,* a different tactic was taken, and the construction of a knock-in mutant was attempted. We were unable to clone the intact *veg* gene with its own promoter in *E. coli*. Different approaches were thought, such as the use of a promotor inducible with anhydro-tetracycline (P_{tet}), but the use of this compound strongly impaired biofilm formation. A form to curve this challenge was to couple *veg* to a promotor constitutively expressed in *C. difficile*. With this purpose, the promotor of *spo0A* was chosen since it is expressed at all stages of growth in *C. difficile* and in most cells in the population (Deakin et al., 2012b). To do this, we used genomic DNA from strain 630Δ*erm* of *C. difficile* to PCR-amplify the *veg* open reading frame (275 bp) using the oligonucleotide primers Veg-EcoRI-Fw and Rw-Veg-XhoI and P_{spo0A} (490 bp) using primers PspoA-EcoRI and SOE Pspo0A-Veg. The two resulting DNA fragments were fused by SOE-PCR (splicing by overlap extension PCR) using Pspo0A-Veg and

Rw-Veg-XhoI as the forward and reverse primers, respectively. The amplified fused product (795 bp), named P*spo0A*-*veg* was digested with EcoRI and XhoI and inserted between the same sites of the multicopy plasmid pMTL84121 (Heap et al., 2009) to yield pLB2 (*veg*^{MC}; table 4; figure 3.1A). After confirmation through agarose gel electrophoresis of the digested products (figure 3.1B) and sequencing, pLB2 was transferred into *C. difficile* by conjugation. The recipient strains were 630Δ*erm*, a lab control strain, and a clinical isolate, strain 1800 (RT126), a strong biofilm producer. This yielded strains AHCD1648 and AHCD1649, respectively. Two additional derivatives of the same strains were constructed for control purposes, AHCD543 and AHCD1650, respectively, harbouring pMTL84121 (empty plasmid, ep).

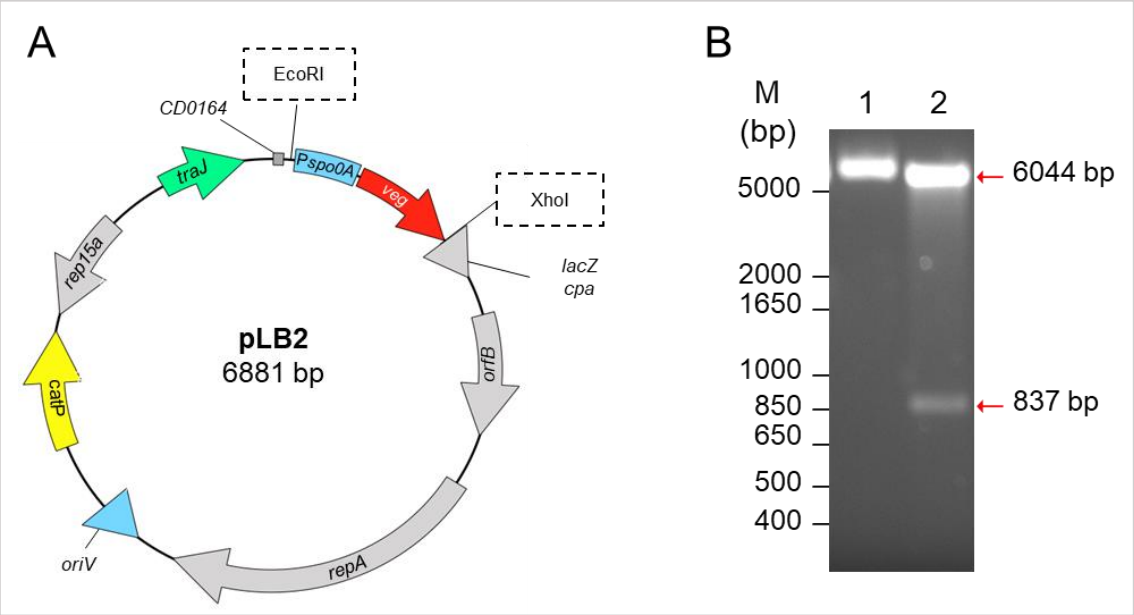

Figure 3.1 – Construction of *veg* under the control of the *spo0A* promotor in a multicopy vector, pLB2. For the functional characterization of Veg we started with construction of a plasmid in which *veg* is under the control of the *spo0A* promoter. pLB2 was inserted in the lab strain 630Δ*erm* and in the epidemic strain 1800. **(A)** Map of pLB2, the multicopy vector for *veg* multicopy (*veg*^{MC}). The inserted fragment is represented in blue and red, blue representing the *spo0A* promotor and red the *veg* coding region. Dotted lines represent the cleavage site of the endonucleases. **(B)** An agarose gel (1%) of the undigested plasmid, lane 1 and of the digested plasmid; digestion was performed with EcoRI and XhoI endonucleases, lanes 2.

3.1.2 Characterization of the *veg* multicopy strains

This multicopy alleles may act as gain of function mutations and if so, these alleles can provide insight into gene function. To characterize the strains carrying the *veg*^{MC} allele, assays for growth rates, biofilm production and sporulation were carried out.

3.1.2.1 Growth rate

Biofilm production and sporulation are stationary phase processes. To rule out that the differences in sporulation and biofilm formation between strains resulted from differences during vegetative growth, growth rates were calculated as described in 2.1.4. Briefly, a starter culture with a fixed OD_{600} were incubated and the OD_{600} was measured at hourly intervals for 14 hours.

In the 630Δ*erm* background, no significant differences were observed between the strain carrying the empty plasmid or the *veg*^{MC} allele (0.462 h^{-1} and 0.526 h^{-1}, respectively; figure 3.2A). However significant differences ($p < 0.05$) were observed in the 1800 background between the two strains (figure 3.2A). The presence of *veg*^{MC} increased the growth rate to 0.634 compared to 0.559 h^{-1} for the strain carrying the empty plasmid. These results indicate that *veg* may also have a role during growth, at least in some genomic backgrounds.

3.1.2.2 Biofilm production

As explained in section 1.2.4, in *B. subtilis,* Veg has an impact on biofilm formation by repressing SinR which in turn represses the matrix operon genes (Lei et al., 2013a). To analyse a possible role of Veg on biofilm formation in *C. difficile* we tested the ability of the strains bearing a *veg*^{MC} allele to form a biofilm in the presence of glucose and deoxycholate (DOC), using the crystal violet assay, as described in sections 2.1.2 and 2.1.3.

As expected, at 48 hours strains from the 1800 background produced around five times more biofilm than the 630Δ*erm* strains (Louro, 2020). Strain 630Δ*erm* and its *veg*^{MC} and empty plasmid derivatives showed no significant differences in biofilm formation ($p > 0.05$) at 48 hours of incubation; this was not the case for biofilm formation of strain 630Δ*erm* upon 24 hours of incubation, where the strain bearing the *veg*^{MC} allele produces significantly more biofilm than the one with the empty plasmid ($p < 0.05$). (figure 3.2B). The 1800 strains showed no significant differences in biofilm formation in the two incubation periods ($p > 0.05$) (figure 3.2B). These results suggest that the *veg* gene may positively affect the onset of biofilm formation in the lab strain (630Δ*erm*) but has no impact on the already strong biofilm producer 1800.

3.1.2.3 Sporulation

For an obligate anaerobe like *C. difficile*, spores are essential to survive in the aerobic environment. These dormant cells, resistant to oxygen and antibiotics, are a vehicle for transmission and dissemination as well as a way to persist inside the host (Paredes-Sabja et al., 2014b). Since spores are a key component of CDI, we analysed the ability of the various strains to sporulate (figure 3.2C). For that, we analysed their ability to form heat resistant spores after 24 h of incubation, as described in section 2.1.5.

Strains from the 1800 background produced more spores than the 630Δ*erm* strains (figure 3.2C). The 630Δ*erm* strains showed a normal and homogenous distribution between the strain with the empty plasmid and the *veg*^{MC} strain, and no significant differences were detected ($p > 0.05$), with an average sporulation efficiency of 45 % (figure 3.2C). In the 1800 strains, even though no significant differences are evident between sporulation efficiency of the two strains ($p > 0.05$), the distribution of the measurements for both samples is non-homogenous ($p < 0.05$),

but the statistical precision is higher in the *veg*^{MC} strain that in the control strain ($\sigma_C > \sigma_{Veg}$). Thus, concerning sporulation efficiency the *veg*^{MC} alelle in the 1800 background reduces variance when compared to the other strains tested (figure 3.2C).

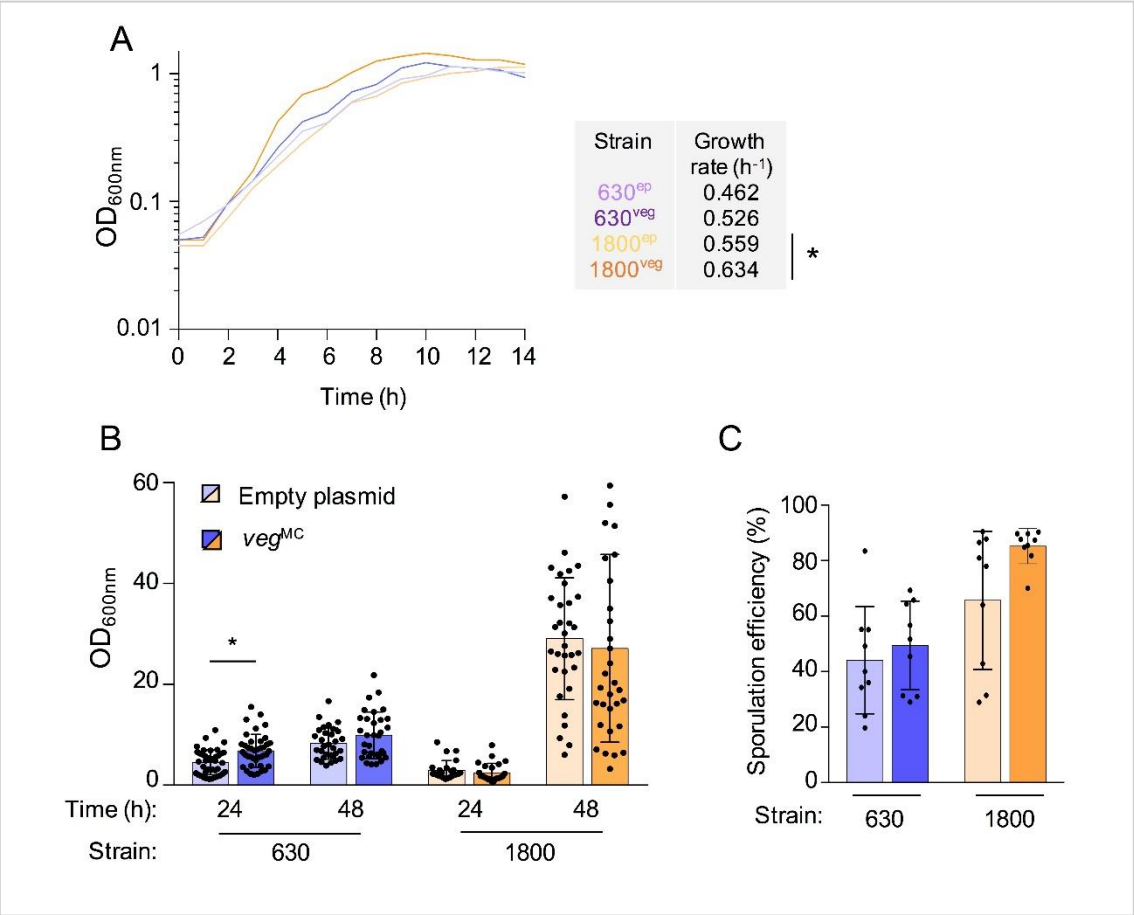

Figure 3.2 – Characterization of *veg* multicopy strains. The following assays were conducted: growth curves **(A)**, biofilm formation **(B)** and sporulation efficiency **(C)**. **(A) Growth curves upon Veg overproduction (OD_{600}).** Blue lines represent strain 630Δ*erm* and orange lines strain 1800. Darker bars represent strains containing *veg*^{MC} and brighter bars represent control strains carrying the empty plasmid. The * symbol represents significant differences between conditions ($p < 0.05$). **(B) Biofilm formation of strains bearing *veg*^{MC} (OD_{600})**. Biofilm formation was assessed as previously described (see the Material and Methods section), after 24 and 48 hours of incubation. Blue bars represent strain 630 Δ*erm* and orange bars strain 1800. Darker bars represent strains containing *veg*^{MC} and brighter bars represent control strains containing the empty plasmid. Error bars represent standard deviation. Statistical analysis was performed using a one-way ANOVA test when the samples were normal and homogeneous, if this was not verified, a non-parametric test, Kruskal-Wallis, was performed. The * symbol represents significant differences between conditions ($p < 0.05$). **(C) Efficiency of sporulation of strains bearing *veg*^{MC} (%).** Blue bars represent strain 630Δ*erm* and orange bars strain 1800. Darker bars represent strains containing *veg*^{MC} and brighter bars represent control strains containing the empty plasmid. Error bars represent standard deviation. Statistical analysis was performed using a one-way ANOVA test when the samples were normal and homogeneous, if this was not verified, a non-parametric test, Kruskal-Wallis, was performed. The * symbol represents significant differences between conditions ($p < 0.05$).

3.2 *In vitro* characterization of Veg

3.2.1 Overproduction and purification of Veg

To gain a deeper understanding into the mechanism of action of Veg we overproduced Veg-His_6 in *E. coli* (DE3), and purified the protein through affinity chromatography (see 2.3.2). The plasmid for overproduction of Veg-His_6 (pBG1) was described previously (Gonçalves, 2020).

Veg-His_6 production was auto-induced in 1 L cultures of *E. coli* (DE3) harbouring the pBG1 plasmid (table 4), in auto-induction media. After collected by centrifugation and resuspended in Start Buffer, whole cell lysates were prepared through mechanical disruption in a French Press. The soluble fraction was then loaded into a Superdex75 100/300 column pre-equilibrated with Start buffer, coupled to an ÄKTA system. Before an increasing gradient of imidazole was applied (Elution buffer), the chromatogram shows a first peak of absorbance at an elution volume of about five mL that corresponds to remaining unspecific binding to the column's matrix (figure 3.3A). When the imidazole gradient was applied, a single peak is observed, at an elution volume of approximately 47 mL with a maximum absorbance of 667 mAU. The maximum absorbance of the peak corresponds to a concentration of approximately 210 mM of imidazole (figure 3.3A). Veg elutes at 210 mM of imidazole and is partially purified, with the peak corresponding to the elution fractions 15-17.

Samples from each step of the purification was analysed by 15 % SDS-PAGE (figure 3.3B). Briefly, Veg-His_6 is produced as a single species of about 11 kDa, close to the expected mass of the protein (10.9 kDa) (figure 3.3B; red arrow). Veg is present in the soluble fraction (S) as well as in the insoluble fraction (P). Veg is still present in the Flow-Through of the column (FT) probably due to saturation of the column's matrix or insufficient contact with the matrix. Veg elutes partially purified in the peak fractions, 15, 16 and 17 (figure 3.3B).

The concentration of Veg was determined using a Nanodrop system with the estimated MW and Molar extinction (5960 M^{-1} cm^{-1}) coefficient of Veg. Veg elutes at a concentration of 83 µM (0.9 mg/mL). Veg was then dialysed into the desired buffer, usually CHES buffer (table 6), concentrated to 450 µM, and kept at - 80 °C until needed for further assays.

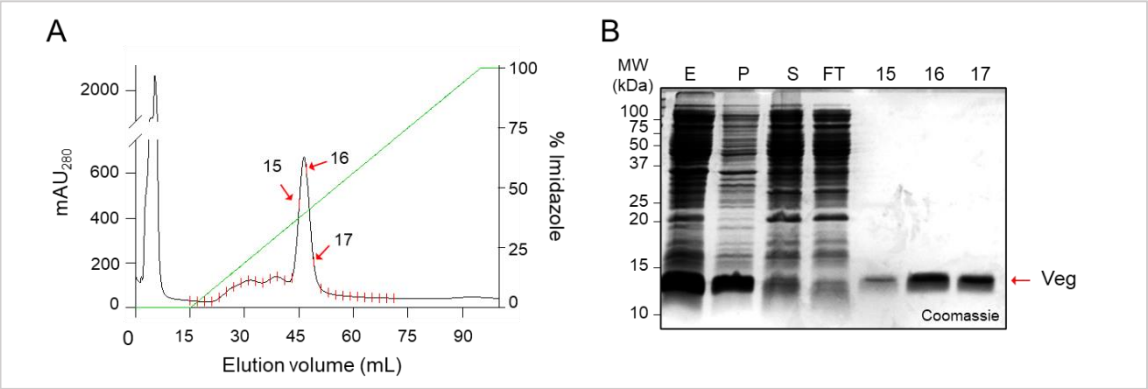

Figure 3.3 – Purification of Veg by affinity chromatography in a Ni^{2+}-His-trap column. (A) Chromatogram of elution of a Ni^{2+}-His-trap 5 mL column after application of the soluble cellular fraction. Black line represents the absorbance from the elution of the column at a wavelength of 280 nm and the green line represents the percentage of imidazole in the elution buffer. Red lines count the collected fractions of the elution. Fractions associated with a peak were resolved by SDS-PAGE**. (B)** SDS-PAGE gel representative of the process of Veg purification: **Ext-** total cellular extract from cell disruption; **P-** Insoluble fraction of the extract; **S-** Soluble fraction of the extract; **FL-T-** Flow through of the column after sample application; fractions, **15**, **16** and **17** correspond to the collected elution fractions marked in the chromatogram. Samples were resolved by 15 % SDS-PAGE.

3.2.2 Buffer screen using a Thermofluor reveals increased stability of Veg in high salt molarity and pH.

The Thermofluor assay has the ability to screen in parallel a multiplicity of buffer conditions and requires only microgram amounts of protein. The Thermofluor assay determines the protein melting temperature (T_m) and is based on the interaction between a dye and the hydrophobic regions of the protein, which are exposed upon protein thermal denaturation. The assay consists initially, on the non-interaction between the dye and the native folded protein. As the temperature increases the protein undergoes thermal denaturation, exposing its hydrophobic patches and thus causing a drastic increase in the fluorescence signal due to the strong interaction of the dye with those regions. Thermofluor has been shown to be quite effective in identifying the best buffer formulation to be used for protein purification and storage (Santos et al., 2012).

The data derived from Thermofluor assays were analysed taking into account two main factors: the midpoint temperature of the protein-unfolding transition (T_m – melting temperature) and the transition slope. The ideal buffer formulation is that where only one transition (folded-unfolded state) is observed with a T_m increase relative to the initial purification buffer. This increase means the protein is more stable and therefore more energy (heat) is required to unfold it. Since protein unfolding is a cooperative process, the unfolding of a small protein region should induce the immediate unfolding of the remaining protein core; thus, an optimal protein stabilization buffer should present a sharp and fast thermal denaturation transition between the folded and unfolded states, detected through high transition slopes, in parallel with a higher T_m. A preliminary screen was performed in order to choose the best signal-to-noise ratio, where different solutions

of dye were diluted into 50 mM HEPES pH 8.0 (1, 2, 5 and 10-fold) and using a protein concentration of 0.05 mg/mL.

The different Veg denaturation curves were analysed and the first derivative was calculated in order to determine the protein melting temperature for each assay. Based on these assay conditions, a buffer screen was carried out, consisting of 96 different buffers, each with a concentration of 10 mM and a pH range from 3 to 10. Each buffer was prepared with different salt (NaCl) concentrations: 0, 150 and 500 mM. The T_m values determined from each condition were then compared with the T_m value from the control experiment. From the 96 conditions tested, the lower pHs, 3-4, gave no measurable transitions, probably due to the very low pH induce protein destabilization, which was independent of the concentration of salt in the buffer (data not shown). Figure 3.4 shows the curves from the fluorescence data comparing the three buffers used in this study in combination with different salt concentrations. For the establishment of the optimal pH, three buffers were chosen with a pH lower than, around as, and higher than the isoelectric point (pI) of Veg (9.2): 10 mM Tris-HCl, pH 7.5, 10 mM CHES, pH 9.5, and 10 mM CAPS, pH 10, respectively. The three buffers have the same NaCl concentration (500 mM) (figure 3.4A). The lower pH shows no thermal transition, indicative of protein instability in the buffer. The higher pHs tested, 9.5 and 10, showed the generation of a melting transition starting around the 50 °C. From this, we can conclude that the protein is more stable in a pH closer or higher than the proteins pI. From here, we carried the remaining assays for stability with CHES buffer since it showed the narrowest melting transition. For the establishment of the NaCl concentration, the buffer and pH were fixed, 10 mM CHES, pH 9.5 (figure 3.4B). The same buffer was tested with different NaCl concentrations, 0 mM, 150 mM, and 500 mM. 0 mM showed a high absorbance at low temperatures, indicative that Veg was already unfolded and unstable. The same was not observed in the 150 mM concentration, were there was a clear melting transition, with a defined curve. However, the highest salt concentration had the best defined melting transition, with the highest T_m, indicative of the better stability of the three concentrations tested. A third comparison was performed, between the buffers further used in this study (see below; figure 3.4C). All three buffers had a concentration of NaCl of 150 mM NaCl. In CHES and CAPS a melting transition curve is observed, indicative of protein stability. On the contrary, in Tris-HCl buffer, the chromatogram did not show a melting curve, indicative of protein instability.

Considering the melting temperature and protein-unfolding transient slope of Veg, we conclude that Veg is better stabilized by a 10 mM CHES, pH 9.5 buffer with addition of 150-500 mM NaCl (figure 3.4 B;C).

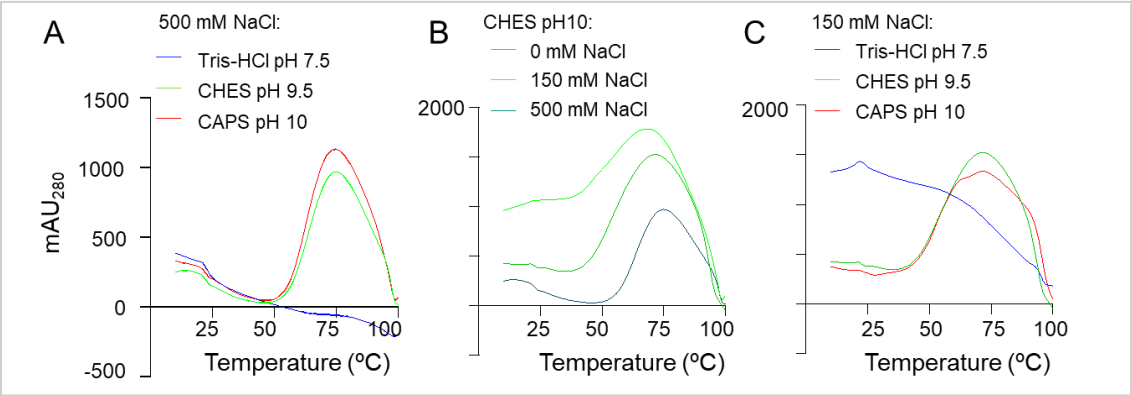

Figure 3.4 – Buffer screen for protein stability using a Thermofluor assay. pH and NaCl concentrations were the parameters varied. **(A)** The three buffers have fixed NaCl concentration (500 mM), a pH lower than, closer to, and higher than, Veg isoelectric point (pI ~ 9,2) were tested. 10 mM Tris-HCl, 500 mM NaCl, pH 7,5; 10 mM CHES, 500 mM NaCl, pH 9,5; and 10 mM CAPS, 500 mM NaCl, pH 10, were tested, respectively. **(B)** Three concentrations of salt were tested in the same buffer, 10 mM CHES, X mM NaCl, pH 9,5. **C:** Stability of the protein in the three buffers used for Gel Filtration Assays.

3.2.3 Veg behaves as a monomer by Size Exclusion Chromatography

Veg contains an Sm-like domain (Lei et al., 2013a), which are known to form multimers with one another and bind, as heteromeric complexes to various RNAs, recognizing primarily short U-rich stretches. This family of proteins include regulatory proteins that form a ring-shape homo-oligomer of 6 or 7 units, which bind to nucleic acids (Achsel et al., 2001b; Sauer et al., 2002). In some cases, it was shown that the oligomeric behaviour was dependent on pH and on the presence of a nucleic acid ligand (Kilic et al., 2006). As explained in section 1.2.4, the homology model constructed for Veg shows the characteristics of the Sm-like domain and a possible heptameric oligomer (Bandyra et al., 2013; Mura et al., 2013)

To better understand the oligomeric state of Veg-His_6 in solution, size exclusion chromatography assays, hereinafter termed SEC, were performed. Different conditions were used to test for oligomer formation: protein concentration, pH, and the presence of a short oligo-U ssRNA or a short oligo-T dsDNA (figures 3.5 and 3.6).

We first determine that Veg-His_6 was more stable in high ionic force (500 mM NaCl) (see above, section 3.2.2). Nonetheless, increasing the ionic force decreases ionic interactions (Curtis & Lue, 2006; Deller et al., 2016; Ibragimova & Wade, 1998); therefore, after assessing the influence of Veg concentration and the presence of each nucleic acid, we repeated, the assays at a lower ionic force (100 mM NaCl). At a first approach, the effect of Veg concentration at different pH on its oligomeric state was tested (figure 3.5 A-C). A range of Veg concentrations (between 50 and 250 µM) at different pH values (7.5, 10 and 11.5) were injected into a superdex75 increase 100/300 column. Higher concentrations were not tested due to protein aggregation and precipitation. Dextran blue (DB; MW ~ 670 kDa) was also injected for column internal control. Dextran blue elutes from the column at an elution volume of 8 mL. At pH 7.5 and

10 only one major SEC peak is observed with the approximate molecular weight of 9.4 kDa and 10.6 kDa, respectively. This peak contains the protein as shown by immunoblotting of the peak fractions (figure 3.5 D-F). We conclude that in these conditions, Veg-His_6 behaves as a monomer (MW= 10.5 kDa; 88 amino acids). At pH 7.5, peaks corresponding to low molecular weight species, in which the protein was not detected by immunoblot, may result from proteolysis of Veg-His_6. Tris-HCL pH 7.5 was the buffer in which the protein was less stable (see above). At pH 11.5 SEC of purified Veg-His_6 shows two main peaks (figure 3.5C). The larger peak corresponds to a species of 18.3 kDa closer to that of a Veg dimer. The smaller peak has a calculated size of 7.9 kDa closer to that of monomeric Veg-His_6. These values, close to the expected sizes of the dimer and monomer, represent a 2:1 ratio, respectively. In these conditions, with a pH above the pI, the protein has a global negative surface charge, which is in agreement with the homology model. These findings suggest that Veg may change its oligomeric state in a concentration-independent but pH-dependent manner. Interestingly, immunoblot of the purified Veg-His_6 shows four species that exhibited molecular weights compatible with a monomer (the most represented species), a dimer, traces of a possible trimer and also a high molecular form (figure 3.5D).

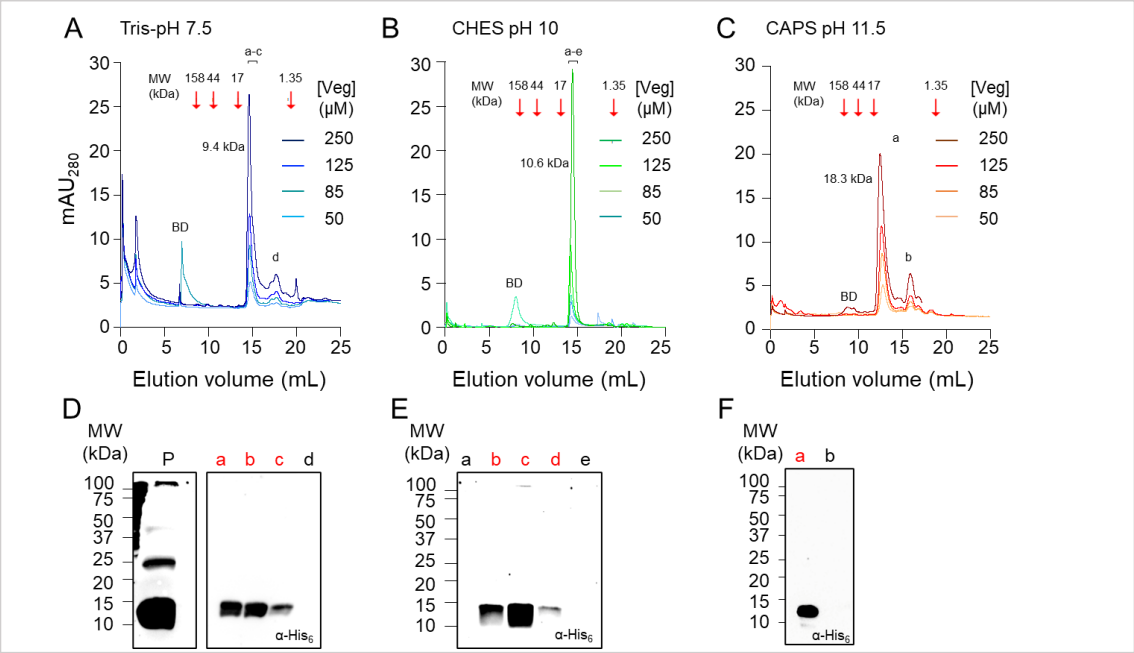

Figure 3.5 – Analysis of the oligomeric state of Veg by gel filtration. (A) A Superdex75 column was used to analyse the behaviour of Veg at different concentrations. Veg in 10 mM Tris-HCl, 500 mM NaCl, pH 7,5 (**B**). Veg in 10 mM CHES, 100 mM NaCl, pH 9,5 (**C**) Veg in 10 mM CAPS, 100 mM NaCl, pH 10. mAU corresponds to milli-units of absorbance at 280 nm and MW to Molecular Weight in kDa. Each peak was analysed through SDS-PAGE and Western Blot, **(D-F)**: Western blot analysis of the peaks of each chromatogram (a-e). P lane represents the total sample injected into the gel filtration system and each letter (a-e) a different peak. mAU corresponds to milli-units of absorbance at 280 nm and MW refers to molecular weight in kDa. Presence of Veg in the peak was indicated in red.

We next examined the oligomeric state of Veg-His_6 in the presence of a short oligo-U ssRNA or a short oligo-T dsDNA (figure 3.6). These assays were performed in the same conditions as the previous (figure 3.5 A-B). Veg-His_6 at pH10 and 7.5 was incubated with each

nucleic acid for 10 minutes before injection into the column and each nucleic acid was injected in separate as a control. SEC of purified Veg-His_6 incubated with either nucleic acid at both pH shows one main peak (figure 3.6). This single peak is localized at the same elution volume as the protein alone (monomer) and also corresponds to the nucleic acid when injected in the absence of Veg. This major peak contains the protein as shown by immunoblotting of the peak fractions (figure 3.6C). Taken together, our results suggest that at pH lower or similar to the pI, Veg-His_6 behaves as a monomer. Moreover, the presence of short DNA or RNA oligos does not influences its oligomeric state. Unfortunately, due to time constrains we were not able to perform these assays at pH 11.5.

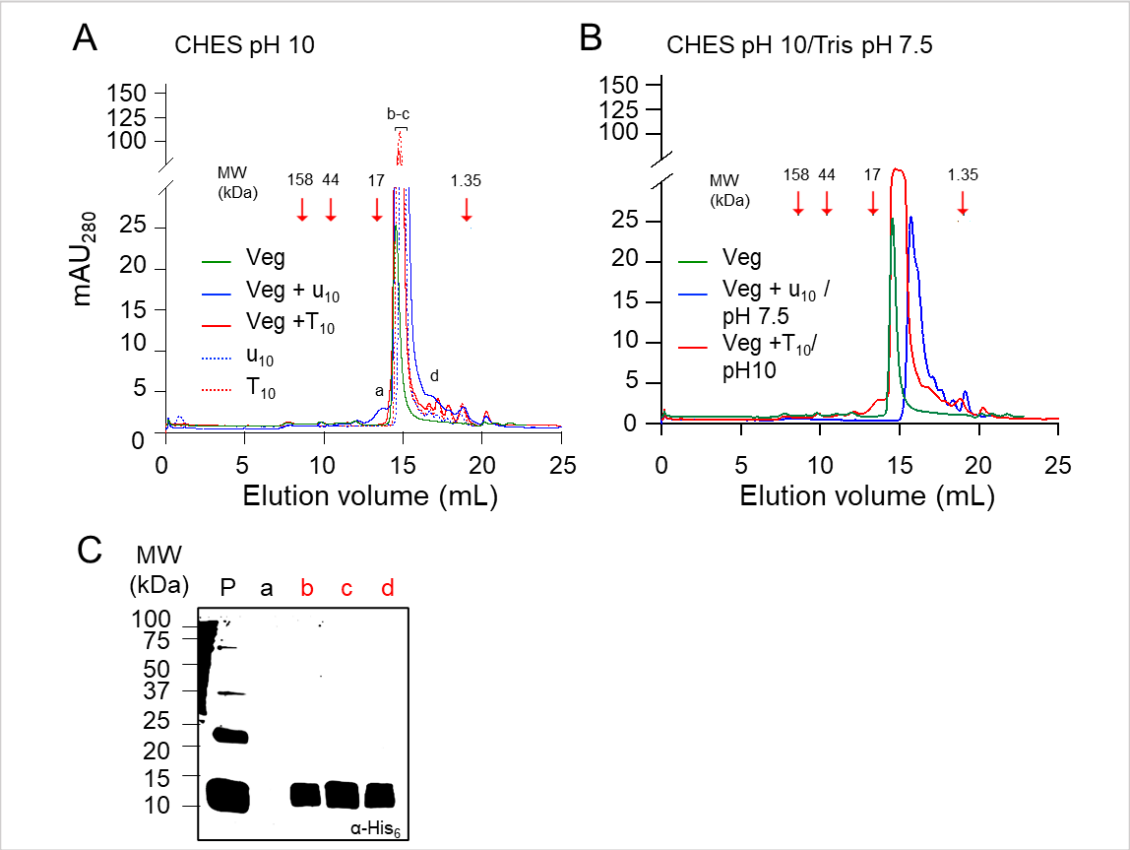

Figure 3.6 – Analysis of the oligomeric state of Veg through gel filtration in the presence of nucleic acids. For these studies, a Superdex75 column was used. (A) In the presence of nucleic acids (T_{10} and u_{10}), at high (10 mM CHES, 500 mM NaCl, pH 9,5) (B) and low ionic force (10 mM CHES, 100 mM NaCl, pH 9,5 and 10 mM Tris-HCl, 100 mM NaCl, pH 7.5). **(C)** Each peak was analysed through SDS-PAGE and Western Blot analysis of chromatogram A (a-d). P lane represents the total sample injected into the gel filtration system. mAU corresponds to milli-units of absorbance at 280 nm and MW refers to molecular weight in kDa. The position of Veg in the peak is indicated in red.

3.2.4 Veg oligomerization

As an independent test of oligomer formation by Veg, we used glutaraldehyde crosslinking assays. Purified Veg-His_6 was subject to *in vitro* crosslinker experiments with glutaraldehyde (GA) and analysed by immunoblotting with an anti-Histag antibody. GA is a

crosslinker with a short arm that covalently binds two proteins that are close to each other and presumably interacting with each other. Crosslinkers with short arms are indicative of a close proximity between two elements, making the covalent bonds formed more specific.

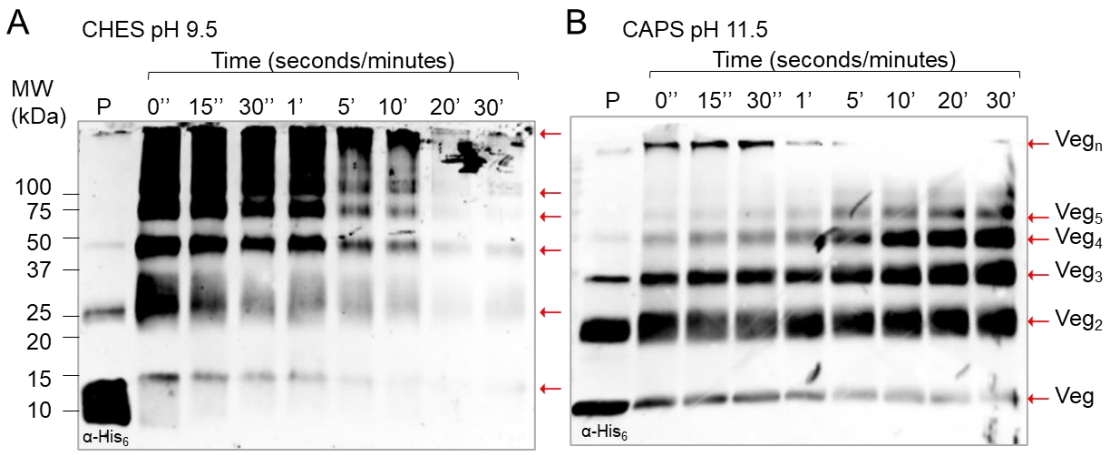

Figure 3.7 – Glutaraldehyde crosslinking of Veg. Crosslinking of Veg was performed with glutaraldehyde (0.05 %) along the time to evaluate the formation of oligomeric forms in different conditions. **(A)** Crosslinking in 10 mM CHES, 100 mM NaCl, pH 9.5. **(B)** Crosslinking in 10 mM CAPS, 100 mM NaCl, pH 11.5. Each time point was then resolved through SDS-PAGE and subject to immunoblot analysis with anti-Histag antibodies. P represents the purified protein (without crosslinker). The position of molecular weight (MW, in kDa) markers is shown.

Crosslinking was performed with purified Veg-His_6 at two different pH values, 9.5 and 11.5. GA was added in a final concentration of 0.05 % (v/v) and incubated for 30 minutes with samples taken at the indicated time points (figure 3.7). Each time point was then quenched using a final concentration of Tris-HCl 1 M and analysed in a 15 % SDS-PAGE followed by immunoblotting. At pH 9.5 (figure 3.7A), a fast disappearance of the band corresponding to the monomer (predominant at time 0) is observable with the concomitant detection of bands that may correspond to dimers and trimers (figure 3.7A). Aggregation of Veg in the microtube was observed during the experiment, forming a cloud with a cottage cheese-like texture and appearance. In later time points, 10'-30', absence of protein in the gel could be due to the large aggregates unable to enter the gel. Consistent with the SEC experiments (see 3.2.3), at pH 11 the predominant form of Veg is a dimer (figure 3.7B). However, three higher order protein complexes became apparent during crosslinking, which includes a trimer, a tetramer and a pentamer (figure 3.7B). Additionally, the aggregation visible in the microtube at pH 9.5, was not observed in these conditions, possible due to an increased stability of the protein at this pH as a result of the global negative surface charge. In any case the pattern of bands detected in both conditions is similar possibly corresponding to complexes already present in solution. Although is not known whether Veg forms higher order complexed *in vivo*, these results indicates that Veg is capable of forming a series of higher order oligomers.

3.2.5 Veg binds to DNA

As mentioned previously, *B. subtilis* Veg is unstable in some culturing conditions. It was previously suggested that a protease, yet to be identified, would be involved in proteolysis of Veg (Ramalhete, 2021). The synteny of *veg* and *yabG* (figure 1.5A) suggest that the specific protease contributing to Veg instability could be YabG. YabG is a cysteine protease required in *C. difficile* for processing of several spore surface proteins (Shrestha et al., 2019). YabG is also essential for the transcription of at least two genes, *cdeM* and *cotA* (Marini, 2020). CdeM and CotA are key determinant of the spore surface layers assembly (Pizarro-Guajardo et al., 2016, 2020). One hypothesis is that YabG removes by proteolysis a repressor of *cotA* and *cdeM* expression. We postulate that Veg could be the putative repressor. Although Sm domains, as the one present in Veg, have been showed to bind RNA (Achsel et al., 2001b), we decided to test direct binding of Veg to DNA, specifically to the promoter region of *cdeM* (figure 3.8).

To assess if Veg interacts with the promoter region of *cdeM*, a Electrophoretic Mobility Shift Assay (EMSA) was performed. In this assay, a DNA fragment marked with a fluorophore (CY3), is incubated with increasing concentrations of Veg-His_6, and the mix is resolved through a nondenaturating polyacrylamide gel (see 2.3.7).

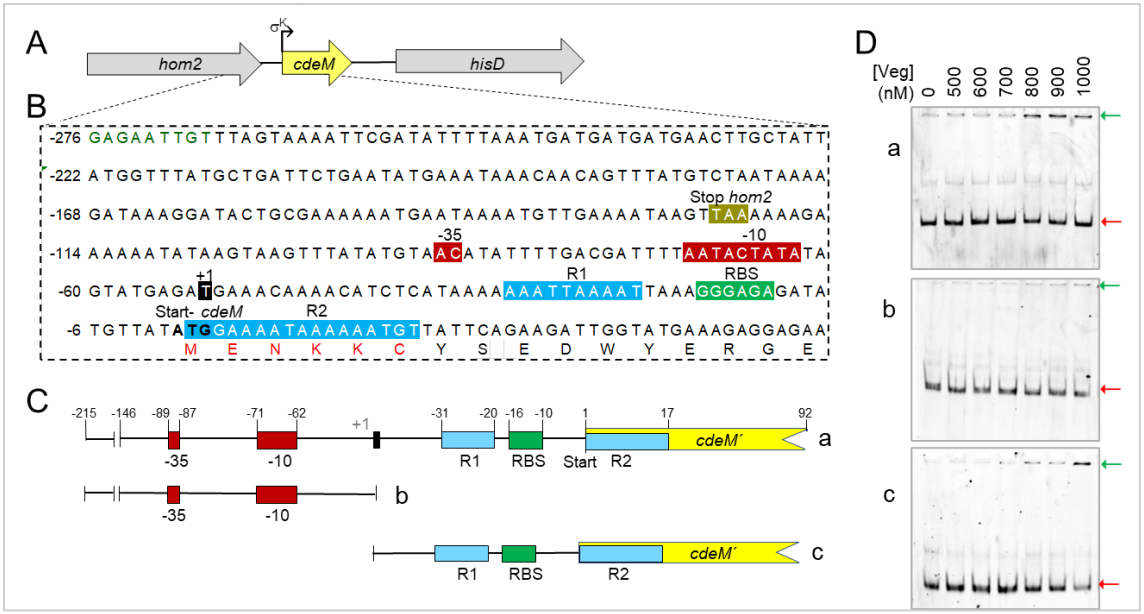

Figure 3.8 – *cdeM* locus and interaction with Veg. (A) Genomic region surrounding the *cdeM* gene in the *C. difficile* genome. **(B)** Sequence of the promotor region of *cdeM*. Sigma K binding region is represented in red (-35 to -10). Ribosome Binding Site (RBS) is represented in green. R1 and R2 are two regions AT rich, where is though that Veg would interact, and are represented in blue. **(C)** Map of the three fragments tested in EMSA assays. The first fragment is the full length, containing both the σ^K consensus binding site and the two AT repeats regions (-215 to +92) (a). The second fragment only includes the σ^K consensus binding site (-215 to +1) (b). The third fragment only includes the two AT repeat regions (R1 and R2) (-3 to +145) (c). **(D)** Electrophoretic mobility shift assay (EMSA) of Veg with the *cdeM* promotor region. EMSA assays were performed in increased concentrations of Veg ranging from 500 nM up to 1000 nM. All nucleic acid had a 5'-CY3 fluorophore and were added at 3 ng/well. Promotor fragments of the *cdeM* region were used. The regions of the *cdeM* promotor were the indicated in panel **C** (a, b, and c). 0 represents the control where no protein was added.

A 301 bp DNA fragment containing the promotor region of *cdeM* (-215 to + 92, relative to the start codon was PCR-amplified using genomic DNA from strain 630Δ*erm* and primer pairs PCdeM-CY3-Fw/PCdeM-CY3-Rv (figure 4.1; fragment a). The DNA fragment was incubated without protein or with increasing concentrations of Veg-His_6 from 500-1000 nM. No DNA shift was detected at Veg-His_6 concentrations ranging from 500 to 700 nM. At 800 nM and higher concentrations the DNA shifted and was now retained in the well of the gel (figure 4.1Da). This result shows that Veg directly binds to the promoter region of *cdeM*. A threshold concentration of Veg may be needed for binding to occur, but once this concentration is reached, other Veg molecules bind, possibly spreading along the DNA, from one or more nucleation sites, a through protein-protein interactions, which may lead to the promoter being almost completely covered by Veg.

In an effort to characterize the binding sites of Veg more closely in the promoter of *cdeM*, we did EMSA with two shorter DNA fragments: one including the core σ^K consensus binding site (-215 to -35, relative to the star codon, fragment b) and a second one including two AT-rich direct repeats (-3 to + 145, relative to the star codon; figure 4.1; R1 and R2, fragment c). The 175 bp and 148 bp DNA fragments were PCR-amplified using genomic DNA from strain 630Δ*erm* and primer pairs PCdeM-CY3-Fw/PCdeM-Reg-HindIII-Rv and PCdeM-Reg-EcoRI-Fw/PCdeM-Reg-HindIII-Rv , respectively (figure 4.1B; fragment b and c). No binding of Veg-His_6 to the fragment bearing the core σ^K promoter was detected (figure 4.1Db). In contrast, a DNA shift was detected in the fragment containing the repeat regions at 1000 nM concentration of Veg-His_6 (figure 4.1Dc). Together these results suggest that Veg directly binds to the *cdeM* promoter at the AT-rich R1 and/or R2 direct repeats.

4. Discussion

In *C. difficile* little is known about the regulatory network of interactions among diverse types of cell and molecules including DNA, RNA, proteins, and metabolites (Antunes et al., 2012a; Malabirade et al., 2018; Mani & Dupuy, 2001b; Müh et al., 2019; Pereira et al., 2013) that controls cell differentiation during biofilm formation. In this study we focused on the characterization of Veg, a small basic Sm-like protein of *C. difficile* previously implicated in sporulation and biofilm formation.

4.1 *In vivo* characterization of Veg

We were unable to construct a *veg* deletion mutant in the lab strain 630Δ*erm* and one possibility is that the gene is essential. A previous screen, using random transposon mutagenesis of the epidemic strain *C. difficile* R20291, however, did not identify the *veg* gene (CDR20291_3405) as essential (Dembek et al., 2015). These results may indicate different requirements for *veg* during growth, between strains. In future work, to test if *veg* is an essential gene, at least in 630Δ*erm*, we consider silencing this gene using a xylose-inducible CRISPR-interference vector (Müh et al., 2019). In this approach, fusion of antisense nucleotides to the guide RNA of Cas9 blocks gene transcription in a xylose depended on manner, enabling the function of essential genes to be studied.

In the impossibility to perform a knock-out of *veg,* a different approach was taken, and we constructed a knock-in mutant. Repeated attempts to create a construct that would allow expression of *veg* from its native promoter in a multicopy plasmid failed, likely due to the toxicity of Veg in *E. coli* when expressed from its own promoter (cloning intermediate). This may indicate that the *veg* promoter is recognized by the σ^A for RNA polymerase from *E. coli*, leading to Veg accumulation and toxicity. Thus, we decided to fuse the *veg* gene to a different promoter, that of the *spo0A* gene (Deakin et al., 2012b). The multicopy plasmid was inserted in the lab strain 630Δ*erm* and in a clinical isolate, strain 1800. Strain 1800 is a strong biofilm-producing strain belonging to the livestock-associated ribotype RT126 (Louro, 2020).

As observed in previous studies of the lab, the clinical isolate produced more biofilm and spores than the control stain 630Δ*erm* that that is used in many laboratory as a standard laboratory strain ((Louro, 2020), figure 3.2). It was suggested that the 630Δ*erm* strain has become "domesticated," losing its ability to carry out many behaviours' characteristic of its wild ancestors, including biofilm formation. This has been proven for other pathogenic bacteria like *Escherichia coli*, *Salmonella*, and *S. aureus* (Liu et al., 2017). Interestingly, introduction of a *veg* multicopy allele increased early biofilm formation only in the strain 630Δ*erm* (figure 3.2). This result suggests that overexpression of Veg may compensate for some lost genetic traits during domestication in the lab strain. A possible explanation might derive from the repression of SinR by Veg as it occurs in *B. subtilis* (Mandic-Mulec et al., 1992). In fact, in a 630Δ*erm* background, a *sinR* mutant was

shown to produce a denser biofilm (Poquet et al., 2018). The direct interaction between Veg and SinR and its inhibitory effect on SinR binding to DNA should be checked in order to test this hypothesis.

In sporulation assays, no significant differences were found between the *veg*^{MC} strains and its control strains (figure 3.2). However, in strain 1800, overproduction of Veg decrease the heterogeneity found between sporulation assays (figure 3.2). At the moment, we cannot distinguish if heterogeneity between different replicate populations results from the number of cells entering sporulation or the number of sporulating cells able to form heat resistant spores. To discriminate between these two possibilities, the different populations should be followed during sporulation by optical microscopy. These experiments would also give us a clue about the stage at which Veg is acting, at least in terms of reducing heterogeneity of sporulation. In any case, overexpression of Veg seems to decrease the noise in the formation of a heat resistant spores in the clinical isolate 1800. The decision to form spores is a major commitment of time and resources. Heterogeneity of cell fate can be thought of as a strategy to prevent the entire population of cells from entering into sporulation in response to conditions that may be transient.

4.2 *In vitro* characterization of Veg

The presence of an Sm-like domain may give a hint about Veg behaviour *in vitro* and *in vivo*. Other proteins where this ancient RNA-binding motif is present include Hfq from *B. subtilis*, which forms a homo-hexamer around an RNA motif; and AF-Sm1 and AF-Sm2 from *Archaeoglobus fulgidus,* which are small ribonuclearproteins, shown to form an homo-heptamer ring structure similar to those found in other archaeal Sm1-type proteins (Törö et al., 2002). These proteins have an RNA-binding motif and show preferential binding to U-rich oligonucleotides (Törö et al., 2002) whose oligomerisation is dependent on the presence of small nucleic acids. Because of the similarity of Veg to these proteins, assays were conducted to understand if Veg would follow the same behaviour.

We found that in solution Veg is present as a monomer at pH below or similar to the pI (9.2) and as dimer at a pH above the pI (figure 3.5). Unlike other Sm-like proteins, the oligomeric state of Veg does not change in the presence of short oligomeric nucleic acids (figure 3.6). However, crosslinking assays show that Veg is also capable of forming a series of higher order oligomers that differ between each other by one monomer (figure 3.7). The reason why these oligomeric forms are not detected by SEC is not presently understood.

Interestingly, the pI of Veg is higher than the pI of other Sm-like protein (Kilic et al., 2006) and even from *B. subtilis* Veg, which has a calculated pI of 4.8. The fact that *C. difficile* Veg is a small basic protein, rich in positively charged residues, able to form long oligomers prompted as to hypothesize that Veg may behave as a nucleoid-associated protein (NAP). NAPs were shown to have a role in chromosomal organization and transcriptional regulation (Dahlke & Sing, 2018).

Transcriptional silencing can be achieved by DNA bridging, in which NAP binds two sites that are further apart and causes the DNA between them to form a loop or binding to two sites close together results in DNA coating blocking the RNA polymerase (Dahlke & Sing, 2018; Dorman, 1991; Fang et al., 2020; Malabirade et al., 2018).

In all sporeformers the *veg* gene is just downstream of the *yabG* gene, which encodes for a cysteine protease required in *C. difficile* for processing of several spore surface proteins (Shrestha et al., 2019) (figure 1.5). YabG is also essential for the transcription of at least two genes, *cdeM* and *cotA* (Marini, 2020). CdeM and CotA are key determinants for the assembly of the spore surface layers (Pizarro-Guajardo et al., 2016, 2020). One possibility is that YabG proteolytically removes a repressor of *cotA* and *cdeM* expression. Consistent with this model, work from the lab have shown that the presence of the *cdeM* promoter region in a multicopy plasmid bypasses the need for YabG for *cdeM* expression, most likely by titrating the repressor (Olivença, 2021). We thus considered the possibility that Veg could be the putative repressor of *cdeM* transcription. Using EMSA we show that Veg binds to DNA, specifically to a region containing two direct AT-rich direct repeats in the *cdeM* promoter (figure 3.8). A particular behaviour is observed in the EMSA gels, in that the shift is not gradual but rather rapid, and a low-mobility complex is formed that remains in the well, which suggests a large complex. There are descriptions of similar behaviours in other proteins, for instance the histone-like nucleoid Structuring protein (H-NS), which is a negative regulator of the lateral flagellar system in a *Shewanella piezotolerans* (Jian et al., 2016; Jiang et al., 2000; Zhao et al., 2021). Veg may bind to DNA, presumably to the repeats in the *cdeM* promoter, and this first binding may serve to nucleate Veg polymerization leading to efficient gene silencing through transcriptional repression (figure 4.1).

If so, it is unclear what form of Veg binds to DNA. A monomer could bind to DNA and spread laterally through Veg-Veg interactions, or the binding unit could be a dimer. Although our model of Veg suggests the possibility that it may, like other Sm-like proteins, form an hexamer or heptamer, we feel it is unlikely that Veg binds to the DNA as such, because in no condition were we able to observe species larger than the dimer.

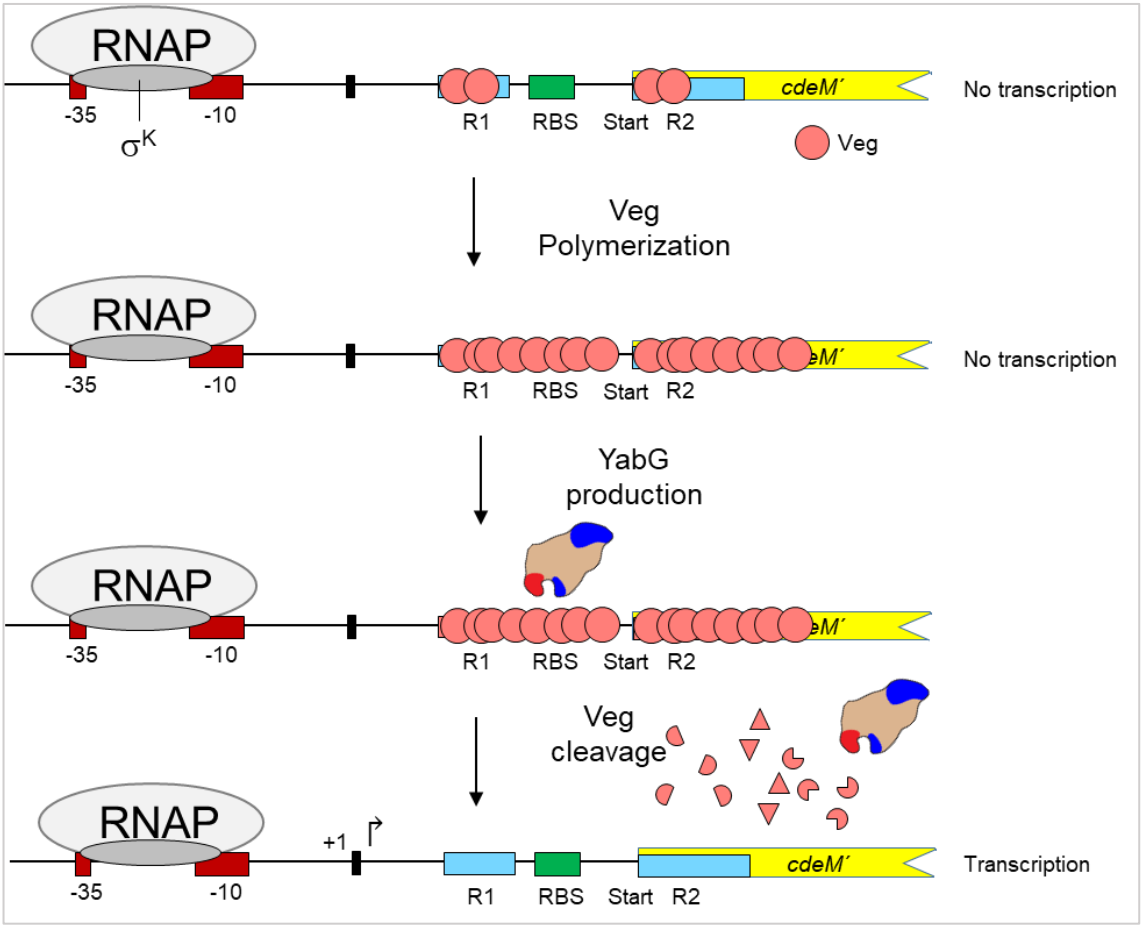

Figure 4.1 – Interaction model of Veg and YabG with the *cdeM* locus. Veg is thought to interact with R1 and R2 (see figure 3.2) which are AT rich. After binding to these regions, Veg is thought to be recruited to form a long oligomer along *cdeM locus*. In late sporulation (see figure 3.2), after Veg cleavage by YabG, *cdeM* transcription would occur.

Although we cannot exclude that Veg binds to SinR regulating gene expression by directly antagonizing the function of the latter, in the light of the *in vitro* results with the *cdeM* promoter we prefer a model in which Veg does not exerts its function through interaction with other proteins but rather by interaction to DNA, leading to changes on gene expression that could impact on growth, sporulation and biofilm formation. The construction of an interference mutant (see above) would be pivotal to test the role of Veg as a repressor *in vivo*. In case of success RNAseq analysis of the mutant could reveal other genes under Veg control. To identify other DNA binding sites of Veg we could perform chromatin immunoprecipitation assays coupled with sequencing analysis (ClP-seq).

A way to regulate the NAP is proteolytic cleavage (Dole et al., 2004; Fang et al., 2020; Malabirade et al., 2018; Zhang et al., 2014). This irreversible mode of regulation is made possible by an abundance of lysine and arginine residues in these proteins, which are also involved in DNA binding (Dahlke & Sing, 2018; Wang et al., 2011). YabG was shown to cleave at arginine residues and during sporulation in the mother cell it is likely a regulator of Veg by proteolysis

(Olivença, C. 2021; figure 4.1). *C. difficile* Veg has 5 arginine residues, two at the borders of a disorder region (figure 1.5). Single substitutions of each arginine to alanines and/or to lysine (also positively charged) would allow to identify the residues important for DNA binding and for cleavage by YabG. Under different growth conditions, Veg may be regulated by other mechanisms; for instance, binding of other proteins to the DNA could change its topology leading to the release of Veg (Dorman, 1991; Fang et al., 2020; Malabirade et al., 2018).

5. Concluding remarks

Veg shows structural similarity to Sm-like proteins but did not show the typical behaviour of Sm proteins. It did not form an oligomer in the presence of the short, rich(U)-RNA, at least under the conditions tested. Veg formed a monomer and a dimer in solution, but has a set of different inherent characteristics, such as its isoelectric point (pI), which can explain why no higher order oligomers were detected.

We propose that Veg may produce a polymer-like oligomer, instead of a ring-shaped complex, and therefore may require a longer nucleic acid fragment to do so. Structural studies would permit to gain a deeper understanding into the mechanism of action of Veg. We show that Veg binds to the regulatory region of the *cdeM* gene, most likely at two direct AT-rich repeats, and this is, to our knowledge, the first evidence for a Sm-like protein binding DNA.

veg might be an essential gene that represses late-phase genes that may be lethal if expressed in early developmental stages. If this is the case Veg would be a potential target for new therapeutic treatments of CDI.

Table 1 – *C. difficile* strains used in this work.

Strain	Ribotype	Genotype; Phenotype	Reference
630Δ*erm*	12	Strain 630Δ*erm*	(Hussain et al., 2005)
1800	126	Human isolate. *tcdA* (+), *tcdB* (+), *tcdD* (-).	INSA
AHCD543	12	Strain 630Δ*erm* containing an empty pMTL-84121 for negative control; $Thia^R$	This work
AHCD772	12	Strain 630Δ*erm*Δ*pyrE*	(Heap et al., 2012)
AHCD1217	12	strain 630Δ*erm* containing a derived plasmid from pMTL84121 with the gene *gusA* under the control of the promotor P*Tcda(A)* (with *ccpa* box); $Thia^R$	This work
AHCD1218	12	Strain 630Δ*erm* containing a derived plasmid from pMTL84121 with the gene *gusA* under the control of the promotor P*fliC*; $Thia^R$	"
AHCD1234	12	Strain 630Δ*erm* containing a derived plasmid from pMTL84121 with the gene *gusA* under the control of the promotor *PTcda(B)* (without *ccpa* box); $Thia^R$	"
AHCD1235	12	Strain 630Δ*erm* containing a derived plasmid from pMTL84121 with the gene *gusA* under the control of the promotor P*sspa;* $Thia^R$	"
AHCD1648	12	Strain 630Δ*erm* containing the pLB2 plasmid for Veg overproduction; $Thia^R$	"
AHCD1649	126	Strain 1800 containing the pLB2 plasmid for Veg overproduction; $Thia^R$	"
AHCD1650	126	Strain 1800 containing an empty pMTL-84121 for negative control; $Thia^R$	"
AHCD1651	12	Strain 630Δ*erm* containing a derived plasmid from pMTL84121 with the gene *gusA* without a promoter region; $Thia^R$	"

Table 2 – *E. coli* strains used in this work.

Strain	Genotype; Phenotype	Reference
DH5α	F– Φ80lacZΔM15 Δ(*lac*ZYA*arg*F) U169 *rec*A1 *end*A1 *hsd*R17	Bethesda Research laboratories
BL21 (DE3)	F- *ompT gal dcm lon hsdSB*(*rB–mB–*) λ(DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*]) [*malB*+]K-12(λS)	Novagen
HB101	supE44 *aa*14 *galK*2 *lacY*1 D(*gpt-proA*) 62 *rpsL*20 *(Str*R*) xyl*-5 *mtl*-1 *recA*13 D*(mcrC-mrr) hsdSB*(rB-mB-) RP4	(el Meouche et al., 2013)
AHEC203	HB101 (pRP4) (pMTL-YN1)	(Ng et al., 2013)
AHEC204	HB101 (pRP4) (pMTL-YN1C)	"
AHEC205	HB101 (pRP4) (pMTL-YN3)	"
AHEC1100	BL21 (DE3) carrying pBG1, a plasmid derived from pET16b, carrying the *veg* gene (from *Clostridioides difficile*) with the His-tag, for overproduction of Veg; Amp^R	This work
AHEC1296	HB101 (pRP4) carrying pLB2, a plasmid derived from pMTL24121, containig the *veg* gene (from *Clostridioides difficile)* under the control of the *spo0A* promotor; Amp^R; Cm^R	"
AHEC1126	HB101 (pRP4) carrying pSR86, a plasmid derived from pYN1, containing a truncated version of the *veg* gene (from *Clostridioides difficile*); Amp^R; Cm^R	"
AHEC738	HB101 (pRP4) containing a derived plasmid from pMTL84121 with the gene *gusA* under control of the promotor *tcda(A)* (with *ccpa* box) ; Cm^R	"
AHEC739	HB101 (pRP4) containing a derived plasmid from pMTL84121 with the gene *gusA* under the control of the *tcda(B)* promotor (without *ccpa* box); Cm^R	"
AHEC741	HB101 (pRP4) containing a derived plasmid from pMTL84121 with the gene *gusA* under the control of the *sspa* promotor*;* Cm^R	"
AHEC1298	HB101 (pRP4) containing a derived plasmid from pMTL84121 with the gene *gusA* without a promoter region; Cm^R	"

Table 3 – Oligonucleotide sequence used in this work

Primer	**Oligonucleotide sequence (5' → 3')**
Veg600Fwd	CATGCCATGGCCGTGGCTACTGTTCAAACTC
Veg764Rv	CCGCTCGAGTTAGTGATGATGATGATGATGACTTATTTGTAATTTATCTTG
Veg-EcoRI-Fw	CCGGAATTCGATGTAGTAGTTCCTATTGATG
PcdeC-Veg-XhoI-Ver	GGCTCGAGTTAACTTATTTGTAATTTATCTTG
Veg-Veg-Fw	GTGCAGGAGCAAACTATGCAAG
Veg-Veg-Rv	GCTCCTACTTTAATCTTTCTC
YN3-vef-Fw	CATCAAGAAGAGCGACTTCG
YN3-vef-Rev	TTCTTTCTATTCAGCACTGTTATGC
PspoA-EcoRI	CGGAATTCAATGACAGCAATTTAATGGGTAATTTC
Pspo0A-Veg	GAGTTTGAACAGTAGCCATTAAAAAACATCTTCTTATTAC
SOE Pspo0A-Veg	GAGTTTGAACAGTAGCCATTAAAAAACATCTTCTTATTAC
Fw Veg	ATGGCTACTGTTCAAACTCTAGATAAG
Rw Veg XhoI	GGCTCGAGTTAACTTATTTGTAATTTATCTTG
PCdeM-CY3-Fw	GCTGATTCTGAATATGAAATAAAC
PCdeM-CY3-Rv	CTTTCATATTCTTCTCTATCATTTTGGA
PCdeM-Reg-EcoRI-Fw	GGGGAATTCGAGATGAAACAAAACATCTC
PCdeM-Reg-HindIII-Rv	CCCAAGCTTGAGATGTTTTGTTTCATCTC
u10	uuuuuuuuuu
T10	TTTTTTTTTT

Note – Restriction sites are underlined

Table 4 – Plasmids used in this work

Plasmid	Relevant properties	Reference
pBG1	Plasmid derived from pET16b carrying the *veg* gene (from *Clostridioides difficile*) with the His6-tag, for overproduction of Veg-6-His-Tag N-terminal ; Amp^R	(Gonçalves, 2020)
pET16b	Protein expression vector; Amp^R	Novagen
pLB2	Plasmid for protein overproduction derived from pMTL84121 carrying the *veg* gene (from *Clostridioides difficile*) under the control of the *spo0A* promotor. Cm^R;$Thia^R$.	This work
pMTL84121	*Clostridium* modular plasmid containing *catP* Cm^R;$Thia^R$	(Heap et al., 2009)
pSR86	Plasmid derived from pMTL-YN3 for Allele Coupled Exchange for *veg* truncation containing the *pyrE* allele; Cm^R;$Thia^R$	(Ramalhete, 2021)
pMS35	Plasmid derived from pMTL84121 containing the gene *gusA* under the control of no promotor for negative control; Cm^R;$Thia^R$.	This work
pMS537	Plasmid derived from pMTL84121 for over expression of the transcriptional fusion of *gusA* under the control of the *tcdA* with the *ccpa* box promotor; Cm^R;$Thia^R$.	“
pMS538	Plasmid derived from pMTL84121 for over expression of the transcriptional fusion of *gusA* under the control of the *tcdA* without the *ccpa* box promotor; Cm^R;$Thia^R$.	“
pMS539	Plasmid derived from pMTL84121 for over expression of the transcriptional fusion of *gusA* under the control of the *sspa* promotor; Cm^R;$Thia^R$.	“
pMS540	Plasmid derived from pMTL84121 for over expression of the transcriptional fusion of *gusA* under the control of the *fliC* promotor; Cm^R;$Thia^R$.	“

Table 5 – Culture media used in this work.

Media	Composition (for 100 mL)
Luria-Bertani Broth	1 g tryptone, 0.5 g yeast extract, 0.05 g NaCl.
Auto-induction medium	1 mM MgSO4, 1 X 5052, 1 X NPS, LB medium 92.9% (v/v) (Studier, 2005).
BHI	3.7 g BHI.
BHISG-DOC	3.7 g BHI; 300 µL L-cystein hydrochloride (1.27 M); 10 mL of D-glucose (1 M); 240 µL sodium deoxycholate (100 mM).
BHISG-DOC-RUG	BHISG-DOC with Thiamphenicol 15 µg/µL and Rug 10 µg/µL.
CDMM	10 mg/mL casamino acids, 0.5 mg/mL L-tryptophan, 0.5 mg/mL L-cystein, 5 mg/mL Na_2HPO_4, 5 mg/mL $NaHCO_3$, 0.9 mg/mL KH_2PO_4, 0.9 mg/mL NaCl, 10 mg/mL D-Glucose, 0.4 mg/mL $(NH_4)_2SO_4$, 0.026 mg/mL $CaCl_2$•2H2O, 0.02 mg/mL $MgCl_2$•$6H_2O$, 0.01 mg/mL $MnCl_2$•$4H_2O$, 0.001 mg/mL $CoCl_2$•$6H_2O$, 0.004 mg/mL $FeSO_4$•$7H_2O$, 0.001 mg/mL D-biotin, 0.001 mg/mL Calcium-D-panthothenate, 0.001 mg/mL Pyridoxine.
70:30	6.3 g bactopeptone, 0.35 g proteose-peptone, 1.11 g BHI, 0.15 g yeast extract, 0.07 g $(NH_4)_2SO_4$, 0.11 g tris-base, 0.1 g of L-cystein hydrochloride.

Table 6 – Solutions used in this work

Solution	Composition
Comassie solution	0.5 g/mL Coomassie Brilliant Blue R-250; 80 % absolute ethanol; 20% acetic acid.
Destaining solution	30% absolute ethanol; 10% acetic acid.
HEPES buffer	10 mM HEPES, 500 NaCl, pH 8.
Loading buffer 10x	50 mL Tris-HCl; 7 mL β-mercaptoethanol; 20 g SDS; 0.2 g bromophenol blue; remaining volume made up of glycerol 86%.
Orange G Loading buffer	0.2 % orange G, 0.37 % EDTA, 50 % glycerol.
PBS	137 mM NaCl; 2.7 mM KCl; 4.3 mM Na_2HPO_4; 1.4 mM KH_2PO_4.
PBS-Tween	PBS 1x with 0,1% Tween-20.
RF1 Buffer	12 mg/mL RbCl, 9.9 mg/mL MnCl2, 1.5 mg/mL CaCl2, 11 % glycerol, 3 % KAc 1 M pH 7.46.
RF2 Buffer	1.2 mg/mL RbCl, 8.3 mg/mL CaCl2, 10 % glycerol, 2 %, MOPS 0.5 M, pH 6.8.
SDS-PAGE Resolving gel 15%	35% distilled water;25% 4x lower Tris buffer; 37.67% bis-acrylamide; 0.1% SDS; 0.1% APS; 0.005% TEMED.
SDS-PAGE Stacking gel	61.2 % distilled water; 25.5 % 4x upper Tris buffer; 10.2 % bis-acrylamide; 0.1 % SDS; 0.1 % APS; 0.1 % TEMED.
Transfer buffer	14.4 g/L glycine; 3.02 g/L Tris base; 10 % EtOH.
TAE-Buffer	40 mM Tris, 20 mM acetic acid, 1 mM EDTA.
Start Buffer	20 mM Phosphate, 500 mM NaCl, 20 mM Imidazole pH 7.4.
STET buffer	8 % sucrose, 0.5 % Triton X-100 (v/v), 50 mM of EDTA, 10 mM of Tris HCl, pH 8.0.
EMSA buffer	10 mM Tris-HCl pH 7.4, 50 mM NaCl, 5 mM EDTA pH 8, 1 mM DTT, 50 % glycerol, 50 µg/ mL Bovine serum albumin.
Elution buffer	20 mM Phosphate, 500 mM NaCl, 500 mM Imidazole pH 7.4.
CHES buffer	10 mM CHES, 150 mM NaCl, pH 9.5.
CAPS buffer	10 mM CAPS, 150 mM NaCl, pH 11.5.
Tris-HCl buffer	10 mM Tris-HCl, 150 mM NaCl, pH 7.5.
Quenching buffer	2 M Tris-HCl, 150 mM NaCl, pH 8.

6. References

Abecasis, A. B., Serrano, M., Alves, R., Quintais, L., Pereira-Leal, J. B., & Henriques, A. O. (2013). A genomic signature and the identification of new sporulation genes. *Journal of Bacteriology*, *195*(9), 2101–2115. https://doi.org/10.1128/JB.02110-12/SUPPL_FILE/ZJB999092534SO1.PDF

Achsel, T., Brahms, H., Kastner, B., Bachi, A., Wilm, M., & Lührmann, R. (1999). A doughnut-shaped heteromer of human Sm-like proteins binds to the 3'-end of U6 snRNA, thereby facilitating U4/U6 duplex formation in vitro. *The EMBO Journal*, *18*(20), 5789–5802. https://doi.org/10.1093/emboj/18.20.5789

Achsel, T., Stark, H., & Lührmann, R. (2001a). The Sm domain is an ancient RNA-binding motif with oligo(U) specificity. *Proceedings of the National Academy of Sciences*, *98*(7), 3685–3689. https://doi.org/10.1073/PNAS.071033998

Achsel, T., Stark, H., & Lührmann, R. (2001b). The Sm domain is an ancient RNA-binding motif with oligo(U) specificity. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(7), 3685. https://doi.org/10.1073/PNAS.071033998

Åkerlund, T., Persson, I., Unemo, M., Norén, T., Svenungsson, B., Wullt, M., & Burman, L. G. (2008). Increased sporulation rate of epidemic Clostridium difficile type 027/NAP1. *Journal of Clinical Microbiology*, *46*(4), 1530–1533. https://doi.org/10.1128/JCM.01964-07/SUPPL_FILE/FIGS1_FIGS2_TABLES1_V2.ZIP

Alvarez-Ordóñez, A., Coughlan, L. M., Briandet, R., & Cotter, P. D. (2019). Biofilms in Food Processing Environments: Challenges and Opportunities. *Https://Doi.Org/10.1146/Annurev-Food-032818-121805*, *10*, 173–195. https://doi.org/10.1146/ANNUREV-FOOD-032818-121805

Antunes, A., Camiade, E., Monot, M., Courtois, E., Barbut, F., Sernova, N. v., Rodionov, D. A., Martin-Verstraete, I., & Dupuy, B. (2012a). Global transcriptional control by glucose and carbon regulator CcpA in Clostridium difficile. *Nucleic Acids Research*, *40*(21), 10701–10718. https://doi.org/10.1093/NAR/GKS864

Antunes, A., Camiade, E., Monot, M., Courtois, E., Barbut, F., Sernova, N. v., Rodionov, D. A., Martin-Verstraete, I., & Dupuy, B. (2012b). Global transcriptional control by glucose and carbon regulator CcpA in Clostridium difficile. *Nucleic Acids Research*, *40*(21), 10701–10718. https://doi.org/10.1093/NAR/GKS864

Antunes, A., Martin-Verstraete, I., & Dupuy, B. (2011). CcpA-mediated repression of Clostridium difficile toxin gene expression. *Molecular Microbiology*, *79*(4). https://doi.org/10.1111/j.1365-2958.2010.07495.x

Arnaouteli, S., Bamford, N. C., Stanley-Wall, N. R., & Kovács, Á. T. (2021a). Bacillus subtilis biofilm formation and social interactions. *Nature Reviews Microbiology 2021 19:9*, *19*(9), 600–614. https://doi.org/10.1038/s41579-021-00540-9

Arnaouteli, S., Bamford, N. C., Stanley-Wall, N. R., & Kovács, Á. T. (2021b). Bacillus subtilis biofilm formation and social interactions. In *Nature Reviews Microbiology* (Vol. 19, Issue 9, pp. 600–614). Nature Research. https://doi.org/10.1038/s41579-021-00540-9

Awad, M. M., Johanesen, P. A., Carter, G. P., Rose, E., & Lyras, D. (2014). Clostridium difficile virulence factors: Insights into an anaerobic spore-forming pathogen. *Gut Microbes*, *5*(5), 579. https://doi.org/10.4161/19490976.2014.969632

Babakhani, F., Bouillaut, L., Gomez, A., Sears, P., Nguyen, L., & Sonenshein, A. L. (2012). Fidaxomicin inhibits spore production in Clostridium difficile. *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America*, *55 Suppl 2*(Suppl 2). https://doi.org/10.1093/CID/CIS453

Bäckhed, F., Fraser, C. M., Ringel, Y., Sanders, M. E., Sartor, R. B., Sherman, P. M., Versalovic, J., Young, V., & Finlay, B. B. (2012). Defining a Healthy Human Gut Microbiome: Current Concepts, Future Directions, and Clinical Applications. *Cell Host & Microbe*, *12*(5), 611–622. https://doi.org/10.1016/J.CHOM.2012.10.012

Bai, U., Mandic-Mulec, I., & Smith, I. (1993). SinI modulates the activity of SinR, a developmental switch protein of Bacillus subtilis, by protein-protein interaction. *Genes & Development*, *7*(1), 139–148. https://doi.org/10.1101/GAD.7.1.139

Bandyra, K. J., Bouvier, M., Carpousis, A. J., & Luisi, B. F. (2013). The social fabric of the RNA degradosome. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, *1829*(6–7), 514–522. https://doi.org/10.1016/J.BBAGRM.2013.02.011

Begley, M., Gahan, C. G. M., & Hill, C. (2005). The interaction between bacteria and bile. *FEMS Microbiology Reviews*, *29*(4). https://doi.org/10.1016/j.femsre.2004.09.003

Bhattacharjee, D., & Sorg, J. A. (2020). Factors and Conditions That Impact Electroporation of Clostridioides difficile Strains. *MSphere*, *5*(2). https://doi.org/10.1128/MSPHERE.00941-19/ASSET/690CDDBC-23BC-45E6-A41A-44DAF9AE2F75/ASSETS/GRAPHIC/MSPHERE.00941-19-F0008.JPEG

Blaut, M., Collins, M. D., Welling, G. W., Doré, J., van Loo, J., & de Vos, W. (2002). Molecular biological methods for studying the gut microbiota: the EU human gut flora project. *British Journal of Nutrition*, *87*(S2), S203–S211. https://doi.org/10.1079/bjn/2002539

Bonneville, L., Maia, V., Barroso, I., Martínez-Suárez, J. v., & Brito, L. (2021). Lactobacillus plantarum in Dual-Species Biofilms With Listeria monocytogenes Enhanced the Anti-Listeria Activity of a Commercial Disinfectant Based on Hydrogen Peroxide and Peracetic Acid. *Frontiers in Microbiology*, *12*, 631627. https://doi.org/10.3389/FMICB.2021.631627/FULL

Boudry, P., Gracia, C., Monot, M., Caillet, J., Saujet, L., Hajnsdorf, E., Dupuy, B., Martin-Verstraete, I., & Soutourina, O. (2014). Pleiotropic role of the RNA chaperone protein Hfq in the human pathogen Clostridium difficile. *Journal of Bacteriology*, *196*(18), 3234–3248. https://doi.org/10.1128/JB.01923-14/ASSET/4D27959C-6BFB-485F-A10E-815DE11D8851/ASSETS/GRAPHIC/ZJB9990932880004.JPEG

Boudry, P., Piattelli, E., Drouineau, E., Peltier, J., Boutserin, A., Lejars, M., Hajnsdorf, E., Monot, M., Dupuy, B., Martin-Verstraete, I., Gautheret, D., Toffano-Nioche, C., & Soutourina, O. (2021). Identification of RNAs bound by Hfq reveals widespread RNA partners and a sporulation regulator in the human pathogen Clostridioides difficile. *RNA Biology*, *18*(11), 1931–1952. https://doi.org/10.1080/15476286.2021.1882180

Branda, S. S., Vik, Å., Friedman, L., & Kolter, R. (2005). Biofilms: the matrix revisited. *Trends in Microbiology*, *13*(1), 20–26. https://doi.org/10.1016/J.TIM.2004.11.006

Brown, K. A., Khanafer, N., Daneman, N., & Fisman, D. N. (2013). Meta-analysis of antibiotics and the risk of community-associated Clostridium difficile infection. *Antimicrobial Agents and Chemotherapy*, *57*(5), 2326–2332. https://doi.org/10.1128/AAC.02176-12

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*(6). https://doi.org/10.1016/j.anaerobe.2010.10.001

Calabi, E., Calabi, F., Phillips, A. D., & Fairweather, N. F. (2002). Binding of Clostridium difficile surface layer proteins to gastrointestinal tissues. *Infection and Immunity*, *70*(10), 5770–5778. https://doi.org/10.1128/IAI.70.10.5770-5778.2002

Carpentier, B., & Cerf, O. (1993). Biofilms and their consequences, with particular reference to hygiene in the food industry. *Journal of Applied Bacteriology*, *75*(6), 499–511. https://doi.org/10.1111/J.1365-2672.1993.TB01587.X

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*(1), 58–64. https://doi.org/10.4161/GMIC.1.1.10768

Carter, G. P., Rood, J. I., & Lyras, D. (2012). The role of toxin A and toxin B in the virulence of Clostridium difficile. *Trends in Microbiology*, *20*(1). https://doi.org/10.1016/j.tim.2011.11.003

Chai, Y., Chu, F., Kolter, R., & Losick, R. (2008). Bistability and biofilm formation in Bacillus subtilis. *Molecular Microbiology*, *67*(2), 254–263. https://doi.org/10.1111/J.1365-2958.2007.06040.X

Chastanet, A., Vitkup, D., Yuan, G. C., Norman, T. M., Liu, J. S., & Losick, R. M. (2010). Broadly heterogeneous activation of the master regulator for sporulation in Bacillus subtilis. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(18), 8486–8491. https://doi.org/10.1073/PNAS.1002499107/-/DCSUPPLEMENTAL/STXT01.DOC

Chen, B., Wen, J., Zhao, X., Ding, J., & Qi, G. (2020). Surfactin: A Quorum-Sensing Signal Molecule to Relieve CCR in Bacillus amyloliquefaciens. *Frontiers in Microbiology*, *11*, 631. https://doi.org/10.3389/FMICB.2020.00631/BIBTEX

Costerton, J. W., Lewandowski, Z., Caldwell, D. E., Korber, D. R., & Lappin-Scott, H. M. (1995). Microbial biofilms. *Annual Review of Microbiology*, *49*, 711–745. https://doi.org/10.1146/ANNUREV.MI.49.100195.003431

Cowardin, C. A., Buonomo, E. L., Saleh, M. M., Wilson, M. G., Burgess, S. L., Kuehne, S. A., Schwan, C., Eichhoff, A. M., Koch-Nolte, F., Lyras, D., Aktories, K., Minton, N. P., & Petri, W. A. (2016). The binary toxin CDT enhances Clostridium difficile virulence by suppressing protective colonic eosinophilia. *Nature Microbiology*, *1*(8). https://doi.org/10.1038/NMICROBIOL.2016.108

Cummings, J. H. (2009). Probiotics: better health from 'good' bacteria? *Nutrition Bulletin*, *34*(2), 198–202. https://doi.org/10.1111/J.1467-3010.2009.01746.X

Curtis, R. A., & Lue, L. (2006). A molecular approach to bioseparations: Protein–protein and protein–salt interactions. *Chemical Engineering Science*, *61*(3), 907–923. https://doi.org/10.1016/J.CES.2005.04.007

Dahlke, K., & Sing, C. E. (2018). Force-extension behavior of DNA in the presence of DNA-bending nucleoid associated proteins. *The Journal of Chemical Physics*, *148*(8), 84902. https://doi.org/10.1063/1.5016177

Daou, N., Wang, Y., Levdikov, V. M., Nandakumar, M., Livny, J., Bouillaut, L., Blagova, E., Zhang, K., Belitsky, B. R., Rhee, K., Wilkinson, A. J., Sun, X., & Sonenshein, A. L. (2019). Impact of CodY protein on metabolism, sporulation and virulence in Clostridioides difficile ribotype 027. *PLOS ONE*, *14*(1), e0206896. https://doi.org/10.1371/JOURNAL.PONE.0206896

Dapa, T., Leuzzi, R., Ng, Y. K., Baban, S. T., Adamo, R., Kuehne, S. A., Scarselli, M., Minton, N. P., Serruto, D., & Unnikrishnan, M. (2013). Multiple factors modulate biofilm formation by the anaerobic pathogen Clostridium difficile. *Journal of Bacteriology*, *195*(3), 545–555. https://doi.org/10.1128/JB.01980-12

Dawson, L. F., Peltier, J., Hall, C. L., Harrison, M. A., Derakhshan, M., Shaw, H. A., Fairweather, N. F., & Wren, B. W. (2021). Extracellular DNA, cell surface proteins and c-di-GMP promote biofilm formation in Clostridioides difficile. *Scientific Reports*, *11*(1). https://doi.org/10.1038/S41598-020-78437-5

Dawson, L. F., Valiente, E., Faulds-Pain, A., Donahue, E. H., Wren, B. W., & Popoff, M. R. (2012). Characterisation of Clostridium difficile Biofilm Formation, a Role for Spo0A. *PLoS ONE*, *7*(12). https://doi.org/10.1371/journal.pone.0050527

de Roo, A. C., Regenbogen, S. E., de Roo, A. C., & Regenbogen, S. E. (2020). Clostridium difficile Infection: An Epidemiology Update. *Clinics in Colon and Rectal Surgery*, *33*(2), 49–57. https://doi.org/10.1055/s-0040-1701229

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. (2012a). The *Clostridium difficile spo0A* Gene Is a Persistence and Transmission Factor. *Infection and Immunity*, *80*(8). https://doi.org/10.1128/IAI.00147-12

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. (2012b). The Clostridium difficile spo0A gene is a persistence and transmission factor. *Infection and Immunity*, *80*(8), 2704–2711. https://doi.org/10.1128/IAI.00147-12/SUPPL_FILE/ZII999099744SO2.PDF

Deller, M. C., Kong, L., & Rupp, B. (2016). Protein stability: A crystallographer's perspective. In *Acta Crystallographica Section:F Structural Biology Communications* (Vol. 72, pp. 72–95). International Union of Crystallography. https://doi.org/10.1107/S2053230X15024619

Dembek, M., Barquist, L., Boinett, C. J., Cain, A. K., Mayho, M., Lawley, T. D., Fairweather, N. F., & Fagan, R. P. (2015). High-throughput analysis of gene essentiality and sporulation in Clostridium difficile. *MBio*, *6*(2). https://doi.org/10.1128/MBIO.02383-14

Dineen, S. S., Villapakkam, A. C., Nordman, J. T., & Sonenshein, A. L. (2007). Repression of Clostridium difficile toxin gene expression by CodY. *Molecular Microbiology*, *66*(1). https://doi.org/10.1111/j.1365-2958.2007.05906.x

Dole, S., Nagarajavel, V., & Schnetz, K. (2004). The histone-like nucleoid structuring protein H-NS represses the Escherichia coli bgl operon downstream of the promoter. *Molecular Microbiology*, *52*(2), 589–600. https://doi.org/10.1111/J.1365-2958.2004.04001.X

Domínguez-Manzano, J., Olmo-Ruiz, C., Bautista-Gallego, J., Arroyo-López, F. N., Garrido-Fernández, A., & Jiménez-Díaz, R. (2012). Biofilm formation on abiotic and biotic surfaces during Spanish style green table olive fermentation. *International Journal of Food Microbiology*, *157*(2), 230–238. https://doi.org/10.1016/J.IJFOODMICRO.2012.05.011

Donelli, G., Vuotto, C., Cardines, R., & Mastrantonio, P. (2012). Biofilm-growing intestinal anaerobic bacteria. *FEMS Immunology and Medical Microbiology*, *65*(2), 318–325. https://doi.org/10.1111/J.1574-695X.2012.00962.X

Donlan, R. M. (2002). Biofilms: microbial life on surfaces. *Emerging Infectious Diseases*, *8*(9), 881–890. https://doi.org/10.3201/EID0809.020063

Dorman, C. J. (1991). MINIREVIEW DNA Supercoiling and Environmental Regulation of Gene Expression in Pathogenic Bacteria. *INFECTION AND IMMUNITY*, *59*(3), 745–749. https://journals.asm.org/journal/iai

Driks, A., & Eichenberger, P. (2016a). The Spore Coat. *Microbiology Spectrum*, *4*(2). https://doi.org/10.1128/MICROBIOLSPEC.TBS-0023-2016

Driks, A., & Eichenberger, P. (2016b). The Spore Coat. *Microbiology Spectrum*, *4*(2). https://doi.org/10.1128/MICROBIOLSPEC.TBS-0023-2016

Dubois, T., Tremblay, Y. D. N., Hamiot, A., Martin-Verstraete, I., Deschamps, J., Monot, M., Briandet, R., & Dupuy, B. (2019a). A microbiota-generated bile salt induces biofilm formation in Clostridium difficile. *Npj Biofilms and Microbiomes 2019 5:1*, *5*(1), 1–12. https://doi.org/10.1038/s41522-019-0087-4

Dubois, T., Tremblay, Y. D. N., Hamiot, A., Martin-Verstraete, I., Deschamps, J., Monot, M., Briandet, R., & Dupuy, B. (2019b). A microbiota-generated bile salt induces biofilm formation in Clostridium difficile. *Npj Biofilms and Microbiomes 2019 5:1*, *5*(1), 1–12. https://doi.org/10.1038/s41522-019-0087-4

Dupuy, B., & Sonenshein, A. L. (1998). Regulated transcription of *Clostridium difficile* toxin genes. *Molecular Microbiology*, *27*(1). https://doi.org/10.1046/j.1365-2958.1998.00663.x

Eckburg, P. B., Bik, E. M., Bernstein, C. N., Purdom, E., Dethlefsen, L., Sargent, M., Gill, S. R., Nelson, K. E., & Relman, D. A. (2005). Diversity of the human intestinal microbial flora. *Science (New York, N.Y.)*, *308*(5728), 1635–1638. https://doi.org/10.1126/SCIENCE.1110591

el Meouche, I., Peltier, J., Monot, M., Soutourina, O., Pestel-Caron, M., Dupuy, B., & Pons, J. L. (2013). Characterization of the SigD Regulon of C. difficile and Its Positive Control of Toxin Production through the Regulation of tcdR. *PLoS ONE*, *8*(12). https://doi.org/10.1371/JOURNAL.PONE.0083748

Fadouloglou, V. E., Kokkinidis, M., & Glykos, N. M. (2008). Determination of protein oligomerization state: Two approaches based on glutaraldehyde crosslinking. *Analytical Biochemistry*, *373*(2), 404–406. https://doi.org/10.1016/j.ab.2007.10.027

Fang, Y., Akimoto, M., Mayanagi, K., Hatano, A., Matsumoto, M., Matsuda, S., Yasukawa, T., & Kang, D. (2020). Chemical acetylation of mitochondrial transcription factor A occurs on specific lysine residues and affects its ability to change global DNA topology. *Mitochondrion*, *53*, 99–108. https://doi.org/10.1016/J.MITO.2020.05.003

Fimlaid, K. A., Jensen, O., Donnelly, M. L., Siegrist, M. S., & Shen, A. (2015). Regulation of Clostridium difficile Spore Formation by the SpoIIQ and SpoIIIA Proteins. *PLOS Genetics*, *11*(10), e1005562. https://doi.org/10.1371/JOURNAL.PGEN.1005562

Fimlaid, K. A., & Shen, A. (2015). Diverse mechanisms regulate sporulation sigma factor activity in the Firmicutes. *Current Opinion in Microbiology*, *24*, 88–95. https://doi.org/10.1016/J.MIB.2015.01.006

Flemming, H. C., & Wuertz, S. (2019). Bacteria and archaea on Earth and their abundance in biofilms. *Nature Reviews Microbiology*, *17*(4), 247–260. https://doi.org/10.1038/s41579-019-0158-9

Frostid, L. R., Chengid, J. K. J., & Unnikrishnanid, M. (2021). Clostridioides difficile biofilms: A mechanism of persistence in the gut? *PLOS Pathogens*, *17*(3), e1009348. https://doi.org/10.1371/JOURNAL.PPAT.1009348

Fukushima, T., Ishikawa, S., Yamamoto, H., Ogasawara, N., & Sekiguchi, J. (2003). Transcriptional, Functional and Cytochemical Analyses of the veg Gene in Bacillus subtilis. *The Journal of Biochemistry*, *133*(4), 475–483. https://doi.org/10.1093/JB/MVG062

Gamier, T., & Cole, S. T. (1988). Studies of UV-inducible promoters from Clostridium perfringens in vivo and in vitro. *Molecular Microbiology*, *2*(5), 607–614. https://doi.org/10.1111/J.1365-2958.1988.TB00069.X

Girinathan, B. P., Ou, J., Dupuy, B., & Govind, R. (2018a). *Pleiotropic roles of Clostridium difficile sin locus*. https://doi.org/10.1371/journal.ppat.1006940

Girinathan, B. P., Ou, J., Dupuy, B., & Govind, R. (2018b). Pleiotropic roles of Clostridium difficile sin locus. *PLOS Pathogens*, *14*(3), e1006940. https://doi.org/10.1371/JOURNAL.PPAT.1006940

Gominet, M., Compain, F., Beloin, C., & Lebeaux, D. (2017). Central venous catheters and biofilms: where do we stand in 2017? *APMIS*, *125*(4), 365–375. https://doi.org/10.1111/APM.12665

Gonçalves, B. A. A. (2020). *Studies on the relationship between assembly of the spore coat and the biofilm*. https://repositorio.ul.pt/handle/10451/48586

Govind, R., & Dupuy, B. (2012). Secretion of Clostridium difficile Toxins A and B Requires the Holin-like Protein TcdE. *PLoS Pathogens*, *8*(6). https://doi.org/10.1371/journal.ppat.1002727

Gupta, A., & Khanna, S. (2014). Community-acquired Clostridium difficile infection: an increasing public health threat. *Infection and Drug Resistance*, *7*, 63. https://doi.org/10.2147/IDR.S46780

Hammond, G. A., & Johnson, J. L. (1995). The toxigenic element of Clostridium difficile strain VPI 10463. *Microbial Pathogenesis*, *19*(4). https://doi.org/10.1016/S0882-4010(95)90263-5

Heap, J. T., Pennington, O. J., Cartman, S. T., & Minton, N. P. (2009). A modular system for Clostridium shuttle plasmids. *Journal of Microbiological Methods*, *78*(1), 79–85. https://doi.org/10.1016/J.MIMET.2009.05.004

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*(10), 2779–2787. https://doi.org/10.1099/mic.0.26145-0

Hennequin, C., Porcheray, F., Waligora-Dupriet, A. J., Collignon, A., Barc, M. C., Bourlioux, P., & Karjalainen, T. (2001). GroEL (Hsp60) of Clostridium difficile is involved in cell adherence. *Microbiology (Reading, England)*, *147*(Pt 1), 87–96. https://doi.org/10.1099/00221287-147-1-87

Henriques, A. O., & Moran, C. P. (2007). Structure, assembly, and function of the spore surface layers. *Annual Review of Microbiology*, *61*, 555–588. https://doi.org/10.1146/ANNUREV.MICRO.61.080706.093224

Hundsberger, T., Braun, V., Weidmann, M., Leukel, P., Sauerborn, M., & Eichel-Streiber, C. (1997). Transcription Analysis of the Genes tcdA-E of the Pathogenicity Locus of Clostridium Difficile. *European Journal of Biochemistry*, *244*(3). https://doi.org/10.1111/j.1432-1033.1997.t01-1-00735.x

Hunt, J. J., & Ballard, J. D. (2013). Variations in Virulence and Molecular Biology among Emerging Strains of Clostridium difficile. *Microbiology and Molecular Biology Reviews*, *77*(4). https://doi.org/10.1128/MMBR.00017-13

Huq, A., Whitehouse, C. A., Grim, C. J., Alam, M., & Colwell, R. R. (2008). Biofilms in water, its role and impact in human disease transmission. *Current Opinion in Biotechnology*, *19*(3), 244–247. https://doi.org/10.1016/J.COPBIO.2008.04.005

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. *Journal of Medical Microbiology*, *54*(Pt 2), 137–141. https://doi.org/10.1099/JMM.0.45790-0

Ibragimova, G. T., & Wade, R. C. (1998). Importance of Explicit Salt Ions for Protein Stability in Molecular Dynamics Simulation. *Biophysical Journal*, *74*(6), 2906–2911. https://doi.org/10.1016/S0006-3495(98)77997-4

Isidro, J., Mendes, A. L., Serrano, M., Henriques, A. O., & Oleastro, M. (2017). Overview of Clostridium difficile Infection: Life Cycle, Epidemiology, Antimicrobial Resistance and Treatment. In *Clostridium Difficile - A Comprehensive Overview*. InTech. https://doi.org/10.5772/intechopen.69053

Janoir, C. (2016). Virulence factors of Clostridium difficile and their role during infection. *Anaerobe*, *37*. https://doi.org/10.1016/j.anaerobe.2015.10.009

Jian, H., Xu, G., Gai, Y., Xu, J., & Xiao, X. (2016). *The Histone-Like Nucleoid Structuring Protein (H-NS) Is a Negative Regulator of the Lateral Flagellar System in the Deep-Sea Bacterium Shewanella piezotolerans WP3*. https://doi.org/10.1128/AEM.00297-16

Jiang, M., Shao, W., Perego, M., & Hoch, J. A. (2000). Multiple histidine kinases regulate entry into stationary phase and sporulation in Bacillus subtilis. *Molecular Microbiology*, *38*(3), 535–542. https://doi.org/10.1046/J.1365-2958.2000.02148.X

Kambach, C., Walke, S., Young, R., Avis, J. M., de La Fortelle, E., Raker, V. A., Lührmann, R., Li, J., & Nagai, K. (1999). Crystal structures of two Sm protein complexes and their implications for the assembly of the spliceosomal snRNPs. *Cell*, *96*(3), 375–387. https://doi.org/10.1016/S0092-8674(00)80550-4

Karasawa, T., Ikoma, S., Yamakawa, K., & Nakamura, S. (1995). A defined growth medium for Clostridium difficile. *Microbiology (Reading, England)*, *141 (Pt 2)*(2), 371–375. https://doi.org/10.1099/13500872-141-2-371

Karlsson, S., Lindberg, A., Norin, E., Burman, L. G., & Åkerlund, T. (2000). Toxins, Butyric Acid, and Other Short-Chain Fatty Acids Are Coordinately Expressed and Down-Regulated by Cysteine in *Clostridium difficile*. *Infection and Immunity*, *68*(10). https://doi.org/10.1128/IAI.68.10.5881-5888.2000

Kearns, D. B., Chu, F., Branda, S. S., Kolter, R., & Losick, R. (2005). A master regulator for biofilm formation by Bacillus subtilis. *Molecular Microbiology*, *55*(3), 739–749. https://doi.org/10.1111/J.1365-2958.2004.04440.X

Kevorkian, Y., Shirley, D. J., & Shen, A. (2016). Regulation of Clostridium difficile spore germination by the CspA pseudoprotease domain. *Biochimie*, *122*, 243–254. https://doi.org/10.1016/J.BIOCHI.2015.07.023

Kilic, T., Sanglier, S., Dorsselaer, A. van, & Suck, D. (2006). Oligomerization behavior of the archaeal Sm2-type protein from Archaeoglobus fulgidus. *Protein Science : A Publication of the Protein Society*, *15*(10), 2310. https://doi.org/10.1110/PS.062191506

Kobayashi, K., & Iwano, M. (2012). BslA(YuaB) forms a hydrophobic layer on the surface of Bacillus subtilis biofilms. *Molecular Microbiology*, *85*(1), 51–66. https://doi.org/10.1111/J.1365-2958.2012.08094.X

Kociolek, L. K., & Gerding, D. N. (2016). Breakthroughs in the treatment and prevention of Clostridium difficile infection. *Nature Reviews Gastroenterology & Hepatology*, *13*(3). https://doi.org/10.1038/nrgastro.2015.220

Kolenbrander, P. E., Palmer, R. J., Periasamy, S., & Jakubovics, N. S. (2010). Oral multispecies biofilm development and the key role of cell–cell distance. *Nature Reviews Microbiology 2010 8:7*, *8*(7), 471–480. https://doi.org/10.1038/nrmicro2381

le Grice, S. F. J., Shih, C. C., Whipple, F., & Sonenshein, A. L. (1986). Separation and analysis of the RNA polymerase binding sites of a complex Bacillus subtilis promoter. *Molecular and General Genetics MGG 1986 204:2*, *204*(2), 229–236. https://doi.org/10.1007/BF00425503

Lei, Y., Oshima, T., Ogasawara, N., & Ishikawa, S. (2013a). Functional analysis of the protein Veg, which stimulates biofilm formation in Bacillus subtilis. *Journal of Bacteriology*, *195*(8), 1697–1705. https://doi.org/10.1128/JB.02201-12

Lei, Y., Oshima, T., Ogasawara, N., & Ishikawa, S. (2013b). Functional analysis of the protein veg, which stimulates biofilm formation in bacillus subtilis. *Journal of Bacteriology*, *195*(8), 1697–1705. https://doi.org/10.1128/JB.02201-12

Li, X., Romero, P., Rani, M., Keith Dunker, A., & Obradovic, Z. (n.d.). *Predicting Protein Disorder for N-,C-and Internal Regions*.

Linder, J. A., Huang, E. S., Steinman, M. A., Gonzales, R., & Stafford, R. S. (2005). Fluoroquinolone prescribing in the United States: 1995 to 2002. *The American Journal of Medicine*, *118*(3). https://doi.org/10.1016/j.amjmed.2004.09.015

Liu, B., Eydallin, G., Maharjan, R. P., Feng, L., Wang, L., & Ferenci, T. (2017). Natural Escherichia coli isolates rapidly acquire genetic changes upon laboratory domestication. *Microbiology (Reading, England)*, *163*(1), 22–30. https://doi.org/10.1099/MIC.0.000405

Louro, M. C. (2020). *Characterization of biofilm formation in clinical isolates of Clostridioides difficile*. https://repositorio.ul.pt/handle/10451/47816

Macpherson, A. J. (2006). IgA Adaptation to the Presence of Commensal Bacteria in the Intestine. *Current Topics in Microbiology and Immunology*, *308*, 117–136. https://doi.org/10.1007/3-540-30657-9_5

Malabirade, A., Partouche, D., el Hamoui, O., Turbant, F., Geinguenaud, F., Recouvreux, P., Bizien, T., Busi, F., Wien, F., & Arluison, V. (2018). Revised role for Hfq bacterial regulator on DNA topology. *Scientific Reports 2018 8:1*, *8*(1), 1–12. https://doi.org/10.1038/s41598-018-35060-9

Mandic-Mulec, I., Gaur, N., Bai, U., & Smith, I. (1992). Sin, a stage-specific repressor of cellular differentiation. *Journal of Bacteriology*, *174*(11), 3561–3569. https://doi.org/10.1128/JB.174.11.3561-3569.1992

Mani, N., & Dupuy, B. (2001a). Regulation of toxin synthesis in Clostridium difficile by an alternative RNA polymerase sigma factor. *Proceedings of the National Academy of Sciences*, *98*(10). https://doi.org/10.1073/pnas.101126598

Mani, N., & Dupuy, B. (2001b). Regulation of toxin synthesis in Clostridium difficile by an alternative RNA polymerase sigma factor. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(10), 5844–5849. https://doi.org/10.1073/PNAS.101126598

Marini, E. (2020). *Assembly of the Clostridioides difficile spore surface layers*. https://run.unl.pt/handle/10362/105568

Marsh, J. W., Arora, R., Schlackman, J. L., Shutt, K. A., Curry, S. R., & Harrison, L. H. (2012). Association of relapse of Clostridium difficile disease with BI/NAP1/027. *Journal of Clinical Microbiology*, *50*(12), 4078–4082. https://doi.org/10.1128/JCM.02291-12

Matamouros, S., England, P., & Dupuy, B. (2007). Clostridium difficile toxin expression is inhibited by the novel regulator TcdC. *Molecular Microbiology*, *64*(5). https://doi.org/10.1111/j.1365-2958.2007.05739.x

McDonald, L. C., Gerding, D. N., Johnson, S., Bakken, J. S., Carroll, K. C., Coffin, S. E., Dubberke, E. R., Garey, K. W., Gould, C. v, Kelly, C., Loo, V., Shaklee Sammons, J., Sandora, T. J., & Wilcox, M. H. (2018). Clinical Practice Guidelines for Clostridium difficile Infection in Adults and Children: 2017 Update by the Infectious Diseases Society of America (IDSA) and Society for Healthcare Epidemiology of America (SHEA). *Clinical Infectious Diseases*, *66*(7). https://doi.org/10.1093/cid/cix1085

Meza-Torres, J., Auria, E., Dupuy, B., & Tremblay, Y. D. N. (2021). Wolf in Sheep's Clothing: Clostridioides difficile Biofilm as a Reservoir for Recurrent Infections. *Microorganisms*, *9*(9). https://doi.org/10.3390/microorganisms9091922

Mielich-Süss, B., & Lopez, D. (2015a). Molecular mechanisms involved in Bacillus subtilis biofilm formation. *Environmental Microbiology*, *17*(3), 555–565. https://doi.org/10.1111/1462-2920.12527

Mielich-Süss, B., & Lopez, D. (2015b). Molecular mechanisms involved in Bacillus subtilis biofilm formation. *Environmental Microbiology*, *17*(3), 555. https://doi.org/10.1111/1462-2920.12527

Mowat, A. M. I. (2003). Anatomical basis of tolerance and immunity to intestinal antigens. *Nature Reviews Immunology 2003 3:4*, *3*(4), 331–341. https://doi.org/10.1038/nri1057

Müh, U., Pannullo, A. G., Weiss, D. S., & Ellermeier, C. D. (2019). A Xylose-Inducible Expression System and a CRISPR Interference Plasmid for Targeted Knockdown of Gene Expression in Clostridioides difficile. *Journal of Bacteriology*, *201*(14). https://doi.org/10.1128/JB.00711-18

Mura, C., Randolph, P. S., Patterson, J., & Cozen, A. E. (2013). Archaeal and eukaryotic homologs of Hfq. *Http://Dx.Doi.Org/10.4161/Rna.24538*, *10*(4), 636–651. https://doi.org/10.4161/RNA.24538

Neumann-Schaal, M., Jahn, D., & Schmidt-Hohagen, K. (2019). Metabolism the difficile way: The key to the success of the pathogen Clostridioides difficile. *Frontiers in Microbiology*, *10*(FEB), 219. https://doi.org/10.3389/FMICB.2019.00219/BIBTEX

Newman, J. A., Rodrigues, C., & Lewis, R. J. (2013). Molecular basis of the activity of SinR protein, the master regulator of biofilm formation in Bacillus subtilis. *The Journal of Biological Chemistry*, *288*(15), 10766–10778. https://doi.org/10.1074/JBC.M113.455592

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*(2), e56051. https://doi.org/10.1371/JOURNAL.PONE.0056051

Oezguen, N., Power, T. D., Urvil, P., Feng, H., Pothoulakis, C., Stamler, J. S., Braun, W., & Savidge, T. C. (2012). Clostridial toxins. *Gut Microbes*, *3*(1). https://doi.org/10.4161/gmic.19250

Olling, A., Seehase, S., Minton, N. P., Tatge, H., Schröter, S., Kohlscheen, S., Pich, A., Just, I., & Gerhard, R. (2012). Release of TcdA and TcdB from Clostridium difficile cdi 630 is not affected by functional inactivation of the tcdE gene. *Microbial Pathogenesis*, *52*(1). https://doi.org/10.1016/j.micpath.2011.10.009

Ollington, J. F., Haldenwang, W. G., Huynh, T. v., & Losick, R. (1981). Developmentally regulated transcription in a cloned segment of the Bacillus subtilis chromosome. *Journal of Bacteriology*, *147*(2), 432–442. https://doi.org/10.1128/JB.147.2.432-442.1981

Pandit, A., Adholeya, A., Cahill, D., Brau, L., & Kochar, M. (2020). Microbial biofilms in nature: unlocking their potential for agricultural applications. *Journal of Applied Microbiology*, *129*(2), 199–211. https://doi.org/10.1111/JAM.14609

Paredes-Sabja, D., Shen, A., & Sorg, J. A. (2014a). Clostridium difficile spore biology: sporulation, germination, and spore structural proteins. *Trends in Microbiology*, *22*(7), 406–416. https://doi.org/10.1016/J.TIM.2014.04.003

Paredes-Sabja, D., Shen, A., & Sorg, J. A. (2014b). Clostridium difficile spore biology: Sporulation, germination, and spore structural proteins. *Trends in Microbiology*, *22*(7), 406–416. https://doi.org/10.1016/J.TIM.2014.04.003

Pereira, F. C., Saujet, L., Tomé, 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 Genetics*, *9*(10), e1003782. https://doi.org/10.1371/JOURNAL.PGEN.1003782

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. F., &

Cutting, S. M. (2014). The spore-associated protein BclA1 affects the susceptibility of animals to colonization and infection by Clostridium difficile. *Molecular Microbiology*, *92*(5), 1025–1038. https://doi.org/10.1111/MMI.12611/SUPPINFO

Pizarro-Guajardo, M., Calderón-Romero, P., & Paredes-Sabja, D. (2016). Ultrastructure Variability of the Exosporium Layer of Clostridium difficile Spores from Sporulating Cultures and Biofilms. *Applied and Environmental Microbiology*, *82*(19), 5892–5898. https://doi.org/10.1128/AEM.01463-16

Pizarro-Guajardo, M., Calderón-Romero, P., Romero-Rodríguez, A., & Paredes-Sabja, D. (2020). Characterization of Exosporium Layer Variability of Clostridioides difficile Spores in the Epidemically Relevant Strain R20291. *Frontiers in Microbiology*, *11*, 1345. https://doi.org/10.3389/FMICB.2020.01345/BIBTEX

Purcell, E. B., McKee, R. W., Bordeleau, E., Burrus, V., & Tamayo, R. (2016). Regulation of Type IV Pili contributes to surface behaviors of historical and epidemic strains of Clostridium difficile. *Journal of Bacteriology*, *198*(3), 565–577. https://doi.org/10.1128/JB.00816-15

Putnam, E. E., Nock, A. M., Lawley, T. D., & Shen, A. (2013). SpoIVA and SipL are Clostridium difficile spore morphogenetic proteins. *Journal of Bacteriology*, *195*(6), 1214–1225. https://doi.org/10.1128/JB.02181-12

Rabi, R., Turnbull, L., Whitchurch, C. B., Awad, M., & Lyras, D. (2017). Structural Characterization of Clostridium sordellii Spores of Diverse Human, Animal, and Environmental Origin and Comparison to Clostridium difficile Spores. *MSphere*, *2*(5). https://doi.org/10.1128/MSPHERE.00343-17

Ramalhete, S. (2021). *Insights into the production of toxins and the assembly of spores in Clostridioides difficile*. https://run.unl.pt/handle/10362/123468

Ramos-Silva, P., Serrano, M., & Henriques, A. O. (2019). From root to tips: sporulation evolution and specialization inBacillus subtilisand the intestinal pathogenClostridioides difficile. *BioRxiv*. https://doi.org/10.1101/473793

Ribis, J. W., Fimlaid, K. A., & Shen, A. (2018). Differential requirements for conserved peptidoglycan remodeling enzymes during Clostridioides difficile spore formation. *Molecular Microbiology*, *110*(3), 370–389. https://doi.org/10.1111/MMI.14090

Rodriguez-Palacios, A., & LeJeune, J. T. (2011). Moist-heat resistance, spore aging, and superdormancy in Clostridium difficile. *Applied and Environmental Microbiology*, *77*(9), 3085–3091. https://doi.org/10.1128/AEM.01589-10/SUPPL_FILE/AEM01589V2_SUPERDORMANCY_SUPLEMENT_SUB_FEB.DOC

Rupnik, M. (2007). Is Clostridium difficile-associated infection a potentially zoonotic and foodborne disease? *Clinical Microbiology and Infection*, *13*(5), 457–459. https://doi.org/10.1111/J.1469-0691.2007.01687.X

Rupnik, M., Wilcox, M. H., & Gerding, D. N. (2009). Clostridium difficile infection: new developments in epidemiology and pathogenesis. *Nature Reviews Microbiology*, *7*(7). https://doi.org/10.1038/nrmicro2164

Santos, S. P., Bandeiras, T. M., Pinto, A. F., Teixeira, M., Carrondo, M. A., & Romão, C. v. (2012). Thermofluor-based optimization strategy for the stabilization and crystallization of Campylobacter jejuni desulforubrerythrin. *PROTEIN EXPRESSION AND PURIFICATION*, *81*, 193–200. https://doi.org/10.1016/j.pep.2011.10.001

Sára, M., & Sleytr, U. B. (2000). S-Layer Proteins. *Journal of Bacteriology*, *182*(4), 859. https://doi.org/10.1128/JB.182.4.859-868.2000

Sarker, M. R., & Paredes-Sabja, D. (2012). Molecular basis of early stages of *Clostridium difficile* infection: germination and colonization. *Future Microbiology*, *7*(8). https://doi.org/10.2217/fmb.12.64

Sauer, E., Schmidt, S., Weichenrieder, O., & Doudna, J. A. (n.d.). *Small RNA binding to the lateral surface of Hfq hexamers and structural rearrangements upon mRNA target recognition*. https://doi.org/10.1073/pnas.1202521109

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. *Journal of Bacteriology*, *193*(13), 3186–3196. https://doi.org/10.1128/JB.00272-11

Sebaihia, M., Wren, B. W., Mullany, P., Fairweather, N. F., Minton, N., Stabler, R., Thomson, N. R., Roberts, A. P., Cerdeño-Tárraga, A. M., Wang, H., Holden, M. T. G., Wright, A., Churcher, C., Quail, M. A., Baker, S., Bason, N., Brooks, K., Chillingworth, T., Cronin, A., … Parkhill, J. (2006). The multidrug-resistant human pathogen Clostridium difficile has a highly mobile, mosaic genome. *Nature Genetics*, *38*(7), 779–786. https://doi.org/10.1038/NG1830

Semenyuk, E. G., Poroyko, V. A., Johnston, P. F., Jones, S. E., Knight, K. L., Gerding, D. N., & Driks, A. (2015). Analysis of bacterial communities during Clostridium difficile infection in the mouse. *Infection and Immunity*, *83*(11), 4383–4391. https://doi.org/10.1128/IAI.00145-15

Serrano, M., Kint, N., Pereira, F. C., Saujet, L., Boudry, P., Dupuy, B., Henriques, A. O., & Martin-Verstraete, I. (2016). A Recombination Directionality Factor Controls the Cell Type-Specific Activation of σK and the Fidelity of Spore Development in Clostridium difficile. *PLoS Genetics*, *12*(9). https://doi.org/10.1371/JOURNAL.PGEN.1006312

Shrestha, R., Cochran, A. M., & Sorg, J. A. (2019). The requirement for co-germinants during Clostridium difficile spore germination is influenced by mutations in yabG and cspA. *PLoS Pathogens*, *15*(4). https://doi.org/10.1371/JOURNAL.PPAT.1007681

Slimings, C., & Riley, T. v. (2014). Antibiotics and hospital-acquired Clostridium difficile infection: update of systematic review and meta-analysis. *The Journal of Antimicrobial Chemotherapy*, *69*(4), 881–891. https://doi.org/10.1093/JAC/DKT477

Sorg, J. A., & Sonenshein, A. L. (2008). Bile salts and glycine as cogerminants for Clostridium difficile spores. *Journal of Bacteriology*, *190*(7). https://doi.org/10.1128/JB.01765-07

Stoodley, P., Sauer, K., Davies, D. G., & Costerton, J. W. (2002). Biofilms as complex differentiated communities. *Annual Review of Microbiology*, *56*, 187–209. https://doi.org/10.1146/annurev.micro.56.012302.160705

Tan, C. H., Lee, K. W. K., Burmølle, M., Kjelleberg, S., & Rice, S. A. (2017). All together now: experimental multispecies biofilm model systems. *Environmental Microbiology*, *19*(1), 42–53. https://doi.org/10.1111/1462-2920.13594

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. *Infection and Immunity*, *69*(12), 7937–7940. https://doi.org/10.1128/IAI.69.12.7937-7940.2001/ASSET/80E19FC1-EBB4-4086-A8ED-8F425339190B/ASSETS/GRAPHIC/II1210521002.JPEG

Theriot, C. M., Koenigsknecht, M. J., Carlson, P. E., Hatton, G. E., Nelson, A. M., Li, B., Huffnagle, G. B., Z. Li, J., & Young, V. B. (2014). Antibiotic-induced shifts in the mouse gut microbiome and metabolome increase susceptibility to Clostridium difficile infection. *Nature Communications*, *5*(1). https://doi.org/10.1038/ncomms4114

Törö, I., Basquin, J., Teo-Dreher, H., & Suck, D. (2002). Archaeal Sm Proteins form Heptameric and Hexameric Complexes: Crystal Structures of the Sm1 and Sm2 Proteins from the Hyperthermophile Archaeoglobus fulgidus. *Journal of Molecular Biology*, *320*(1), 129–142. https://doi.org/10.1016/S0022-2836(02)00406-0

Tremblay, Y. D. N., Durand, B. A. R., Hamiot, A., Martin-Verstraete, I., Oberkampf, M., Monot, M., & Dupuy, B. (2021). Metabolic adaption to extracellular pyruvate triggers biofilm formation in Clostridioides difficile. *The ISME Journal 2021*, 1–13. https://doi.org/10.1038/s41396-021-01042-5

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. *Journal of Bacteriology*, *191*(23), 7296. https://doi.org/10.1128/JB.00882-09

Urlaub, H., Raker, V. A., Kostka, S., & Lührmann, R. (2001). Sm protein–Sm site RNA interactions within the inner ring of the spliceosomal snRNP core structure. *The EMBO Journal*, *20*(1–2), 187–196. https://doi.org/10.1093/EMBOJ/20.1.187

Vishwakarma, V. (2020). Impact of environmental biofilms: Industrial components and its remediation. *Journal of Basic Microbiology*, *60*(3), 198–206. https://doi.org/10.1002/JOBM.201900569

Wilson, G. M., Sun, Y., Lu, H., & Brewer, G. (1999). *Assembly of AUF1 Oligomers on U-rich RNA Targets by Sequential Dimer Association**. http://www.jbc.org

Wilson, K. H., & Perini, F. (1988). Role of competition for nutrients in suppression of Clostridium difficile by the colonic microflora. *Infection and Immunity*, *56*(10), 2610–2614. https://doi.org/10.1128/IAI.56.10.2610-2614.1988

Wörner, K., Szurmant, H., Chiang, C., & Hoch, J. A. (2006). Phosphorylation and functional analysis of the sporulation initiation factor Spo0A from Clostridium botulinum. *Molecular Microbiology*, *59*(3), 1000–1012. https://doi.org/10.1111/J.1365-2958.2005.04988.X

Wright, A., Drudy, D., Kyne, L., Brown, K., & Fairweather, N. F. (2008). Immunoreactive cell wall proteins of Clostridium difficile identified by human sera. *Journal of Medical Microbiology*, *57*(6), 750–756. https://doi.org/10.1099/JMM.0.47532-0/CITE/REFWORKS

Zhang, X., Zhou, H., Xie, Y., Ren, C., Ding, D., Long, J., & Yang, Z. (2014). Rational Design of Multifunctional Hetero-Hexameric Proteins for Hydrogel Formation and Controlled Delivery of Bioactive Molecules. *Advanced Healthcare Materials*, *3*(11), 1804–1811. https://doi.org/10.1002/ADHM.201300660

Zhao, X., Hameed, U. F. S., Kharchenko, V., Liao, C., Huser, F., Remington, J. M., Radhakrishnan, A. K., Jaremko, M., Jaremko, Ł., Arold, S. T., & Li, J. (2021). Molecular basis for the adaptive evolution of environment-sensing by H-NS proteins. *ELife*, *10*, 1–18. https://doi.org/10.7554/ELIFE.57467

Zhao, Y., Wu, J., Li, J. v., Zhou, N.-Y., Tang, H., & Wang, Y. (2013). Gut Microbiota Composition Modifies Fecal Metabolic Profiles in Mice. *Journal of Proteome Research*, *12*(6). https://doi.org/10.1021/pr400263n

Zhu, D., Sorg, J. A., & Sun, X. (2018). Clostridioides difficile biology: Sporulation, germination, and corresponding therapies for C. difficile infection. *Frontiers in Cellular and Infection Microbiology*, *8*(FEB), 29. https://doi.org/10.3389/FCIMB.2018.00029/BIBTEX