CHARLES UNIVERSITY
FACULTY OF PHARMACY IN HRADEC KRÁLOVÉ
DEPARTMENT OF BIOCHEMICAL SCIENCES

Kateřina Houdková Mpharm

Analysis of flagellar proteins in *Clostridium difficile* isolates of clinically relevant PCR-ribotypes

Rigorous thesis

Supervisor: RNDr. Klára Konečná, Ph.D.

Consultant: PharmDr. Jiří Dresler, Ph.D.

Hradec Králové 2022

I hereby declare that this thesis is my original work which I solely composed by myself under the supervision of RNDr. Klára Konečná, Ph.D. All used literature and other sources are summarized in the list of references and properly cited. This work has not been submitted for any different or equal degree.

The results presented in this thesis were published as:
Dresler, J. *et al.*, Analysis of proteomes released from *in vitro* cultured eight *Clostridium difficile* PCR ribotypes revealed specific expression in PCR ribotypes 027 and 176 confirming their genetic relatedness and clinical importance at the proteomic level. Gut Pathog. 9, 45 (2017).

Prohlašuji, že tato práce je mým původním autorským dílem, které jsem vypracovala samostatně pod vedením své školitelky RNDr. Kláry Konečné, Ph.D.
Veškerá literatura a další zdroje, z nichž jsem při zpracování čerpala, jsou uvedeny v seznamu použité literatury a v práci jsou řádně citovány. Práce nebyla využita k získání jiného nebo stejného titulu.

Výsledky prezentované v této práci byly publikovány:
Dresler, J. *et al.,* Analysis of proteomes released from *in vitro* cultured eight *Clostridium difficile* PCR ribotypes revealed specific expression in PCR ribotypes 027 and 176 confirming their genetic relatedness and clinical importance at the proteomic level. Gut Pathog. 9, 45 (2017).

.....
Kateřina Houdková Mpharm

ACKNOWLEDGEMENTS

I would like to thank my supervisor RNDr. Klára Konečná, Ph.D. for her help and guidance over the course of the writing of the thesis. Also, my thanks go to PharmDr. Jiří Dresler, Ph.D. for his continued support, advice and leadership throughout the project.

ABSTRACT

Charles University
Faculty of Pharmacy in Hradec Králové
Department of Biochemical Sciences

Candidate: Kateřina Houdková MPharm
Supervisor: RNDr. Klára Konečná, Ph.D.
Consultant: PharmDr. Jiří Dresler, Ph.D.
Title of rigorous thesis: Analysis of flagellar proteins in *C. difficile* isolates of clinically relevant PCR-ribotypes

Background: Strains of *C. difficile* of known human epidemiologic importance are associated with severe clinical features of *C. difficile* infection (CDI). In this study, a panel of eight different PCR-ribotypes (RTs) with their proteins released *in vitro* were subjected to analysis. The aim of this work is to monitor the relationship between secretions of individual proteins associated with flagellar formation and function in *C. difficile* strains of variable virulence.

Methods: Within our research, a combination of tandem mass spectrometry with liquid chromatography was used. The semi-quantitative analysis employed label free quantification (LFQ) approach.

Results: From the quantifiable proteins, 17 were significantly increased in functional annotations. Among them, several known factors connected with flagellar assembly and other functions were identified. Higher expression of selected flagellar proteins clearly distinguished RTs 027, 176, 005 and 012, confirming the pathogenic role of the assembly in CDI.

Conclusion: The outcome of this work was different observations of individual flagellar proteins in various strains differentiated by increased potential for virulence.

Keywords: *Clostridium difficile*, label-free quantification, virulence factors, toxins A/B, flagellins.

ABSTRAKT

Univerzita Karlova
Farmaceutická fakulta v Hradci Králové
Katedra biochemických věd

Kandidátka: Kateřina Houdková MPharm
Školitelka: RNDr. Klára Konečná, Ph.D.
Konzultant: PharmDr. Jiří Dresler, Ph.D.
Název rigorozní práce: Analýza flagelárních proteinů klinicky reprezentativních ribotypů bakterie *C. difficile*

Cíl práce: Kmeny *C. difficile* známé svou epidemiologickou důležitostí pro člověka jsou spojeny se závažnými klinickými příznaky infekce *C. difficile* (CDI). V této studii byly předmětem analýzy proteiny uvolněné z *in vitro* kultivovaného panelu osmi různých PCR-ribotypů (RT). Cílem této práce je sledovat vztah mezi sekrecí jednotlivých proteinů spojených s výstavbou a funkcí bičíků u kmenů *C. difficile* s variabilní virulencí.

Metoda: V rámci našeho výzkumu byla preferovanou formou zkoumání kombinace shotgun proteomiky a label free quantification (LFQ).

Výsledky: Ze zkoumaných proteinů bylo 17 významně zvýšeno ve funkčních anotacích. Mezi nimi bylo identifikováno několik známých faktorů souvisejících s virulencí, jako jsou proteiny spojené s výstavbou bičíků a dalších funkcí. Vyšší sekrece vybraných bičíkových proteinů jasně odlišila RT 027, 176, 005 a 012, což potvrdilo jejich patogenní roli v CDI.

Závěr: Výsledkem této práce byla různá pozorování u různých kmenů se zvýšeným potenciálem virulence.

Klíčová slova: *Clostridium difficile*, “label-free” kvantifikace, faktory virulence, toxiny A/B, flageliny.

Contents

1 Introduction

Clostridium difficile, currently newly designated as *Clostridioides difficile (C. difficile),* is the most common cause of infective diarrhoea mainly in hospitalised patients and is associated with substantial morbidity and mortality. *C. difficile* associated diarrhoea (CDAD) is responsible for around 10-20% of all cases of antibiotic associated diarrhoea (AAD) [1] and it can occur up to 8 weeks after antibiotic therapy. The incidence depends on a range of factors including the type of antibiotic used. Since the early 2000s, the burden of *C. difficile* infection (CDI) has increased in many European countries, with the annual incidence in Europe estimated at 124,000 cases in 2011–2012 with all-cause mortality rates of 3–30% [2].

CDI also increases treatment costs and length of stay in the acute healthcare setting. This has an overall economic burden of initial and recurrent CDI [3].

Although no particular antibiotics can be ruled out, those most commonly implicated in CDI are cephalosporins (especially second and third generation), quinolones, beta-lactams (namely co-amoxiclav) and aminopenicillins such as ampicillin and amoxicillin, which may be related to their volume of use [4]. Broad-spectrum antibiotics, compared with narrow-spectrum antibiotics, are more likely to significantly change the intestinal flora potentially allowing *C. difficile* and other enteric bacteria to take over [5].

The *C. difficile* strain known as 027, by polymerase chain reaction (PCR) ribotyping and generally referred to as BI/NAP1/027, is currently the single most important epidemic strain causing CDAD in North America and Europe [6]. The emergence of this hypervirulent strain has increased interest in *C. difficile* typing and drives the application of newer genotype-based methods such as PCR-ribotyping (a typing method based on the heterogeneity of ribosomal intergenic spacer region, that has been routinely used to investigate CDI outbreaks), amplified fragment length polymorphism (AFLP), multilocus sequence typing (MLST), multilocus variable-number tandem-repeat analysis (MLVA) and surface layer protein A gene sequence typing (*slpA*ST) [7]. Some methods such as PCR have evidenced major qualities such as easiness, rapidity and reproducibility. The discriminatory power of the aforementioned diverse genotyping methods has to be evaluated to validate its importance as a typing tool for *C. difficile* [8].

Previous studies that focused on the proteomic analysis of *C. difficile* until now were based on investigation of either the whole cell lysates [8, 9]; for example, a study by the team of authors Wright *et al.* (2005) [10] has looked at cell surface proteins, whereas others such as Lawley *et al*. (2009) focused the attention on spore proteins [11] and the insoluble proteome (Jain *et al.* 2010) [9] of the 630 reference strain; or culture supernatants representing secretome *in vitro* [12]. Further comparative proteomic analysis reveals varying protein levels between individual PCR ribotypes, which points to novel mechanisms underlining pathogenicity [13].

The rapid progress in developing an understanding of *C. difficile* virulence factors in the last decade, coupled with the development of new and diverse animal models [14], has facilitated the development of new strategies to combat this bacterium. Although antibiotic therapy is currently still the best treatment option for *C. difficile*, it is likely that studies dissecting the role of spore and vegetative cell factors, as well as toxins, in establishing infection and causing disease will allow prevention and treatment strategies to move beyond antibiotics [15].

C. difficile has been studied in order to understand its virulence and also to try to determine what accounts for the differing virulence between strains [16]. The results of our comparative semi quantitative analysis could be instrumental in the future development of targeted therapy and diagnostic tools.

2 THEORETICAL PART

2.1 *Clostridium difficile*

The recent proposal to restrict the genus *Clostridium* to *Clostridium butyricum* and related species has ramifications for the members of the genera that fall outside this clade that should not be considered as *Clostridium sensu strict* [17]. This includes *Clostridium difficile* and therefore it has been recently reclassified as *Clostridioides difficile*.

Clostridium difficile (C. difficile) is a Gram-positive, anaerobic, spore forming and toxin producing bacterium. Although it is part of the normal intestinal microflora, in 1–3% of healthy adults and 15–20% of infants, it was recognized as an important cause of antibiotic associated diarrhoea (AAD) already in the 1970s [18]. *C. difficile* is spread in the form of spores via the oral-faecal route and infection occurs through infected individuals, nosocomial sources and contaminated environment or health care workers. It is acquired by oral ingestion of spores which are resistant to physical changes and can survive for a long time under the high temperatures and UV light in the environment as well as being tolerant of the acidity of the stomach. During the colonization of the host, *C. difficile* spores are shed by the patient and facilitate the transmission of *C. difficile* further to susceptible hosts [19].

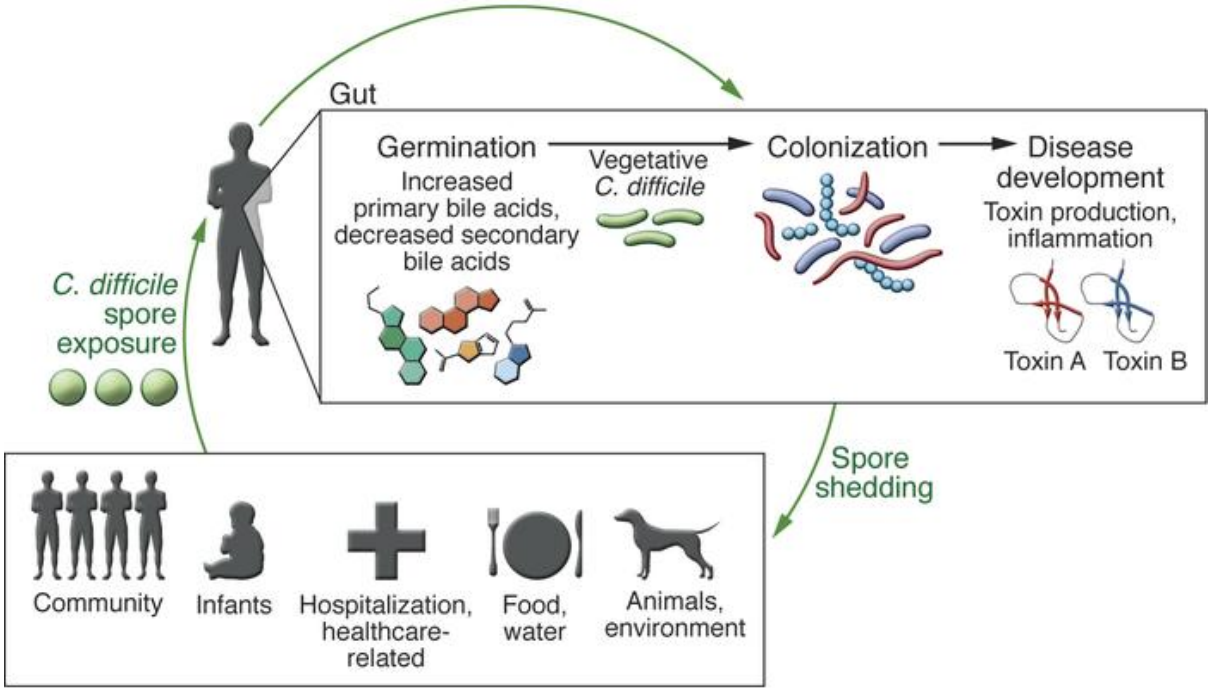

Fig 1. The life cycle of *Clostridium difficile*. The development of the disease is dependent on different stages of the *C. difficile* life cycle. An infected patient is the potential future source of the spores. Scheme taken from Seekatz A.M and Young V.B. (2014) [19].

2.2 *Clostridium difficile* pathogenesis

The normal location for *C. difficile* in humans is the intestinal tract, however, it can be isolated from water, vegetables, long term care environment, and also the intestines of domesticated animals [20]. In a healthy human, when spores are ingested, they are passed through the duodenum and the germination takes place in the jejunum because of the high concentration of the bile (bile salt germinants). The germinated spores pass through the ileum and reach the aerobic environment of the cecum, where the normal microbiome metabolises cholate derivatives to deoxycholate which prevents the vegetative growth of spores. A little germination is of no consequence in the aerobic environment of the cecum [21]–[23].

C. difficile can productively colonize the descending colon of individuals whose normal intestinal flora has been disrupted (e.g. by antibiotic treatment). Colonization likely depends upon adherence of the bacterium to the epithelium, although little is known about the factors that mediate adherence. Toxin producing strains may increase their adherence to intestinal epithelial cells by inducing microtubule protrusions that trap the bacterium. Glycosylating toxin-producing strains stimulate inflammation of the colonic lining by inducing cytoskeletal

changes that compromise the epithelial barrier and inflammatory cytokine production. The pathogenesis of *C. difficile* from ingestion to pseudomembrane formation is depicted in Fig. 2.

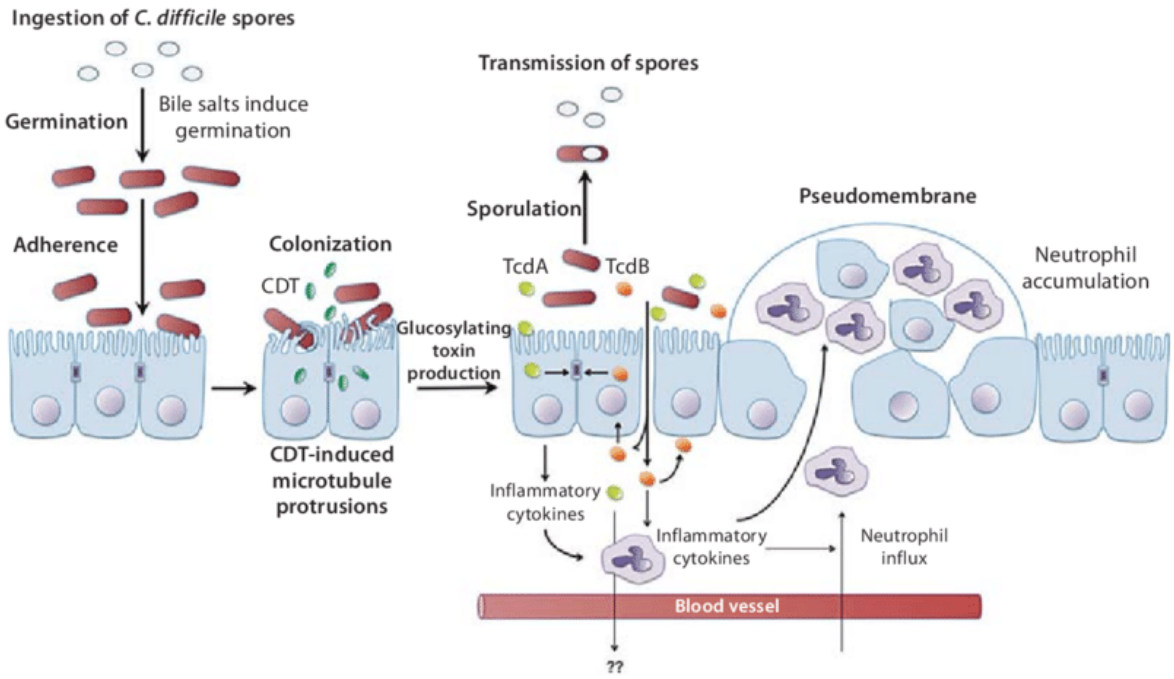

Fig 2. Pathogenesis of *Clostridium difficile*. CDT – Clostridium difficile transferase, TcdA – an enterotoxin, TcdB – a cytotoxin, all produced by *C. difficile*. Adapted from Shen A. *et al* (2012) [24].

Disruption of tight junctions allows the toxins to cross the epithelium, where they can further induce inflammatory cytokine production in lymphocytes and mast cells. This leads to escalation of the inflammatory response due to neutrophil and lymphocyte influx, which can lead to pseudomembrane formation. Whether glycosylating toxins enter the bloodstream remains unclear [16], [24], [25].

2.3 Therapy of *Clostridium* difficile infection (CDI)

An assessment of the severity of the infection is a prerequisite for an effective treatment. Treatment should only be started in patients with CDI symptoms; the presence of the *C. difficile* toxin without symptoms of the infection is not an indication for treatment. The severity of *C. difficile* infection according to Updated guidance on the management and treatment of *C. difficile* infection, Public Health England, 2013, is defined as:

a, **Mild infection** which is not associated with an increased white cell count (WCC) but associated with fewer than 3 episodes of loose stools per day.

b, **Moderate** infection is associated with an increased WCC (but less than 15×10^9 per litre) and typically associated with 3 to 5 loose stools per day.

c, **Severe** infection where WCC is greater than 15×10^9 per litre, or an acutely increased serum creatinine concentration (greater than 50% increase above baseline), or a temperature higher than 38.5 °*C*, or evidence of severe colitis (abdominal or radiological signs). The number of stools may be a less reliable indicator of severity.

d, **Life-threatening** infection: symptoms and signs include hypotension, partial or complete ileus, toxic megacolon or computerised tomography evidence of severe disease.

In 2014, the European Society of Clinical Microbiology and Infectious Diseases (ESCMID) guidelines were published in which two antibiotic agents, metronidazole and vancomycin, were the standards of CDI treatment. Metronidazole was the first line drug in non-severe CDI, while vancomycin was the drug of choice for severe CDI [26]. Since then, the results of two identical, phase 3, multicenter, randomized, double dummy, double-blind, active-controlled, parallel-design efficacy studies showed the superiority of vancomycin relative to metronidazole. However, among patients with severe disease, a statistically significant relationship was not achieved [27].

In 2017, Infectious Diseases Society of America (IDSA) and Society for Healthcare Epidemiology of America (SHEA) updated their guidelines, pointing out that vancomycin and fidaxomicin are now the cornerstones of CDI treatment [26].

C. difficile colonization is far more frequent in the paediatric population than in adults and it is agreed that a positive test for *C. difficile* in children 2 years and under is often because of high carriage rates of the bacteria rather than because of actual infection. Symptoms are rarely reported before that age, even though asymptomatic colonization may represent a source of transmission of the bacillus to others [28]. The choice of therapy for CDI in children and young people under 18 years is based on the recommendation for CDI in adults, as detailed in the Fig. 3, while taking into account licensed indications for children and what products are available.

Treatment	Antibiotic, dosage and course length
First-line antibiotic for a first episode of mild, moderate or severe *C. difficile* infection	Vancomycin: 125 mg orally four times a day for 10 days
Second-line antibiotic for a first episode of mild, moderate or severe *C. difficile* infection if vancomycin is ineffective	Fidaxomicin: 200 mg orally twice a day for 10 days
Antibiotics for *C. difficile* infection if first- and second-line antibiotics are ineffective	Seek specialist advice. Specialists may initially offer: Vancomycin: Up to 500 mg orally four times a day for 10 days With our without Metronidazole: 500 mg intravenously three times a day for 10 days
Antibiotic for a further episode of *C. difficile* infection within 12 weeks of symptom resolution (relapse)	Fidaxomicin: 200 mg orally twice a day for 10 days
Antibiotics for a further episode of *C. difficile* infection more than 12 weeks after symptom resolution (recurrence)	Vancomycin: 125 mg orally four times a day for 10 days OR Fidaxomicin: 200 mg orally twice a day for 10 days
Antibiotics for life-threatening *C. difficile* infection	Seek urgent specialist advice, which may include surgery. Vancomycin: 500 mg orally four times a day for 10 days With Metronidazole: 500 mg intravenously three times a day for 10 days

Fig. 3 Choice of antibiotics in *C. difficile* infection in adult patients aged 18 years and older. Adapted from NICE Guideline (2021) [29].

The role of probiotics in the treatment and prevention of CDI is completely unknown. Although there have been shown some positive effects of probiotics in the way of preventing primary CDI, there is no strong evidence agreeing upon the type or dosing of probiotics, the timing of the therapy and even though it appears that probiotics may be a part of CDI prevention or treatment, there are insufficient data at this time to recommend administration of probiotics for primary, secondary prevention or treatment of CDI due to a lack of properly randomized studies addressing this [26]–[28], [30].

As other modalities such as faecal transplants and a monoclonal antibody, bezlotoxomab, become increasingly part of treatment regimens [31]–[33], the complexities of treating CDI, especially in an outpatient setting may only increase.

Various *C. difficile* strains are already resistant to many antibiotics, particularly in the case of quinolones and the emergent ribotype 027 strains [34] it is of increasing importance to understand the detailed adaptation mechanisms of this pathogen to antibiotic treatment. In

order to do this in a comprehensive manner, it is beneficial to detect not only changes in protein amounts but also in protein synthesis [35], [36].

2.4 Virulence factors

C. difficile produces a number of effector molecules that contribute to its virulence. Some are true aggressins, such as toxins that directly contribute to the pathology associated with infection, while others enable *C. difficile* to colonize and produce these aggressins within the human host. As with all pathogens, not all strains are equally virulent.

2.4.1 Exotoxins, toxin A and toxin B

Only vegetative cells are able to produce the major virulence factors such as the exotoxins, toxin A (TcdA, an enterotoxin) and toxin B (TcdB, a cytotoxin). Extensive studies have demonstrated that these two toxins, TcdA and TcdB, are responsible for severe tissue damage and consequent manifestation of disease [37], [38]. Clinical strains that do not produce at least one of these toxins are largely avirulent in animal models [39]. A comparison of the ability of different strains of *C. difficile* to cause disease in the hamster model was undertaken by Rodriguez J.F. and Esteban M. (1987) [40]; the study clearly demonstrated differences in virulence between strains of *C. difficile* and showed that more virulent, produced more toxin A *in vivo* than less virulent strains. The particular serotypes were associated with reduced virulence [41]. Serotyping is based on the fact that strains of the same species of bacteria can differ in the antigenic determinants (such as flagella or fimbriae) expressed on the cell surface.

TcdA and TcdB bind to target cells and once internalised, the cysteine protease domain of the toxin (CPD) and glucosyltransferase domains (GTD) are activated into the cytosol. There, GTD glucosylates and inactivates Rac1 (among other Rho GTPases) using UDP-glucose as the glucosyl donor. The inactivation causes actin-depolymerization resulting in a loss of structural integrity of the cell and eventually cell death [42].

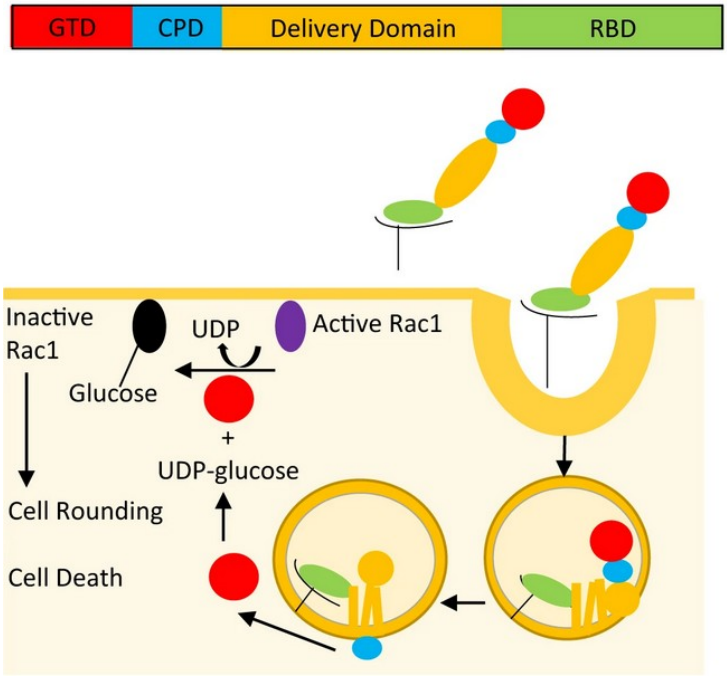

Fig. 4 The schematic of the Tcd toxin domain structure (top) and the mechanism of TcdA and TcdB-induced toxicity of mammalian cells (bottom). N-terminal glucosyltransferase domain GTD (red), cysteine protease domain CPD (blue), delivery domain (orange) and receptor binding domain RBD (green). During the glucosyltransferase reaction catalyzed by TcdB-GTD and TcdA-GTD, GTD glucosylates and inactivates Rho GTPases, including Rac1 among others, in the switch I effector region using UDP-glucose as the glucosyl donor. Adapted from Papparella, A.S. *et al.* (2021) [42].

2.4.2 *Clostridium difficile* transferase

Some strains can also produce a third, large, unrelated toxin, designated *Clostridium difficile* transferase (CDT) otherwise known as a binary toxin. CDT is a virulence-associated determinant whose exact role in disease pathogenesis remains undefined [43]. However, recent studies have shed some light on the possible role of binary toxin with respect to adherence and colonisation of *C. difficile* in the host [44], [45].

2.4.3 Spores formation and germination

Under stress conditions, *C. difficile* produces spores, which are a dormant form of the bacterium. The ability of spores to persist in the environment is attributed to their structure with each spore containing a complete copy of the genome encased in a protective spore core. The core is surrounded by a thick coat of layered peptidoglycan and a loose outer layer, termed the exosporium [11], [46], [47]. This outermost layer, found in spores of a number of

clostridial species, including *C. difficile* [48], [49], is the first point of contact between the pathogen and its environment.

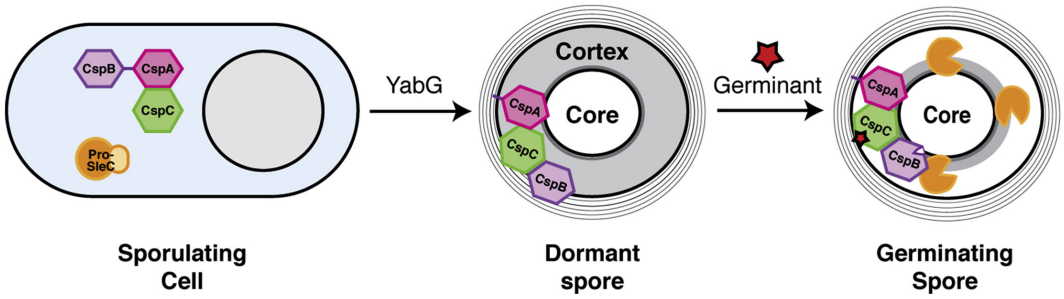

Fig. 5 Main morphogenetic stages of the sporulation of *Clostridium difficile* spore. The serine proteases CspA, CspB and CspC are essential for the germination of spores of *C. difficile* by activating SleC hydrolase for degradation of the peptidoglycan cortex. Adapted from Kevorkian Y. *et al.* (2016) [50].

The diagram above shows that CspA domain controls CspC germinant receptor levels in mature spores and is required particularly when CspA is fused to the CspB protease. During spore formation, the YabG protease separates these domains, although YabG itself is dispensable for germination. The binding of bile salt germinants to the Csp family pseudoprotease CspC triggers a proteolytic signalling cascade consisting of the Csp family protease CspB and the cortex hydrolase SleC. Active SleC degrades the protective cortex layer, allowing spores to resume metabolism and growth [46 - 48]. For a better understanding of *C. difficile* virulence, is crucial to study both the cell and spore states with respect to their genetic, metabolic, and proteomic composition [52].

2.4.4 Adhesion, motility and colonisation

The adherence of spores and vegetative cells to gut epithelial cell lines, and the observation that spores of a ribotype 027 strain adhere better to human epithelial cell monolayers than spores from a ribotype 002 isolate, suggest that adherence may be an important factor in *C. difficile* disease pathogenesis [53], [54]**.** It has been observed that the level of adherence of flagellated strains to the mouse caecum is tenfold higher than the level of adherence of non-flagellated strains [55].

Other putative virulence factors include those that may play a role in adherence and colonisation. Recently, the fibronectin binding protein A (FbpA) was recognised to play a role in *C. difficile* colonization. Other proteins, high and low molecular weight surface layer proteins (SLPs), are predicted to be involved in the adherence of *C. difficile* to host cells during the infection, as well. Cell wall proteins (CWPs) such as Cwp66 and Cwp84 have been shown to be important in the adherence and degradation of the extracellular biofilm matrix [56]. To the adhesion of *C. difficile* to host cells also contribute flagella, fimbriae, and the heat shock protein, GroEL. In addition, it was reported, that opsonisation of *C. difficile* is required for significant phagocytosis, suggesting that there might be an anti-phagocytic factor on its cell surface [57]. While the potency of the toxins in disease may obscure the role of these additional factors in *C. difficile* pathogenesis**,** the role of non-toxin related virulence factors in disease cannot be overlooked. It is vital for the pathogen to be able to hold onto the host cell if the process of pathogenesis is to be successfully started.

Adhesion is an early critical step in colonization but the precise mechanisms by which bacteria adhere to the mucosa and initiate infection remain to be elucidated. Such adhesions include the flagellum [55] and the high-molecular-weight surface layer protein [58].

2.4.5 Role of flagella in virulence

The flagellum is a locomotive organelle for bacterial propulsion. *C. difficile* is known to express peritrichous flagella surrounding the bacterial cell [55].

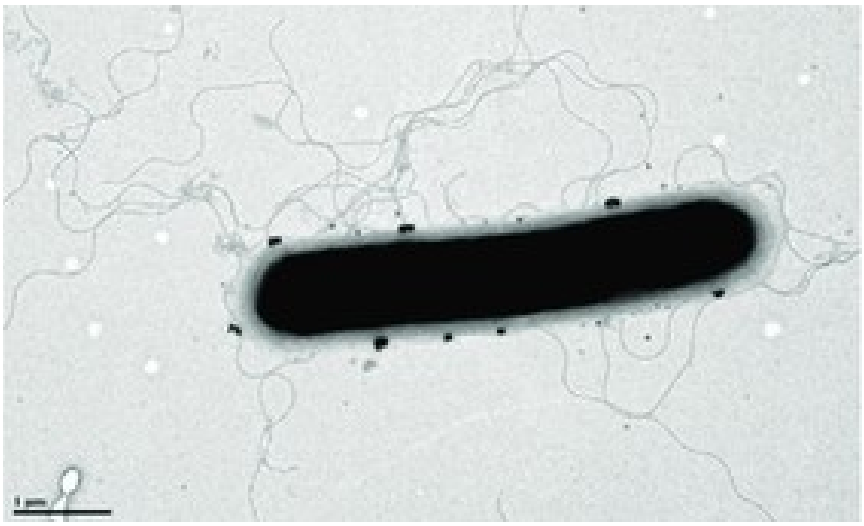

Fig. 6 *Clostridium difficile* forms peritrichous flagella. Flagella (strain LC693) were visualized by transmission electron microscopy. The scale bar represents 1 μM. Adapted from Li C. *et al* (2018) [59].

The external portion of a flagellum is essentially a polymer of a single protein (flagellin) whereas the basal region, protruding through the outer membrane, peptidoglycan layer, and cytoplasmic membrane comprises multiple subunits that anchor and power the organ as presented in Fig. 7. The flagellar motor is powered directly (as opposed to indirectly via adenosine triphosphate) by the proton gradient created across the cytoplasmic membrane by electron transport [60].

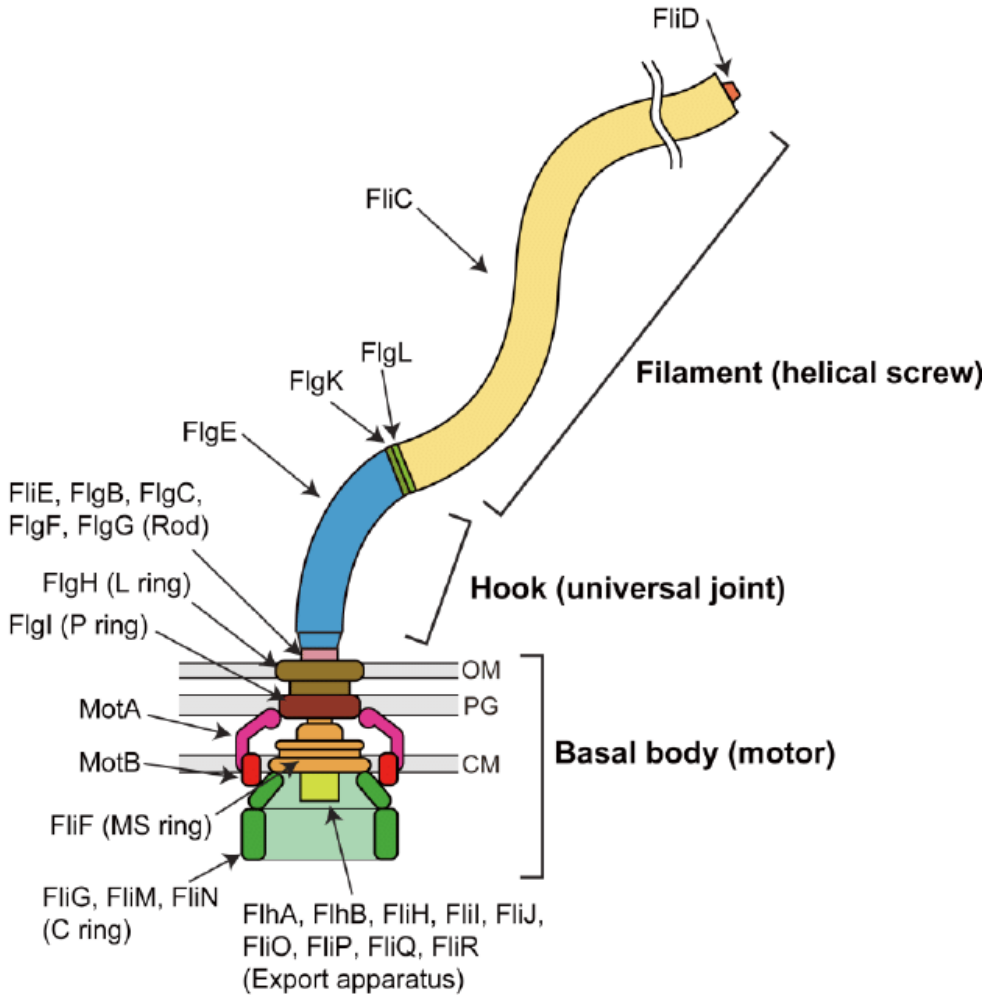

Fig. 7 Schematic diagram of the bacterial flagellum. The flagellum consists of the basal body, which acts as a reversible rotary motor, the hook, which functions as a universal joint and the filament, which works as a helical screw. The flagellar motor consists of a rotor made of FliF, FliG, FliM and FliN flagellins and a dozen stators consisting of MotA and MotB. FliG, FliM and FliN also act as molecular switches, enabling the motor to spin in both directions. OM - outer membrane, PG - peptidoglycan layer, CM - cytoplasmic membrane. Adapted from Morimoto Y. and Minamino T. (2014) [61].

It is tempting to speculate that accumulation of flagella might also increase resistance to antibiotic agents by (i) enhancing cell mobility, (ii) preventing agent incorporation into cell

membranes or (iii) shielding the cell's surface by forming capsular-like structures. All published studies coincide that flagella-mediated motility might contribute to the overall fitness of the bacteria. However, as the heavily flagellated but immotile *sigH* mutant showed, in a study by Maaß A. *et al*. (2021), increased resistance to novelty antibiotic agent, nisin, when compared to the wild type strain, positive effect of increasing the number of flagella during adaptation cannot be justified by enhanced cell mobility, which makes the second and third hypothesis or a combination thereof more likely [62]. Indeed, for most gastrointestinal pathogens, including *C. difficile* 630Δerm, flagella-mediated motility is recognized as essential virulence factor [63] and there is even strong evidence that the expression of flagellar genes in *C. difficile* is coupled to toxin gene regulation [64].

In recent years, a rapidly increasing body of work has described the process of flagellar glycosylation in a diverse number of bacterial species. In some cases, it has been demonstrated that the process of flagellar glycosylation has a role in both flagellar assembly and the interactions between the host and its pathogens [65].

2.4.6 Flagellar proteins

The two best characterized *C. difficile* flagellar proteins include FliC, the major flagellin structural monomer, and FliD, the cap protein [53],[62]. The role of flagella in the pathogenesis of CDI is contentious and appears to be strain dependent [63],[64]. Studies with some isolates, including the epidemic PCR ribotype 027 strain R20291, showed flagella to be important in adherence to cells and intestinal tissue and that there are striking differences between *C. difficile* strains [67], [68].

However, not forgetting proteins *associated* with *flagellar* formation and *function* that are of major interest as facilitators to flagella assembly such as FliS, the flagellin-specific chaperone that binds flagellin and facilitates its export, to prevent premature polymerization of newly synthesized flagellin molecules [69]. Also heat shock protein 90 (HtpG), a genetically conserved member of the heat shock protein family found in eukaryotes and prokaryotes, is involved in a variety of cellular processes including protein folding, repair, and signal transduction [70]. And lastly, enzymes involved in rhamnose biosynthesis were observed. The importance of this flagellin glycosylation was demonstrated for *rmlB* in *Burkholderia cenocepacia* with consequent Human Innate Immune Responses [70].

2.5 Molecular typing of *C. difficile*

Molecular typing such as polymerase chain reaction (PCR) ribotyping is used to monitor the spread of *C. difficile* within Europe. This was determined in an extensive study on 1,196 *C. difficile* isolates from diarrhoeal samples sent to the European coordinating laboratory in 2012/13 and 2013 by 482 participating hospitals from 19 European countries [71]. This was a part of the European, multicentre, prospective, biannual, point-prevalence study of CDI in hospitalised patients with diarrhoea, the largest *C. difficile* epidemiological study of its type, where 125 distinct ribotypes were identified, with considerable intercountry variation in ribotype distribution. Circulating strains of *C. difficile* vary in their susceptibility to antibiotics, and some strains are difficult to treat.

In 2016, the PCR ribotype data were available for 1 326 out of 3 894 (34.1%) cases with enhanced case-based data. The most common PCR ribotypes were RT027 (n=303, 22.9%), RT001 (n=99, 7.5%), RT014 (n=89, 6.7%), RT078 (n=68, 5.1%), RT002 (n=56, 4.2%) and RT 020 (n=56, 4.2%) [72]. European Centre for Disease Prevention and Control (ECDC) encourages European Union/European Economic Area countries to recruit hospitals to collect data compatible with the ECDC surveillance protocol in order to acquire standardised epidemiological and microbiological information on their hospital CDI burden compared to other European hospitals.

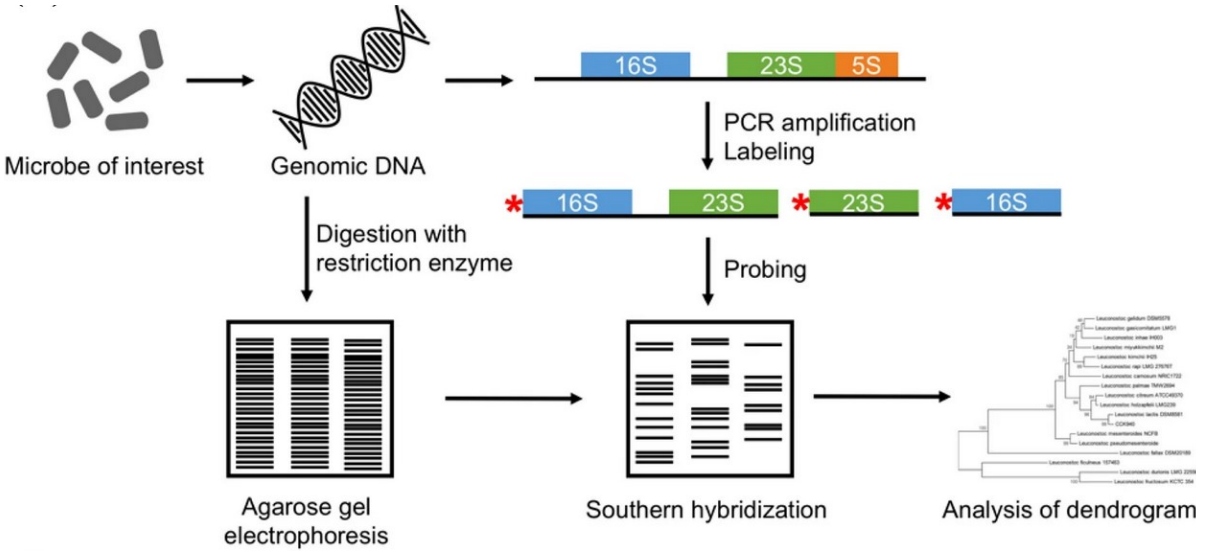

Fig. 8 Schematic diagram for ribotyping. The diagram describes the ribosomal genes present within the bacterial genome that are recognized using nucleic acid probes. This combined detection is based on the digestion of genomic DNA of the microbe of interest with

a restriction endonuclease and Southern hybridization using rDNA cistrons (16S, 23S, and 5S rRNA genes) as the labelled probes. Adapted from Sharma A. *et al.* (2020) [73].

Molecular typing is a crucial part of monitoring the occurrence and spread of *C. difficile*. Over a three-year period from 2013 to 2015, clinical *C. difficile* isolates from 32 Czech hospitals were collected for molecular characterisation [74]. While providing higher discrimination than other typing techniques [75], this method is not fully portable between laboratories, is labour intensive, has a turnaround time of up to a week, and often requires in-house optimization. PCR-ribotyping has been a major typing technique for the past decades, but many clinical laboratories progress to switch toward whole genome sequencing (WGS) based typing [76], [77].

2.6 Proteomics

Proteomic analysis (proteomics) refers to the systematic identification and quantification of the complement of proteins (a system of plasma proteins that can be activated directly by pathogens), also known as the proteome, of a biological system (in this case bacterium *C. difficile*) at a specific point in time. Proteomics are using two major techniques. In the Aebersold R. and Mann M. study, it was mass spectrometry (MS) has become a powerful method for the analysis of complex mixtures of proteins [78].

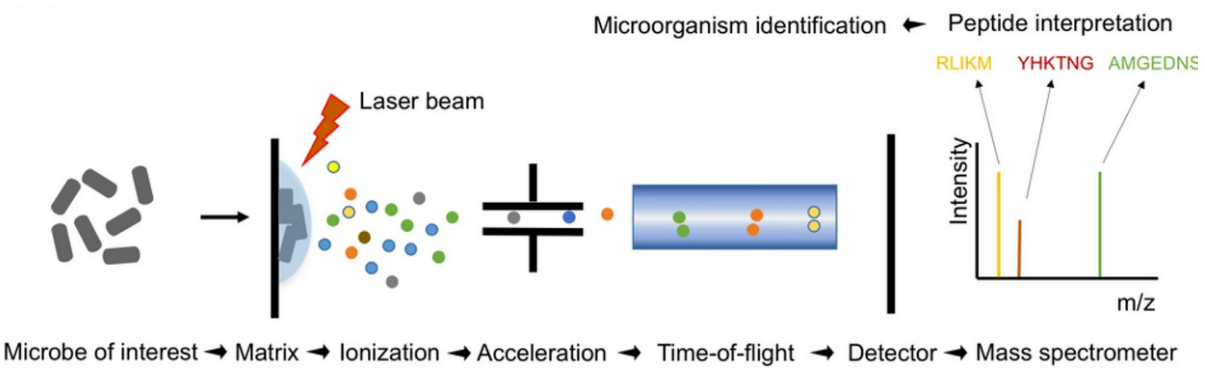

Fig. 9 Diagram for mass spectrometry. Three main components are included in a mass spectrometer: an ion source to form ions (ionization) and transfer them into a gas phase, a mass analyzer is the component of the mass spectrometer that takes ionized masses and separates them based on mass to charge (m/z) ratios and outputs them to the detector to record the number of ions at each m/z value. Adapted from Sharma A. *et al.* (2020) [73].

MS allows the determination of the molecular mass of proteins or peptides based on the m/z ratio of ions in the gas phase. To determine the primary structure of selected proteins, by fragmenting a selected protein in order to deduce the amino-acid sequence, a technique known as tandem MS (MS/MS) is employed. High Performance Liquid Chromatography (HPLC) is an alternative separation technique for proteomic studies, especially in the separation and identification of low molecular weight proteins and peptides [79].

Although there are a number of variants and side steps that are used for some experiments, optimal analysis of shotgun proteomics experimental data will usually involve most or all of these steps: conversion to and processing via open data formats, spectrum identification with a search engine, validation of putative identifications, protein inference, quantification, the organization in local data management systems, interpretation of the protein lists, transfer to public data repositories [80].

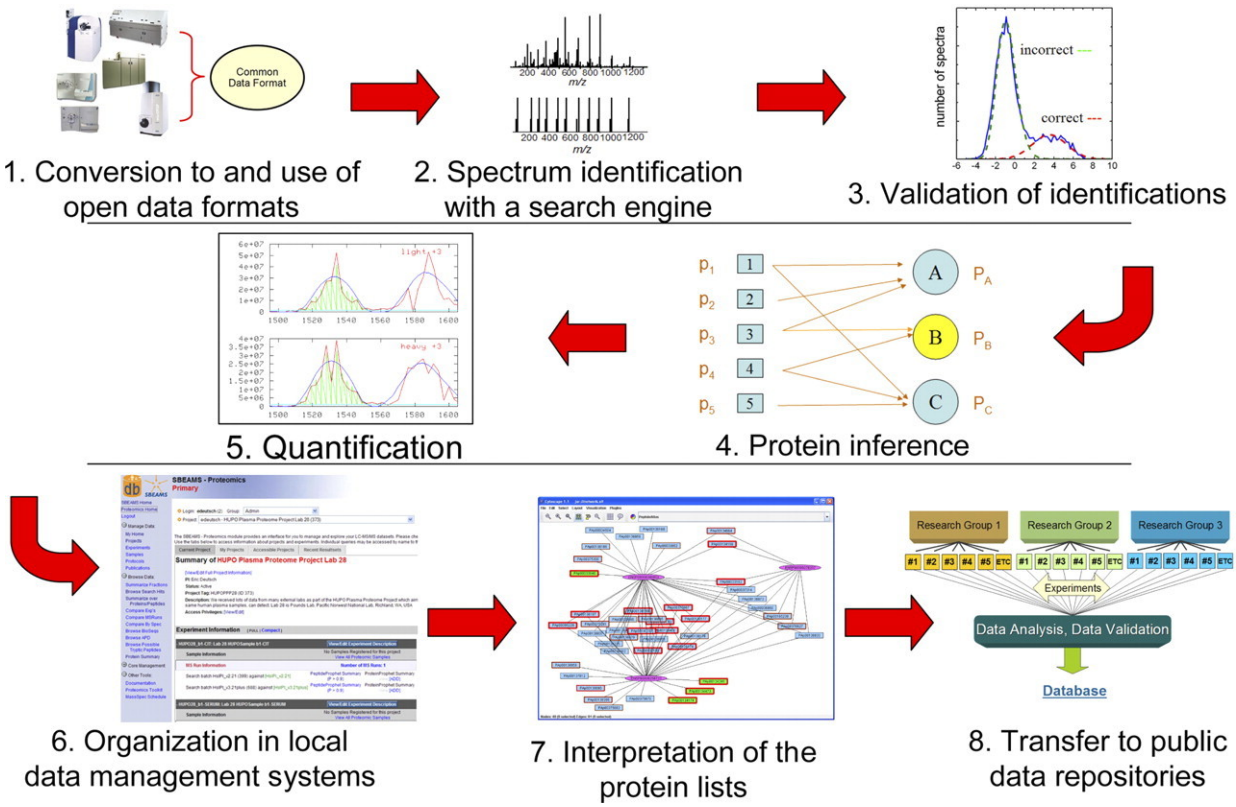

Fig. 10 Typical workflow of the proteomics informatics processing of a data set. Not all data sets go through all steps, but most data sets will benefit from some aspect of processing from most of the steps. Adapted from Deutsch E.W. *et al* (2008) [80].

From the data generated by the MS, the protein is either sequenced *de novo* by manual mass analyses or processed automatically via sequence search engines. These algorithms are developed based on the correlation between experimental and theoretical MS/MS data; the latter being generated from in silico digestion of protein databases [80].

MS-based label-free quantification aims to directly compare relative abundances of proteins across multiple liquid chromatography with tandem mass spectrometry (LC-MS/MS) experiments without utilizing stable isotopes or isotopic tags. There are two commonly used quantitative schemes: mass spectral peak intensities and spectral counting. Label-free quantification is based on the observation that the amount of protein correlates well with peak intensities or spectral counts of peptides unique to a specific protein [81].

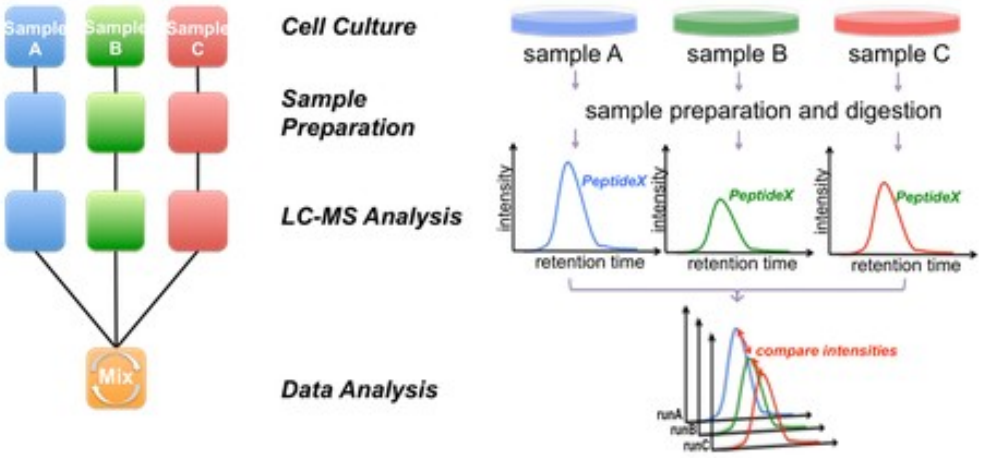

Fig. 11 Schematic of label free quantification. Compared with stable isotopic labeling methods, label-free quantification is cost-effective and it only requires minimal sample preparation and allows comparison across multiple experimental conditions. From Zentrum für Molekulare Biologie der Universität Heidelberg website, accessed on 15.3.2022 [82].

3 Aim of the work

The aim of this work is to monitor the relationship between the secretion of individual flagellar proteins in *C. difficile* strains of variable virulence. The outcome of this work should be different observations in strains with increased potential for virulence, which have not been described at the proteome level. The results of the comparative semi quantitative analysis, could be instrumental in the future development of targeted therapy and diagnostic tools.

4 PRACTICAL PART

4.1 Materials

Acetonitrile (Sigma-Aldrich, USA)

AGC target in the Orbitrap (Thermo Fisher Scientific, DE)

Amicon® Ultra 10 kDa millipore filters (Sigma-Aldrich, USA)

Ammonium bicarbonate (Sigma-Aldrich, USA)

Bicinchoninic acid assay (QuantiPro™ BCA Assay Kit) (Sigma-Aldrich, USA)

C18, 3 µm, 100 Å, 0.075 × 20 mm trap column (Dionex, USA)

Commercial immunochromatographic assay (Vidia, CZ)

Empore™ SPE Cartridges, C18, standard density, bed I.D. 4 mm (Sigma-Aldrich, USA)

Formic acid (Sigma-Aldrich, USA)

Guanidinium chloride (Sigma-Aldrich, USA)

Iodoacetamide (Sigma-Aldrich, USA)

MASCOT (Matrix Science, UK)

Multiskan Spectrum plate reader (Thermo Fisher Scientific, DE)

PepMap RSLC C18, 2 µm, 100 Å, 0.075 × 150 mm analytical column (Dionex, USA)

Proteome Discoverer software v. 1.4 (Thermo Scientific, DE)

Q-Exactive mass spectrometer using a Nanospray Flex ion source (Thermo Scientific, DE)

Sequencing grade trypsin (Promega, CZ)

Schaedler Anaerobe Agar CM0437 (Oxoid Ltd., USA)

Signature for TcdA and TcdB protein (SpikeTides TQL, JPT Peptide Technologies, DE)

Skyline software (MASCOT.dat file)

Sulphuric acid 3M (Sigma-Aldrich, USA)

Thioglycollate medium USP (Oxoid Ltd., USA)

Trifluoroacteic acid (Sigma-Aldrich, USA)

Tris (2-carboxyethyl) phosphine hydrochloride (Sigma-Aldrich, USA)

UltiMate 3000 RSLCnano system (Thermofischer, DE) controlled by Chromeleon software (Dionex, USA)

UniProt/KB (www.foundation.wwpdb.org)

4.2 Cultivation

The *C. difficile* isolates were sourced from the Czech National *C. difficile* strain collection [74]. Seven of these isolates were cultured from stool samples of hospitalized patients with CDI positive for diarrheal glutamate dehydrogenase (GDH) and toxin A/B. As a study negative control, lacking toxigenicity, was included RT 010 sample cultured from diarrheal GDH positive and toxin A/B negative stool sample of a patient with *Candida*-acquired diarrhoea.

4.2.1 Optimisation of sample preparation

Based on the pilot optimisation of targeted mass spectrometry analysis of TcdA and TcdB the representative strains were selected for analysis, see Table 12 below, including RT 010 sample as a negative control lacking toxigenicity.

Table 12 Characterization of *Clostridium difficile* isolates in the study

Nr.	Isolate number*	Year of isolation	Patient age#	Sex	RT	ST	Clade	Presence of toxin genes	*tcdC* truncation
1	2063	2015	82	M	001	3	1	*tcdA*, *tcdB*	no
2	2023	2015	73	M	005	6	1	*tcdA*, *tcdB*	no
3	1107	2014	25	F	010	15	1	non-toxigenic	NA
4	2006	2015	74	M	012	54	1	*tcdA*, *tcdB*	no
5	1120	2014	33	F	014	2	1	*tcdA*, *tcdB*	no
6	854	2014	35	M	027	1	2	*tcdA*, *tcdB*, *cdtA, cdtB*	Δ117
7	2004	2015	54	M	078	11	5	*tcdA*, *tcdB*, *cdtA, cdtB*	C184T
8	2062	2015	84	F	176	1	2	*tcdA*, *tcdB*, *cdtA, cdtB*	Δ117

The isolate number (*) in the previous table refers to the number in the Czech national *C. difficile* strain collection, patient age (#) is the age of the patient at the time of isolation, M–male, F-female, RT - PCR-ribotype, ST - sequence type. Toxin production varies by clade and the *C. difficile* population structure consists of six distinct phylogenetic clades designated 1, 2, 3, 4, 5 and C-I [83]. The presence of toxins was evaluated using Diasorin LIAISON®

commercial test, tcdC truncation is single nucleotide polymorphism (SNP) resulting in TcdC protein truncation to identify *C. difficile* genotypes associated with recurrent CDI.

The *C. difficile* isolates were recovered from the frozen stocks by inoculating onto the Schaedler Anaerobe Agar CM0437 (Oxoid) and cultured for 48 hours at 37 °C under anaerobic conditions. Toxin production of all strains in the study was confirmed using a commercial immunochromatographic assay (Vidia, Czech Republic) for the detection of free toxins A and B in the stool samples previously when the bacterial suspension was investigated as a stool sample. The bacterial mass was resuspended in thioglycolate medium USP (Oxoid) and the amount of bacteria measured as colony forming units (CFU) was evaluated via optical density (OD) analysis at 595 nm (Multiskan Spectrum plate reader, Thermo Fisher Scientific), considering that OD 1 in 1 mL of thioglycolate medium corresponds to 2.4×10^6 CFU. Later, 9 mL of thioglycolate medium was inoculated in triplicate for each representative strain to OD 1.99 and cultivated for 5 days at 37 °C under anaerobic conditions.

OD was also measured at the end of the cultivation and reached comparable values among the cultures. Capillary electrophoresis ribotyping of *C. difficile* isolates was performed, using primers [84] before resuspension in thioglycolate medium USP and after 5 days of culture before proteomic analysis.

4.3 Sample preparation

The processing of samples follows the internal protocols for preparation of stock solutions, sample preparation and solid phase extraction using Empore SPE Cartridges.

1. I prepared stock solutions needed for further treatment of the samples:
    a. 100mM ammonium bicarbonate (ABC), pH 7.8; by dissolving 1.9775 g ABC in 25.0 mL of distilled water.
    b. 6M guanidine buffer; by dissolving 0.8598 g of guanidine hydrochloride into 20.0 ml of distilled water. Once dissolved, complete the volume to 40.0 ml with distilled water.
    c. 100 mM solution of tris(2-carboxyethyl)phosphine (TCEP) in 6M aqueous guanidine hydrochloride; by dissolving 0.7175 g of TCEP and 14.33 g of guanidine hydrochloride in 25.0 mL of distilled water.
    d. 300 mM iodoacetamide; made fresh just prior to use, by resuspending 0.5040 g of iodoacetamide with 15.0 ml of distilled water.
    e. 2.0 µg trypsin /5 µl of solution in acetic buffer; we have a stock solution of bovine sequencing grade trypsin of 0.5 µg/µl in 1mM HCl. Diluted 100x in ABC before use.
2. Following the pelleting of bacterial cells by centrifugation to remove all bacterial cells from the proteomes released from *in vitro* cultures, the pH of the supernatants was adjusted to 3.5 with 3 M sulphuric acid. After an overnight precipitation at 4 °C, the pellets were recovered by centrifugation. On the cooling centrifuge I spun down the samples and removed the supernatant.
3. Pellets were resuspended by washing with 400 µl of 100 mM ABC, spun down for 20 minutes.
4. Saved 50 µl of the sample for the bicinchoninic acid (BCA) analysis. Proteins were quantified by bicinchoninic acid assay. Resuspended pellets were transferred onto Amicon® Ultra 10 kDa filters (Millipore) and washed twice with 100 mM ammonium bicarbonate.
5. Subsequently, the samples were denatured by 6M guanidinium chloride reduced with 100 mM Tris (2-carboxyethyl) phosphine hydrochloride [85] and alkylated with 300 mM iodoacetamide.

6. Added 400 µl of 6M guanidine buffer and spun down for 15 minutes at 18,000 g, added 400 µl of 6M guanidine buffer and spun again for 20 minutes.
7. To reduce S-S bridges, I added 35 µl of 100 mM solution of TCEP in guanidine buffer and reacted at 55°C for 45 minutes.
8. Without spinning, to alkylate the free C, added 35 µl of 300 mM iodoacetamide in guanidine buffer at 37°C for 30 minutes in dark.
9. Spun down for 20 min at 11,000 g.
10. Washed twice with 200 µl of 100mM ABC.
11. Finally, the samples were digested with 2.0 µg of sequencing grade trypsin. Added 2.0 µg of trypsin in 5 µl of solution in acetic buffer, in total 200 µl 50mM ABC, digested overnight at 37 °C.
12. The next day I spun at 12,000 g for 35 minutes into a fresh vial.
13. Added 200 µl 50mM ABC, spun down a filtrated for further processing.
14. Solid phase extraction using Empore SPE Cartridges 4mm/1 ml that are designed for sample pretreatment to remove or minimize sample matrix and other interferences to cleanup a sample prior to analysis.
    a. Condition the disk with 500 µl MeOH and open the vacuum source to about 0.51 bar.
    b. Load and extract the sample twice at a maximal volume 250 µl.
    c. Wash out interferences with 500 µl 5% acetonitrile (ACN)/0.1% trifluoroacteic acid (TFA).
    d. Before the mass spectrometry analysis, the samples were resuspended in 30 µL of 2% ACN/0.1% and loaded into 0.5 ml eppendorf vials.

A set of 10 µl samples was sent to Hradec Kralove where samples were further analyzed by liquid chromatography with tandem mass spectrometry (LC-MS/MS) techniques involving targeted mass spectrometry and label-free quantification (LFQ). Subcellular localization of the proteins released from *in vitro* cultured panel was evaluated by bioinformatic tools (data not shown). The leftover sample material was stored at -80°C.

4.4 Method

4.4.1 Liquid Chromatography with tandem MS analysis (LC-MS/MS)

An UltiMate 3000 RSLCnano system controlled by Chromeleon software (Dionex, USA) was used for chromatography separation. 1 µL of each sample (10x diluted) was loaded onto a PepMap100 C18, 3 µm, 100 Å, 0.075 × 20 mm trap column (Dionex) at 5 µL/min for 5 min. Peptides were separated on a PepMap RSLC C18, 2 µm, 100 Å, 0.075 × 150 mm analytical column (Dionex) by a gradient formed by mobile phase A (0.1% formic acid, FA) and mobile phase B (80% ACN, 0.1% FA), running from 4 to 34% in 68 min, and from 34 to 55% of mobile phase B in 21 min at a flow rate of 0.3 µL/min at 40 °C. Eluted peptides were on-line electrosprayed into Q-Exactive mass spectrometer using a Nanospray Flex ion source (Thermo Scientific, Bremen, Germany). Positive ion full scan MS spectra (*m/z* 350-1650) were acquired using a 1×10^6 AGC target in the Orbitrap at 70 000 resolution. Top 12 precursors of charge state ≥2 and threshold intensity of 5×10^4 counts were selected for HCD fragmentation, with a dynamic exclusion window of 30 s. The isolation window of 1.6 Da and normalized CE of 27 was used. Each MS/MS spectrum was acquired at a resolution of 17,500, with a 10^5 AGC target and a maximum 100 ms injection time.

Higher-energy C-trap dissociation (HCD) MS/MS spectra were searched in Proteome Discoverer software v. 1.4 (Thermo Scientific) using MASCOT (Matrix Science, London, UK). The reference proteome set of *Peptoclostridium difficile* strain R20291 was downloaded from UniProt/KB in July 2015 and merged with common contaminants file downloaded from the MaxQuant web page (http://www.maxquant.org/downloads.htm); the merged database contained 3754 sequences. The tryptic specificity was set with a maximum of 2 missed cleavages. The mass tolerance was set to 10 ppm for precursors and 20 mmu for product ions. Cys carbamidomethylation was set as a fixed modification and methionin oxidation was set as a variable modification. Spectra explained with a cut-off score of 0.95 were extracted by Skyline software from the MASCOT.dat file and stored in an MS/MS library.

4.4.2 Targeted mass spectrometry analysis of TcdA and TcdB

Based on the analyzed tryptic digests of TcdA and TcdB, the candidate peptides were compared with the whole UniProt database (ver. 20140907) using the homemade script running in BioPearl to ensure their specificity. Finally, four tryptic peptides (two for each protein) were considered as signature for TcdA and TcdB protein and purchased as isotopically labeled synthetic equivalents (SpikeTides TQL, JPT Peptide Technologies, Berlin, Germany).

These peptides were subsequently used in known concentrations as spiked standards either for background interferences measurement or for the absolute quantification of these peptides across the individual samples. The synthetic signature peptides were spiked into the pooled sample from toxin negative strains (background sample) in known concentrations ranging from 40-160 fmol/µl per injection. Three replicates for each concentration step were measured in the lowest to the highest concentration order. The coefficient of variation of these three replicates was always below 20 percent. The linear regression analysis on the observed native vs labelled peak area ratios versus concentration was performed to prepare the curve for testing of background versus peptide interactions.

Finally, each sample was spiked by the concentration of 40 fmol of heavy labelled peptides and the ratio of intensities between the analyzed sample and toxin negative control (RT 010 sample) was counted.

4.4.3 Label-free quantification (LFQ) raw data processing of C. difficile proteins

Raw files acquired by LC-MS/MS analysis were further analyzed in MaxQuant (ver.1.5.3.30) and the Andromeda was used as a search engine to search the detected features against the *C. difficile* strain R20291 (RT 027) or against *C. difficile* strain 630 (RT 012) databases downloaded from Uniprot (on 17.6.2015). Only tryptic peptides that were at least seven amino acids in length with up to two missed cleavages were considered. The initial allowed mass tolerance was set to 4.5 ppm at the MS level and 0.5 Da at the MS/MS level. The oxidation of methionine was set as a variable modification and carbamidomethylation of cysteine was set as a fixed modification. A false discovery rate (FDR) of 1 % was imposed for peptide-spectrum matches (PSMs) and protein identification using a target-decoy approach.

Relative quantification was performed using the default parameters of the MaxLFQ algorithm with the minimum ratio count set to 2.

4.4.4 Label-free quantification (LFQ) data analysis of *C. difficile* proteins

The "proteinGroups.txt" MaxQuant output file was uploaded into Perseus (ver.1.5.2.6) for further filtering and statistical analysis. Decoy hits, proteins only identified by site, and potential contaminants were removed. Only those protein groups quantified in at least two replicates of at least one ribotype representative triplicate were considered for further $\log_2$ transformation of LFQ intensities.

4.5 Results

I specifically participated in the sample preparation workflow and prepared the final optimisation set of samples. I did my own sample preparation and bicinchoninic acid assay as described in detail in chapter 4.3, Sample preparation. I processed the data about the flagellar proteins analysed in these experiments.

My colleagues Jiri Dresler, Marcela Krutova and Otokar Nyc designed the experiments. Jiri Dresler, Marcela Krutova, Jana Klimentova and Petr Pajer wrote the manuscript. Alena Fucikova and Jana Klimentova conducted the mass spectrometry analyses, Veronika Hruzova, Miloslava Duracova also participated in the optimization of the sample preparation workflow. Bara Salovska, Jiri Dresler and Libor Pisa performed data analysis. Jana Matejkova assisted with microbiological experiments.

The bacterial stocks were inoculated to 9 ml of Thioglycolate medium and the strains were cultured in triplicate (a total of 24 samples) and the OD analysis was performed at 595 nm (Multiskan Spectrum plate reader, Thermo Fisher Scientific). Based on the previously observed results for *Clostridium botulinum* cultivation (data not shown), the average CFU for final thioglykolate culture 9.83×10^6 CFU/ml corresponds to OD 2.34 (595 nm), the CFU at the beginning of the cultivation was set using above mentioned formula to 8.4×10^6 CFU (data not shown). The processed samples were then quantified by bicinchoninic acid assay for their protein content and the results are summarized in Table 13 overleaf.

Table 13 Measured values of OD and calculated concentrations and protein content of prepared samples.

	Absorbance							
Sample Nr	**150× 1st value**	**150× 2nd value**	**150× 3rd value**	**150× 4th value**	**Average**	**Concentration (µg/ml)**	**c (µg/ml) undiluted**	**Protein content in 350µl (ug)**
1	0.131	0.130	0.163	0.142	0.142	11.672	1750.737	612.758
2	0.190	0.180	0.199	0.117	0.172	13.861	2079.206	727.722
3	0.139	0.126	0.142	0.120	0.132	10.960	1643.985	575.395
4	0.141	0.156	0.183	0.165	0.161	13.113	1966.979	688.443
5	0.166	0.205	0.112	0.148	0.158	12.858	1928.658	675.030
6	0.123	0.139	0.148	0.157	0.142	11.690	1753.474	613.716
7	0.161	0.161	0.171	0.157	0.163	13.204	1980.665	693.233
8	0.172	0.170	0.191	0.194	0.182	14.610	2191.432	767.001
9	0.165	0.177	0.170	0.150	0.166	13.423	2013.512	704.729
10	0.267	0.255	0.255	0.254	0.258	20.157	3023.553	1058.243
11	0.168	0.195	0.185	0.168	0.179	14.409	2161.323	756.463
12	0.172	0.182	0.184	0.200	0.185	14.810	2221.542	777.540
13	0.204	0.216	0.186	0.188	0.199	15.832	2374.827	831.190
14	0.177	0.203	0.196	0.186	0.191	15.248	2287.236	800.532
15	0.122	0.143	0.124	0.149	0.135	11.161	1674.094	585.933
16	0.269	0.190	0.183	0.173	0.204	16.215	2432.309	851.308
17	0.139	0.169	0.148	0.163	0.155	12.639	1895.811	663.534
18	0.137	0.154	0.134	0.140	0.141	11.653	1748.000	611.800
19	0.157	0.198	0.147	0.179	0.170	13.770	2065.519	722.932
20	0.229	0.205	0.207	0.228	0.217	17.201	2580.120	903.042
21	0.163	0.182	0.153	0.163	0.165	14.206	2130.831	745.791
22	0.190	0.299	0.167	0.163	0.205	17.268	2590.137	906.548
23	0.085	0.091	0.097	0.095	0.092	8.527	1279.080	447.678
24	0.190	0.124	0.161	0.184	0.165	14.167	2125.017	743.756

Using the LFQ approach, a total of 662 quantifiable proteins in the study were analyzed (see Additional file 1: Table S3 in Dresler J. *et al.* [86]). The observed quantities of the proteins are depicted as $\log_2$ transformation of LFQ intensities. The values ranging from 22 to 34 reflect the dynamic range of the mass spectrometry based workflow. The LFQ intensities below this value are considered non-analyzable by the implemented qualitative test [87]. At that point, it is not possible to state if the protein is not released or if the quantity released is not detectable.

The most shared proteins were observed in RTs 027 and 176 (n = 563) and the lowest rate was revealed in RT 078 compared to all RTs in the study (n = 454–479). Pathway mapping in KEGG was done using the DAVID classification tool [88] against *C. difficile* strain 630, and several biological processes were annotated to 40.1% of the quantifiable proteins.

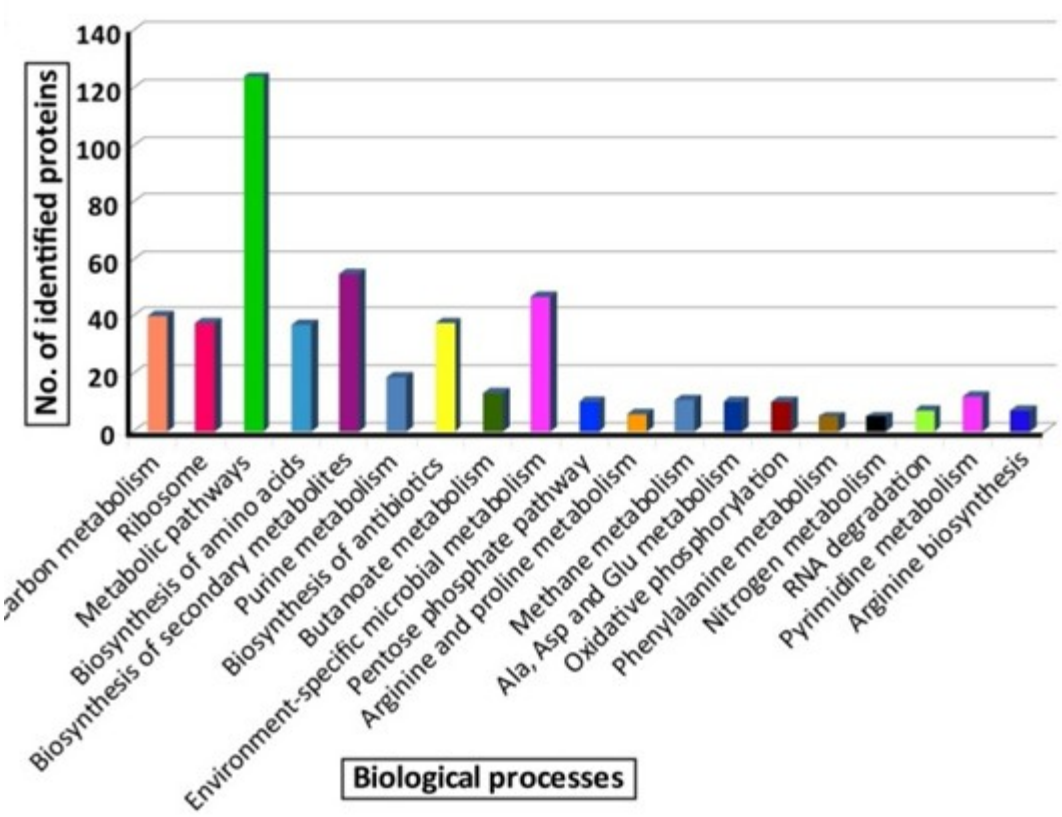

Fig. 14 Quantities of shared proteins among different ribotypes of Clostridium *difficile*. Based on MS analysis the proteins were quantified into groups by their function and pathway mapping in KEGG. Adapted from Dresler J. *et al.* (2017) [13]

4.5.1 Comparison of proteomes of individual RTs

To assess the applicability of the LFQ approach, we examined the similarity of individual proteomes using hierarchical clustering and principal component analysis (PCA). Unsupervised cluster analysis of protein expression profiles was performed using Euclidean distances. Statistical procedures were performed using the computational platform Perseus. Both hierarchical clustering and PCA generated eight distinctive groups encompassing each biological triplicate of analyzed RT representatives and showed the applicability of this workflow. For example, RTs 027 and 176 nearly co-cluster, on the contrary proteomes from RT 078 created a distinctive group.

The selected proteins were chosen as candidates passing through ANOVA statistical test and subsequent Fisher exact test for increase in functional annotations. Furthermore, ANOVA (permutation-based FDR 5%, S0 = 0) was used to identify significant differences in protein expression between the RTs. Only ANOVA-significant hits were used for subsequent hierarchical clustering using Euclidean distances to group proteins with similar expression profiles. Finally, Fisher exact test at 2% Benjamini–Hochberg FDR was applied to determine significantly overrepresented functional annotations for each one of the identified clusters.

In total, 27 proteins were found to be significantly increased in the functional annotations, of which 17 were relevant to my thesis due to their involvement in the assembly of *C. difficile* flagellum. Their LFQ intensities measured are detailed in Table 15 below. The complete LFQ intensities results are shown in the Additional file 1: Table S2 in Dresler J. et al. (2017) [13].

Table 15 LFQ analysis of proteins significantly increased in functional annotations depicted as log2 median protein LFQ intensities for each PCR-ribotype. NaN - represents values not detected.

Proteins involved in assembly of *C. difficile* flagellum			LFQ intensities (medians)							
Protein IDs	**Gene**	**Other name**	**PCR ribotype 010**	**PCR ribotype 014**	**PCR ribotype 012**	**PCR ribotype 005**	**PCR ribotype 001**	**PCR ribotype 078**	**PCR ribotype 027**	**PCR ribotype 176**
C9YI65	*flgE*		NaN	26.9150	NaN	27.8812	NaN	NaN	28.3666	28.8633
C9YI79	*flgG*		NaN	26.0100	NaN	26.5176	NaN	NaN	27.1453	27.7334
C9YI39	*flgK*		NaN	25.8299	NaN	27.1115	NaN	20.8008	27.4089	27.8577
C9YI40	*flgL*		NaN	26.9301	NaN	28.2372	NaN	NaN	28.8744	29.2726
C9YI47	*fliC*		22.6972	31.3676	24.8220	32.0720	23.0162	24.4515	33.4973	33.4187
C9YI45	*fliD*		NaN	28.5824	NaN	295971	NaN	24.4789	30.8483	31.3318
C9YI63	*fliK*		NaN	24.3520	NaN	24.2037	NaN	NaN	25.4507	25.9487
C9YI80	CDR20291_0273	Flagellar basal body protein	NaN	25.7288	NaN	25.6922	NaN	NaN	27.1761	27.5685
C9YI56	*flgC*		NaN	23.3975	NaN	21.6154	NaN	NaN	23.3294	24.2061
C9YI37	*flgM*		NaN	22.7971	NaN	23.0258	NaN	NaN	24.9594	24.9965
C9YI69	*fliL*		NaN	NaN	NaN	24.3746	NaN	NaN	24.1803	24.3687
C9YI57	*fliE*		NaN	NaN	NaN	NaN	NaN	NaN	24.7266	25.1837
C9YI34	CDR20291_0227	Glycosyltransferase	NaN	23.5203	NaN	23.5857	NaN	NaN	25.4568	25.7181
C9YI31	*rmlA*		27.3562	28,0121	26,8858	28.2999	27.3104	29.7989	27.3151	27.0824
C9YI32	*rmlC*		22.8392	22.7973	NaN	23.4473	NaN	23.3886	NaN	23.2446
C9YI44	*fliS2*		NaN	NaN	NaN	23.8684	NaN	NaN	23.7849	24.1603
C9YI84	*htpG*		24.8344	24.2903	24.9792	25.8037	25.1965	25.9804	25.2339	25.1192

4.5.2 Bioinformatic analysis of subcellular localization of the proteins

To identify the subcellular localization of the proteins released from in vitro cultured panel of included RTs, the bioinformatic analyses focused on Sec pathway and alternative secretion modes markers were performed. The sequences of all identified proteins were processed with SignalP 4.1.[89] The secretion type of protein identified was predicted as "classical" with *Signal P Score* > 0.5. Furthermore, the Secretome P 2.0 tool [90] was employed. Using default parameters for gram-positive bacteria and Secretome P Score > 0.5, proteins were predicted as "alternatively secreted". A majority of the proteins found to be significantly increased in functional annotations were predicted to be secreted via sec-dependent secretion pathway or via an alternative secretion system proposing the overlapping of the supernatant proteome and the secretome. The MS data of these proteins significantly increased in the functional annotations see Additional file 1: Table S4 - proteins designated as SigP for classical and SecP for alternative secretion.

However, regarding the flagellar proteins, FliC, essential for flagella formation and motility, and FliL were not predicted to be secreted via any secretion system and are reportedly localized intracellularly. It has been shown that FliL is required for flagellar function but it is not part of the transcriptional hierarchy, supporting the hypothesis that, as is the case for the entries, the regulatory hierarchy responds to assembly cues rather than directly to the expression of flagellar proteins [91].

4.5.3 Proteins involved in the assembly of *C. difficile* flagellum

Out of the proteins significantly increased in functional annotations I focused specifically on the selection of flagellar proteins that are connected with pathogenicity as described in chapter 2.4.6, Role of flagella in virulence and resistance.

A higher expression of flagellar proteins FlgE, G, K, L, Fli C, D, K, and flagellar basal body protein C9YI80 was observed among RTs 027, 176, 005, and 014 except for FlgK, flagellar hook-associated protein, and FliD, flagellar cap protein, quantifiable also in RT 078. Moreover, FlgC, flagellar basal body protein, levels were increased only in RTs 027, 176, and 014, whilst FlgM, a regulatory protein that couples transcription of the flagellar class 3 promoters to completion, and FliL were found to be increased in RTs 027, 176, and 005. Interestingly, FliE, a rod adapter protein, was expressed only in RTs 027 and 176, and

glycosyl transferase C9YI34, which is involved in the posttranslational modification of flagella, was expressed in RTs 027, 176, 005, and 014.

Observations on flagellin-specific partners or chaperones revealed higher expression of FliS2, the flagellin-specific T3S chaperone, in RTs 005, 027 and 176 and heath shock protein HtpG, a molecular chaperone involved in protein folding and signalling, was found in all RTs. Other proteins RmlA and RmlC, which are required for flagellin glycan, were expressed in all the RTs except for 001, 012 and 027.

4.6 Discussion

The combination of MS based shot gun proteomics with the label free quantification approach, we used in this study, enables for semi-quantitative analysis of a large number of proteins released. This released fraction is less complex than the cellular proteome mentioned elsewhere yet is of high relevance to the studied bacterium.

The prerequisite for the application of LFQ approach is to analyse samples of comparable biological origin and composition. The growth of bacterial strains was comparable and the protein content of prepared triplicates did not differ in orders of magnitude.

The proteins involved in the assembly of *C. difficile* flagellum were present among other proteins revealing higher expression. The quantification of the proteins involved in the assembly of *C. difficile* flagellum pointed toward RTs 027 and 176, the main proteins constituting hook–basal body complex and the rotating filament were observed as overexpressed. The discriminatory proteins with expression characteristic only for RTs 027 and 176 involved FliE protein and glycosyltransferase C9YI34. FliE participates in the normal export of other substrates that play role in the assembly. However, a very low basal level of export function was previously described even in the absence of FliE. This argues against a vital role for FliE in export and proposes the primary role of FliE as a structural adapter between the annular symmetry of the membrane and supramembrane ring and the helical symmetry of the rod and all subsequent axial structures [92].

The earlier mentioned chaperones were relatively abundant in the representative samples. The FlgM works as an inhibitor of FliA, a flagellum-specific RNA polymerase responsible for flagellin transcription, forming a complex where FliA is readily displaced by

FliS as demonstrated in the study on *Salmonella enteric* serovar Typhimurium [93]. Therefore the ribotypes with observable FlgM have FliS2 in tow, although the levels in RT 078 are lower compared to other RTs. A similar role in flagellar assembly plays the aforementioned HtpG [94] that was observed in coinciding ribotypes as FlgM and FliS2.

The involvement of a homologue of glycosyltransferase in the glycosylation process C9YI34 (CD0240 in *C. difficile* 630) was proven. Inactivation of CD0240 led to a loss of the surface-associated flagellin protein and rendered the strain non-motile. However, the strain still produced truncated polymerized flagella filaments [65], [95]. In the study, this protein was also observed in RTs 014 and 005. However, the expression levels were lower. Thus, flagellin glycosylation was confirmed to be important in *C. difficile* flagellum assembly and virulence.

In addition, enzymes involved in rhamnose biosynthesis were observed. The importance of this flagellin glycosylation was demonstrated for *rmlB* in *Burkholderia cenocepacia* with consequent Human Innate Immune Responses [70]. The homologs involved in the same operon (RmlA and RmlC) [96] were also observable in our study.

Regarding RT 078, the absence of differential expression of most proteins involved in the assembly of *C. difficile* flagellum (with exception of FliK and FliD) could be explained by the previously published genomic study which confirmed the complete loss of the F3 flagellar region while retaining the F1 region (containing *fliK* and *fliD* genes). This has been corroborated using microarray data from phylogenetic studies [97], [98]. The low protein expression of the FliC, FliD and a putative glycosyltransferase (in comparison with RTs 027 and 176), is in agreement with the studies on non-flagellated *C. difficile* serotypes retaining transcription of *fliC* and *fliD* genes reporting the absence of its protein products [55], [66]. Yet, for *C. difficile* strain 630Δerm it was found that mutants in *fli*C and *fli*D adhered more strongly to Caco-2 cells than the wild-type [67], [68].

The question of a correlation between particular flagellins and toxin levels among RTs 027, 176, and 005 could be raised. However, the lower levels of TcdA and TcdB in RT 014 and the inability to detect the flagellar proteins by the mass spectrometry, propose a greater complexity of the *C. difficile* virulence factors. After identifying the key flagellar proteins in virulent strains the next logical step is to undertake studies with gene deletion mutants and test these for pathogenicity compared to the wild-type strain. A number

of studies are pursuing this notion with a knock-out of virulence related genes that can be used as a strategy to produce attenuated bacterial vaccines, for example.

Another avenue of interest is a change in protein expression following exposure to a known agent or stress. As the proteomic response of a pathogen to stress aims to counteract the effects of this stress, it can be expected that the pattern of a cell's responses to an agent will be dependent on its mechanism of action [62].

5 Conclusion

The study set out to monitor the relationship between secretion of the proteins associated with flagellar formation and function in *C. difficile* strains of variable virulence. The outcome of this work was different observations in strains with increased potential for virulence. However, with the present state of knowledge, this remains speculative and awaits experimental validation. The proteins associated with the bacterial cell surface or secreted into the environment represent important virulence factors and may also be directly involved in the recognition, targeting, and detoxification of antimicrobials. Hence, it may also be useful to supplement the available data sets in follow-up studies with comprehensive views on other subcellular proteome fractions than the soluble proteins in the cell.

The increasing incidence of *C. difficile* infection and the occurrence and spread of antibiotic-resistant strains emphasizes the need for detailed knowledge of the protein as a marker. The availability of protein signature libraries in *C. difficile* will not only enhance the knowledge of different mechanisms of action but will also provide a platform that supports the development and evaluation of new therapeutic agents.

6 Abbreviations

AAD	antibiotic associated diarrhoea
ABC	ammonium bicarbonate
AFLP	amplified fragment length polymorphism
AMP	antimicrobial peptide
BCA	bicinchoninic acid
BHI	brain heart infusion
C. difficile	*Clostridium difficile*
CDAD	*C. difficile* associated diarrhoea
CDI	*Clostridioides difficile* infection
CDT	binary toxin
CFU	colony forming units
CM	cytoplasmic membrane
CPD	cysteine protease domain
CWP	cell wall proteins
CYTA	cytotoxic assay test
DGDH	diarrhoeal glutamate dehydrogenase
EIA	enzyme immunoassay
ELISA	enzyme linked immunosorbent assays
ESCMID	European Society of Clinical Microbiology and Infectious Diseases
FASP	filter aided sample preparation
FDR	false discovery rate
FMT	faecal microbiota transplantation
GDH	diarrheal glutamate dehydrogenase
GI	gastrointestinal
GTD	glucosyltransferase domains
HCD	higher energy C-trap dissociation
HCD MS/MS	higher-energy C-trap dissociation mass spectrometry
HPLC	high performance liquid chromatography
IDSA	Infectious Diseases Society of America
LC-MS/MS	liquid chromatography with tandem mass spectrometry
LFQ	label-free quantification
MLST	multilocus sequence typing
MLVA	multilocus variable number tandem repeat analysis
MTZ	metronidazole
NAAT	nucleic acid amplification test
NAP1	North American pulsed-field type 1
OD	optical density
OM	outer membrane
PCR	polymerase chain reaction
PFGE	pulse field gel electrophoresis

PG	peptidoglycan layer
PSM	peptide-spectrum matches
RBD	receptor binding domain
REA	restriction endonuclease analysis
RNA	ribonucleic acid
SHEA	Society for Healthcare Epidemiology of America
SLP	surface layer protein
slpAST	surface layer protein A gene sequence typing
TCEPT	tris(2-carboxyethyl)phosphine
WCC	white cell count
WGS	whole genome sequencing

7 Literature

[1] J. G. Bartlett, 'Antibiotic-Associated Diarrhea', N. Engl. J. Med., vol. 346, no. 5, pp. 334–339, Jan. 2002, doi: 10.1056/NEJMcp011603.

[2] D. Schmid *et al.*, 'All-cause mortality in hospitalized patients with infectious diarrhea: *Clostridium difficile* versus other enteric pathogens in Austria from 2008 to 2010', J. Infect. Public Health, vol. 7, no. 2, pp. 133–144, Mar. 2014, doi: 10.1016/j.jiph.2013.07.010.

[3] S. M. Heimann, M. R. Cruz Aguilar, S. Mellinghof, and M. J. G. T. Vehreschild, 'Economic burden and cost-effective management of *Clostridium difficile* infections', Médecine Mal. Infect., vol. 48, no. 1, pp. 23–29, Feb. 2018, doi: 10.1016/j.medmal.2017.10.010.

[4] C. Slimings and T. V. Riley, 'Antibiotics and hospital-acquired *Clostridium difficile* infection: update of systematic review and meta-analysis', J. Antimicrob. Chemother., vol. 69, no. 4, pp. 881–891, Apr. 2014, doi: 10.1093/jac/dkt477.

[5] R. J. Melander, D. V. Zurawski, and C. Melander, 'Narrow-spectrum antibacterial agents', MedChemComm, vol. 9, no. 1, pp. 12–21, 2018, doi: 10.1039/C7MD00528H.

[6] L. C. McDonald *et al.*, 'An Epidemic, Toxin Gene–Variant Strain of *Clostridium difficile*', N. Engl. J. Med., vol. 353, no. 23, pp. 2433–2441, Dec. 2005, doi: 10.1056/NEJMoa051590.

[7] G. Killgore *et al.*, 'Comparison of Seven Techniques for Typing International Epidemic Strains of *Clostridium difficile* : Restriction Endonuclease Analysis, Pulsed-Field Gel Electrophoresis, PCR-Ribotyping, Multilocus Sequence Typing, Multilocus Variable-Number Tandem-Repeat Analysis, Amplified Fragment Length Polymorphism, and Surface Layer Protein A Gene Sequence Typing', J. Clin. Microbiol., vol. 46, no. 2, pp. 431–437, Feb. 2008, doi: 10.1128/JCM.01484-07.

[8] P. Bidet, F. Barbut, V. Lalande, B. Burghoffer, and J.-C. Petit, 'Development of a new PCR-ribotyping method for *Clostridium difficile* based on ribosomal RNA gene sequencing', FEMS Microbiol. Lett., vol. 175, no. 2, pp. 261–266, Jun. 1999, doi: 10.1111/j.1574-6968.1999.tb13629.x.

[9] C. H. Chilton *et al.*, 'Comparative proteomic analysis of *Clostridium difficile* isolates of varying virulence', J. Med. Microbiol., vol. 63, no. Pt_4, pp. 489–503, Apr. 2014, doi: 10.1099/jmm.0.070409-0.

[10] A. Wright *et al.*, 'Proteomic analysis of cell surface proteins from *Clostridium difficile*', Proteomics, vol. 5, no. 9, pp. 2443–2452, Jun. 2005, doi: 10.1002/pmic.200401179.

[11] T. D. Lawley *et al.*, 'Proteomic and Genomic Characterization of Highly Infectious *Clostridium difficile* 630 Spores', J. Bacteriol., vol. 191, no. 17, pp. 5377–5386, Sep. 2009, doi: 10.1128/JB.00597-09.

[12] A. Boetzkes, K. W. Felkel, J. Zeiser, N. Jochim, I. Just, and A. Pich, 'Secretome analysis of *Clostridium difficile* strains', Arch. Microbiol., vol. 194, no. 8, pp. 675–687, Aug. 2012, doi: 10.1007/s00203-012-0802-5.

[13] J. Dresler *et al.*, 'Analysis of proteomes released from in vitro cultured eight *Clostridium difficile* PCR ribotypes revealed specific expression in PCR ribotypes 027 and 176 confirming their genetic relatedness and clinical importance at the proteomic level', Gut Pathog., vol. 9, p. 45, 2017, doi: 10.1186/s13099-017-0194-9.

[14] M. L. Hutton, K. E. Mackin, A. Chakravorty, and D. Lyras, 'Small animal models for the study of *Clostridium difficile* disease pathogenesis', FEMS Microbiol. Lett., vol. 352, no. 2, pp. 140–149, Mar. 2014, doi: 10.1111/1574-6968.12367.

[15] M. M. Awad, P. A. Johanesen, G. P. Carter, E. Rose, and D. Lyras, '*Clostridium difficile* virulence factors: Insights into an anaerobic spore-forming pathogen', Gut Microbes, vol. 5, no. 5, pp. 579–593, Sep. 2014, doi: 10.4161/19490976.2014.969632.
[16] S. P. Borriello, 'Pathogenesis of *Clostridium difficile* infection', J. Antimicrob. Chemother., vol. 41, no. suppl 3, pp. 13–19, May 1998, doi: 10.1093/jac/41.suppl_3.13.
[17] P. A. Lawson and F. A. Rainey, 'Proposal to restrict the genus *Clostridium Prazmowski to Clostridium butyricum* and related species', Int. J. Syst. Evol. Microbiol., vol. 66, no. 2, pp. 1009–1016, Feb. 2016, doi: 10.1099/ijsem.0.000824.
[18] H. E. Larson, A. B. Price, P. Honour, and S. P. Borriello, '*Clostridium difficile* and the aetiology of pseudomembranous colitis', Lancet Lond. Engl., vol. 1, no. 8073, pp. 1063–1066, May 1978, doi: 10.1016/s0140-6736(78)90912-1.
[19] A. M. Seekatz and V. B. Young, '*Clostridium difficile* and the microbiota', J. Clin. Invest., vol. 124, no. 10, pp. 4182–4189, Oct. 2014, doi: 10.1172/JCI72336.
[20] J. S. Weese, '*Clostridium difficile* in food–-innocent bystander or serious threat?', Clin. Microbiol. Infect., vol. 16, no. 1, pp. 3–10, Jan. 2010, doi: 10.1111/j.1469-0691.2009.03108.x.
[21] M. Makita and W. W. Wells, 'Quantitative analysis of fecal bile acids by gas-liquid chromatography', Anal. Biochem., vol. 5, no. 6, pp. 523–530, Jun. 1963, doi: 10.1016/0003-2697(63)90072-1.
[22] L. A. Thomas, 'Bile acid metabolism by fresh human colonic contents: a comparison of caecal versus faecal samples', Gut, vol. 49, no. 6, pp. 835–842, Dec. 2001, doi: 10.1136/gut.49.6.835.
[23] J. B. Carey and C. J. Watson, 'Isolation of deoxycholic acid from normal human feces', J. Biol. Chem., vol. 216, no. 2, pp. 847–850, Oct. 1955.
[24] A. Shen, '*Clostridium difficile* Toxins: Mediators of Inflammation', J. Innate Immun., vol. 4, no. 2, pp. 149–158, 2012, doi: 10.1159/000332946.
[25] C. Tsigrelis, 'Recurrent *Clostridioides difficile* infection: Recognition, management, prevention', Cleve. Clin. J. Med., vol. 87, no. 6, pp. 347–359, Jun. 2020, doi: 10.3949/ccjm.87gr.20001.
[26] S. B. Debast, M. P. Bauer, E. J. Kuijper, and European Society of Clinical Microbiology and Infectious Diseases, 'European Society of Clinical Microbiology and Infectious Diseases: update of the treatment guidance document for *Clostridium difficile* infection', Clin. Microbiol. Infect. Off. Publ. Eur. Soc. Clin. Microbiol. Infect. Dis., vol. 20 Suppl 2, pp. 1–26, Mar. 2014, doi: 10.1111/1469-0691.12418.
[27] S. Johnson *et al.*, 'Vancomycin, Metronidazole, or Tolevamer for *Clostridium difficile* Infection: Results From Two Multinational, Randomized, Controlled Trials', Clin. Infect. Dis., vol. 59, no. 3, pp. 345–354, Aug. 2014, doi: 10.1093/cid/ciu313.
[28] S. Esposito, G. Umbrello, L. Castellazzi, and N. Principi, 'Treatment of *Clostridium difficile* infection in pediatric patients', Expert Rev. Gastroenterol. Hepatol., vol. 9, no. 6, pp. 747–755, Jun. 2015, doi: 10.1586/17474124.2015.1039988.
[29] '*Clostridioides difficile* infection: antimicrobial prescribing, NICE guideline [NG199] Published: 23 July 2021'.
[30] L. C. McDonald *et al.*, '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)', Clin. Infect. Dis., vol. 66, no. 7, pp. e1–e48, Mar. 2018, doi: 10.1093/cid/cix1085.
[31] M. H. Wilcox *et al.*, 'Bezlotoxumab for Prevention of Recurrent *Clostridium difficile* Infection', N. Engl. J. Med., vol. 376, no. 4, pp. 305–317, Jan. 2017, doi: 10.1056/NEJMoa1602615.

[32] M. Hocquart *et al.*, 'Early Fecal Microbiota Transplantation Improves Survival in Severe *Clostridium difficile* Infections', Clin. Infect. Dis., vol. 66, no. 5, pp. 645–650, Feb. 2018, doi: 10.1093/cid/cix762.
[33] S. S. Hota *et al.*, 'Oral Vancomycin Followed by Fecal Transplantation Versus Tapering Oral Vancomycin Treatment for Recurrent *Clostridium difficile* Infection: An Open-Label, Randomized Controlled Trial', Clin. Infect. Dis., vol. 64, no. 3, pp. 265–271, Feb. 2017, doi: 10.1093/cid/ciw731.
[34] J. Freeman, 'Antibiotic activity against genotypically distinct and indistinguishable *Clostridium difficile* isolates', J. Antimicrob. Chemother., vol. 47, no. 2, pp. 244–246, Feb. 2001, doi: 10.1093/jac/47.2.244.
[35] J. E. Bandow, H. Brötz, L. I. O. Leichert, H. Labischinski, and M. Hecker, 'Proteomic Approach to Understanding Antibiotic Action', Antimicrob. Agents Chemother., vol. 47, no. 3, pp. 948–955, Mar. 2003, doi: 10.1128/AAC.47.3.948-955.2003.
[36] M. Hecker, H. Antelmann, K. Büttner, and J. Bernhardt, 'Gel-based proteomics of Gram-positive bacteria: A powerful tool to address physiological questions', *PROTEOMICS*, vol. 8, no. 23–24, pp. 4958–4975, Dec. 2008, doi: 10.1002/pmic.200800278.
[37] G. P. Carter, J. I. Rood, and D. Lyras, 'The role of toxin A and toxin B in the virulence of *Clostridium difficile*', Trends Microbiol., vol. 20, no. 1, pp. 21–29, Jan. 2012, doi: 10.1016/j.tim.2011.11.003.
[38] D. E. Voth and J. D. Ballard, '*Clostridium difficile* Toxins: Mechanism of Action and Role in Disease', Clin. Microbiol. Rev., vol. 18, no. 2, pp. 247–263, Apr. 2005, doi: 10.1128/CMR.18.2.247-263.2005.
[39] B. Geric *et al.*, 'Binary Toxin–Producing, Large Clostridial Toxin–Negative *Clostridium difficile* Strains Are Enterotoxic but Do Not Cause Disease in Hamsters', J. Infect. Dis., vol. 193, no. 8, pp. 1143–1150, Apr. 2006, doi: 10.1086/501368.
[40] J. F. Rodriguez and M. Esteban, 'Mapping and nucleotide sequence of the vaccinia virus gene that encodes a 14-kilodalton fusion protein', J. Virol., vol. 61, no. 11, pp. 3550–3554, Nov. 1987.
[41] M. DELMeE and V. Avesani, 'Virulence of ten serogroups of *Clostridium difficile* in hamsters', J. Med. Microbiol., vol. 33, no. 2, pp. 85–90, Oct. 1990, doi: 10.1099/00222615-33-2-85.
[42] A. S. Paparella, B. L. Aboulache, R. K. Harijan, K. S. Potts, P. C. Tyler, and V. L. Schramm, 'Inhibition of *Clostridium difficile* TcdA and TcdB toxins with transition state analogues', Nat. Commun., vol. 12, no. 1, p. 6285, Dec. 2021, doi: 10.1038/s41467-021-26580-6.
[43] V. G. Loo *et al.*, 'A Predominantly Clonal Multi-Institutional Outbreak of *Clostridium difficile* –Associated Diarrhea with High Morbidity and Mortality', N. Engl. J. Med., vol. 353, no. 23, pp. 2442–2449, Dec. 2005, doi: 10.1056/NEJMoa051639.
[44] C. Schwan *et al.*, '*Clostridium difficile* Toxin CDT Induces Formation of Microtubule-Based Protrusions and Increases Adherence of Bacteria', PLoS Pathog., vol. 5, no. 10, p. e1000626, Oct. 2009, doi: 10.1371/journal.ppat.1000626.
[45] C. Schwan *et al.*, '*Clostridium difficile* toxin CDT hijacks microtubule organization and reroutes vesicle traffic to increase pathogen adherence', Proc. Natl. Acad. Sci., vol. 111, no. 6, pp. 2313–2318, Feb. 2014, doi: 10.1073/pnas.1311589111.
[46] A. O. Henriques and C. P. Moran, Jr., 'Structure, Assembly, and Function of the Spore Surface Layers', Annu. Rev. Microbiol., vol. 61, no. 1, pp. 555–588, Oct. 2007, doi: 10.1146/annurev.micro.61.080706.093224.
[47] D. L. Popham, J. Helin, C. E. Costello, and P. Setlow, 'Muramic lactam in peptidoglycan of *Bacillus subtilis* spores is required for spore outgrowth but not for

spore dehydration or heat resistance', Proc. Natl. Acad. Sci., vol. 93, no. 26, pp. 15405–15410, Dec. 1996, doi: 10.1073/pnas.93.26.15405.
[48] T. A. Leski *et al.*, 'Identification and Classification of *bcl* Genes and Proteins of *Bacillus cereus* Group Organisms and Their Application in *Bacillus anthracis* Detection and Fingerprinting', Appl. Environ. Microbiol., vol. 75, no. 22, pp. 7163–7172, Nov. 2009, doi: 10.1128/AEM.01069-09.
[49] W. Hodgkiss, Z. J. Ordal, and D. C. Cann, 'The Morphology and Ultrastructure of the Spore and Exosporium of Some *Clostridium* Species', J. Gen. Microbiol., vol. 47, no. 2, pp. 213–225, May 1967, doi: 10.1099/00221287-47-2-213.
[50] Y. Kevorkian, D. J. Shirley, and A. Shen, 'Regulation of *Clostridium difficile* spore germination by the CspA pseudoprotease domain', Biochimie, vol. 122, pp. 243–254, Mar. 2016, doi: 10.1016/j.biochi.2015.07.023.
[51] D. Zhu, J. A. Sorg, and X. Sun, '*Clostridioides difficile* Biology: Sporulation, Germination, and Corresponding Therapies for C. difficile Infection', Front. Cell. Infect. Microbiol., vol. 8, p. 29, Feb. 2018, doi: 10.3389/fcimb.2018.00029.
[52] W. R. Abhyankar, L. Zheng, S. Brul, C. G. de Koster, and L. J. de Koning, 'Vegetative Cell and Spore Proteomes of *Clostridioides difficile* Show Finite Differences and Reveal Potential Protein Markers', J. Proteome Res., vol. 18, no. 11, pp. 3967–3976, Nov. 2019, doi: 10.1021/acs.jproteome.9b00413.
[53] D. Paredes-Sabja and M. R. Sarker, 'Adherence of *Clostridium difficile* spores to Caco-2 cells in culture', J. Med. Microbiol., vol. 61, no. 9, pp. 1208–1218, Sep. 2012, doi: 10.1099/jmm.0.043687-0.
[54] L. T. Joshi, D. S. Phillips, C. F. Williams, A. Alyousef, and L. Baillie, 'Contribution of Spores to the Ability of *Clostridium difficile* To Adhere to Surfaces', Appl. Environ. Microbiol., vol. 78, no. 21, pp. 7671–7679, Nov. 2012, doi: 10.1128/AEM.01862-12.
[55] A. Tasteyre *et al.*, 'Molecular characterization of fliD gene encoding flagellar cap and its expression among *Clostridium difficile* isolates from different serogroups', J. Clin. Microbiol., vol. 39, no. 3, pp. 1178–1183, Mar. 2001, doi: 10.1128/JCM.39.3.1178-1183.2001.
[56] T. Thapa *et al.*, 'Multiple Factors Modulate Biofilm Formation by the Anaerobic Pathogen *Clostridium difficil*e', J. Bacteriol., vol. 195, no. 3, pp. 545–555, Feb. 2013, doi: 10.1128/JB.01980-12.
[57] D. C. Dailey, A. Kaiser, and R. H. Schloemer, 'Factors influencing the phagocytosis of *Clostridium difficile* by human polymorphonuclear leukocytes', Infect. Immun., vol. 55, no. 7, pp. 1541–1546, Jul. 1987, doi: 10.1128/iai.55.7.1541-1546.1987.
[58] E. Calabi *et al.*, 'Molecular characterization of the surface layer proteins from *Clostridium difficile*', *Mol. Microbiol.,* vol. 40, no. 5, pp. 1187–1199, Jun. 2001, doi: 10.1046/j.1365-2958.2001.02461.x.
[59] C. Li *et al.*, 'Characterization of the virulence of a non-RT027, non-RT078 and binary toxin-positive *Clostridium difficile* strain associated with severe diarrhea', Emerg. Microbes Infect., vol. 7, no. 1, pp. 1–11, Dec. 2018, doi: 10.1038/s41426-018-0211-1.
[60] M. Barer and W. L. Irving, Eds., Medical microbiology: a guide to microbial infections: pathogenesis, immunity, laboratory investigation and control, Nineteenth Edition. Amsterdam, Netherlands: Elsevier, 2018.
[61] Y. Morimoto and T. Minamino, 'Structure and Function of the Bi-Directional Bacterial Flagellar Motor', *Biomolecules,* vol. 4, no. 1, pp. 217–234, Feb. 2014, doi: 10.3390/biom4010217.
[62] S. Maaß *et al.*, 'Proteomic Adaptation of *Clostridioides difficile* to Treatment with the Antimicrobial Peptide Nisin', Cells, vol. 10, no. 2, p. 372, Feb. 2021, doi: 10.3390/cells10020372.

[63] K. B. Barken *et al.*, 'Roles of type IV pili, flagellum-mediated motility and extracellular DNA in the formation of mature multicellular structures in *Pseudomonas aeruginosa* biofilms', Environ. Microbiol., vol. 10, no. 9, pp. 2331–2343, Sep. 2008, doi: 10.1111/j.1462-2920.2008.01658.x.
[64] A. Aubry *et al.*, 'Modulation of Toxin Production by the Flagellar Regulon in *Clostridium difficile'*, Infect. Immun., vol. 80, no. 10, pp. 3521–3532, Oct. 2012, doi: 10.1128/IAI.00224-12.
[65] S. M. Twine *et al.*, 'Motility and flagellar glycosylation in *Clostridium difficile*', J. Bacteriol., vol. 191, no. 22, pp. 7050–7062, Nov. 2009, doi: 10.1128/JB.00861-09.
[66] A. Tasteyre *et al.*, 'Phenotypic and genotypic diversity of the flagellin gene (fliC) among *Clostridium difficile* isolates from different serogroups', J. Clin. Microbiol., vol. 38, no. 9, pp. 3179–3186, Sep. 2000.
[67] T. C. Dingle, G. L. Mulvey, and G. D. Armstrong, 'Mutagenic Analysis of the *Clostridium difficile* Flagellar Proteins, FliC and FliD, and Their Contribution to Virulence in Hamsters', Infect. Immun., vol. 79, no. 10, pp. 4061–4067, Oct. 2011, doi: 10.1128/IAI.05305-11.
[68] S. T. Baban *et al.*, 'The role of flagella in *Clostridium difficile* pathogenesis: comparison between a non-epidemic and an epidemic strain', PloS One, vol. 8, no. 9, p. e73026, 2013, doi: 10.1371/journal.pone.0073026.
[69] S. Xu *et al.*, 'FliS modulates FlgM activity by acting as a non-canonical chaperone to control late flagellar gene expression, motility and biofilm formation in *Y ersinia pseudotuberculosis*: FliS escort FlgM as a non-canonical chaperone', Environ. Microbiol., vol. 16, no. 4, pp. 1090–1104, Apr. 2014, doi: 10.1111/1462-2920.12222.
[70] A. Hanuszkiewicz *et al.*, 'Identification of the Flagellin Glycosylation System in *Burkholderia cenocepacia* and the Contribution of Glycosylated Flagellin to Evasion of Human Innate Immune Responses', J. Biol. Chem., vol. 289, no. 27, pp. 19231–19244, Jul. 2014, doi: 10.1074/jbc.M114.562603.
[71] 'Davies KA, Ashwin H, Longshaw CM, Burns DA, Davis GL, Wilcox MH, on behalf of the EUCLID study group. Diversity of *Clostridium difficile* PCR ribotypes in Europe: results from the European, multicentre, prospective, biannual, point-prevalence study of *Clostridium difficile* infection in hospitalised patients with diarrhoea (EUCLID), 2012 and 2013. Euro Surveill 2016;21(29):pii=30294.'.
[72] 'European Centre for Disease Prevention and Control. *Clostridium difficile* infections. In: ECDC. Annual epidemiological report for 2016. Stockholm: ECDC; 2018.'
[73] A. Sharma, S. Lee, and Y.-S. Park, 'Molecular typing tools for identifying and characterizing lactic acid bacteria: a review', Food Sci. Biotechnol., vol. 29, no. 10, pp. 1301–1318, Oct. 2020, doi: 10.1007/s10068-020-00802-x.
[74] M. Krutova, O. Nyc, J. Matejkova, F. Allerberger, M. H. Wilcox, and E. J. Kuijper, 'Molecular characterisation of Czech *Clostridium difficile* isolates collected in 2013–2015', Int. J. Med. Microbiol., Aug. 2016, doi: 10.1016/j.ijmm.2016.07.003.
[75] S. L. Stubbs, J. S. Brazier, G. L. O'Neill, and B. I. Duerden, 'PCR targeted to the 16S-23S rRNA gene intergenic spacer region of *Clostridium difficile* and construction of a library consisting of 116 different PCR ribotypes', J. Clin. Microbiol., vol. 37, no. 2, pp. 461–463, Feb. 1999.
[76] H. M. B. Seth-Smith *et al.*, 'Transition From PCR-Ribotyping to Whole Genome Sequencing Based Typing of *Clostridioides difficile*', Front. Cell. Infect. Microbiol., vol. 11, p. 681518, Jun. 2021, doi: 10.3389/fcimb.2021.681518.
[77] J. Cho *et al.*, '*Clostridioides difficile* Whole-genome Sequencing Differentiates Relapse With the Same Strain From Reinfection With a New Strain', Clin. Infect. *Dis.*, vol. 72, no. 5, pp. 806–813, Mar. 2021, doi: 10.1093/cid/ciaa159.

[78] R. Aebersold and M. Mann, ‘Mass spectrometry-based proteomics’, Nature, vol. 422, no. 6928, pp. 198–207, Mar. 2003, doi: 10.1038/nature01511.
[79] S. Garbis, G. Lubec, and M. Fountoulakis, ‘Limitations of current proteomics technologies’, J. Chromatogr. A, vol. 1077, no. 1, pp. 1–18, Jun. 2005, doi: 10.1016/j.chroma.2005.04.059.
[80] E. W. Deutsch, H. Lam, and R. Aebersold, ‘Data analysis and bioinformatics tools for tandem mass spectrometry in proteomics’, Physiol. Genomics, vol. 33, no. 1, pp. 18–25, Mar. 2008, doi: 10.1152/physiolgenomics.00298.2007.
[81] C. A. Luber *et al.*, ‘Quantitative Proteomics Reveals Subset-Specific Viral Recognition in Dendritic Cells’, Immunity, vol. 32, no. 2, pp. 279–289, Feb. 2010, doi: 10.1016/j.immuni.2010.01.013.
[82] ‘Zentrum für Molekulare Biologie der Universität Heidelberg website, accessed on 15.3.2022’.
[83] B. Elliott, K. E. Dingle, X. Didelot, D. W. Crook, and T. V. Riley, ‘The Complexity and Diversity of the Pathogenicity Locus in *Clostridium difficile* Clade 5’, Genome Biol. Evol., vol. 6, no. 12, pp. 3159–3170, Dec. 2014, doi: 10.1093/gbe/evu248.
[84] K. Rao *et al.*, ‘*Clostridium difficile* Ribotype 027: Relationship to Age, Detectability of Toxins A or B in Stool With Rapid Testing, Severe Infection, and Mortality’, Clin. Infect. Dis., vol. 61, no. 2, pp. 233–241, Jul. 2015, doi: 10.1093/cid/civ254.
[85] P. K. Smith *et al.*, ‘Measurement of protein using bicinchoninic acid’, Anal. Biochem., vol. 150, no. 1, pp. 76–85, Oct. 1985.
[86] J. Cox, M. Y. Hein, C. A. Luber, I. Paron, N. Nagaraj, and M. Mann, ‘Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ’, Mol. Cell. Proteomics MCP, vol. 13, no. 9, pp. 2513–2526, Sep. 2014, doi: 10.1074/mcp.M113.031591.
[87] J. Cox, N. Neuhauser, A. Michalski, R. A. Scheltema, J. V. Olsen, and M. Mann, ‘Andromeda: a peptide search engine integrated into the MaxQuant environment’, J. Proteome Res., vol. 10, no. 4, pp. 1794–1805, Apr. 2011, doi: 10.1021/pr101065j.
[88] D. W. Huang, B. T. Sherman, and R. A. Lempicki, ‘Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources’, Nat. Protoc., vol. 4, no. 1, pp. 44–57, 2009, doi: 10.1038/nprot.2008.211.
[89] T. N. Petersen, S. Brunak, G. von Heijne, and H. Nielsen, ‘SignalP 4.0: discriminating signal peptides from transmembrane regions’, Nat. Methods, vol. 8, no. 10, pp. 785–786, 2011, doi: 10.1038/nmeth.1701.
[90] J. D. Bendtsen, L. Kiemer, A. Fausbøll, and S. Brunak, ‘Non-classical protein secretion in bacteria’, BMC Microbiol., vol. 5, p. 58, 2005, doi: 10.1186/1471-2180-5-58.
[91] U. Jenal, J. White, and L. Shapiro, ‘Caulobacter Flagellar Function, but not Assembly, Requires FliL, a Non-polarly Localized Membrane Protein Present in all Cell Types’, J. Mol. Biol., vol. 243, no. 2, pp. 227–244, Oct. 1994, doi: 10.1006/jmbi.1994.1650.
[92] T. Minamino, S. Yamaguchi, and R. M. Macnab, ‘Interaction between FliE and FlgB, a proximal rod component of the flagellar basal body of *Salmonella*’, J. Bacteriol., vol. 182, no. 11, pp. 3029–3036, Jun. 2000.
[93] A. Galeva, N. Moroz, Y.-H. Yoon, K. T. Hughes, F. A. Samatey, and A. S. Kostyukova, ‘Bacterial Flagellin-Specific Chaperone FliS Interacts with Anti-Sigma Factor FlgM’, J. Bacteriol., vol. 196, no. 6, pp. 1215–1221, Mar. 2014, doi: 10.1128/JB.01278-13.
[94] T. Dong *et al.*, ‘Involvement of the Heat Shock Protein HtpG of *Salmonella Typhimurium* in Infection and Proliferation in Hosts’, Front. Cell. Infect. Microbiol., vol. 11, p. 758898, Nov. 2021, doi: 10.3389/fcimb.2021.758898.

[95] A. Faulds-Pain *et al.*, 'The post-translational modification of the *Clostridium difficile* flagellin affects motility, cell surface properties and virulence', Mol. Microbiol., vol. 94, no. 2, pp. 272–289, Oct. 2014, doi: 10.1111/mmi.12755.
[96] M. Graninger, B. Kneidinger, K. Bruno, A. Scheberl, and P. Messner, 'Homologs of the Rml Enzymes from *Salmonella enterica* Are Responsible for dTDP-β- L -Rhamnose Biosynthesis in the Gram-Positive Thermophile *Aneurinibacillus thermoaerophilus* DSM 10155', Appl. Environ. Microbiol., vol. 68, no. 8, pp. 3708–3715, Aug. 2002, doi: 10.1128/AEM.68.8.3708-3715.2002.
[97] R. A. Stabler *et al.*, 'Comparative genome and phenotypic analysis of *Clostridium difficile* 027 strains provides insight into the evolution of a hypervirulent bacterium', Genome Biol., vol. 10, no. 9, p. R102, 2009, doi: 10.1186/gb-2009-10-9-r102.
[98] R. A. Stabler *et al.*, 'Comparative phylogenomics of *Clostridium difficile* reveals clade specificity and microevolution of hypervirulent strains', J. Bacteriol., vol. 188, no. 20, pp. 7297–7305, Oct. 2006, doi: 10.1128/JB.00664-06.