New tools and insights in physiology and chromosome dynamics of Clostridioides difficile

Oliveira Paiva, A.M.

CHAPTER 7

General Discussion

General Discussion

Clostridioides difficile is one of the major causes of antibiotic-associated diarrhoea [1,2]. Despite being discovered in 1935 [3], the bacterium only gained significant interest with the emergence of outbreak-related strains worldwide [4-6]. The *C. difficile* strains associated with the outbreaks showed increased severity of disease and mortality, making *C. difficile* an important threat for healthcare [7-9]. Over the years, much *C. difficile* research has focused on the main virulence factors that contribute to the development of CDI, such as the production and mechanism of action of the cytotoxic toxins that damage the epithelial cells, or the ability to sporulate, which is crucial for the transmission of *C. difficile* [1]. Recently, the interest in other cellular pathways has increased, not only to understand cell physiology but also to search for new targets for therapeutics, such as DNA replication [10-12]. These developments form the background of this PhD thesis.

In this thesis, we aimed to characterize different aspects of *C. difficile* biology, in order to understand cellular mechanisms, such as the initiation of replication and the chromosome maintenance, but also the controversial role of TcdC in toxin regulation. Along the way, we developed new tools, several of which are based on luminescence, that will facilitate future *C. difficile* research.

C. difficile new tools aid on new challenges

Advances in genetic and molecular tools have been essential to the understanding of *C. difficile* cell biology. Recently, different approaches for the genetic manipulation of *C. difficile* have been developed that allow the use of chromogenic and fluorescent reporters to study gene expression and *C. difficile* cell components.

In *C. difficile*, the β-glucuronidase (GusA) and the alkaline phosphatase (PhoZ) chromogenic reporters have been used [13-15]. These reporters allowed to study the regulation of toxin gene expression, by RstA or CodY for instance [13,16,17]. PhoZ colour development allows the use of this reporter not only in liquid media but also in a plate-based assay, where it can be detected by eye [15]. Even though GusA and PhoZ reporters can provide qualitative or quantitative analysis of the gene expression in *C. difficile*, different drawbacks associated with the reporters have limited their use. The use of the chromogenic reporters is affected by the level of expression, as a low level or heterogenous expression can impair the quantification and visualization of the reporters in the bacterial cells [14,15]. The main limitation of the reporters is the requirement on lysed cells for a quantitative measure of reporter activity [14,15]. This can lead to artefacts, due to inefficient cell lysis, for instance [14,15]. This presented an interesting

opportunity for the development of secreted reporters that could bypass the hurdle of *C. difficile* lysis.

For the development of novel secreted reporters in *C. difficile* we took advantage of the signal sequence of the PPEP-1 protein, which is highly secreted by *C. difficile* [18-20]. In **Chapter 2** we use the PPEP-1 signal sequence and used a synthetic biology approach to engineer two novel secreted reporters, AmyE^{opt} and sLuc^{opt}, both of which were codon-optimized for *C. difficile*.

Our laboratory strain *C. difficile* 630Δ*erm* is not able to degrade starch, which enabled us to easily assess the PPEP-1-mediated secretion of the AmyE^{opt} reporter and successfully validate the constitutive expression of the *B. subtilis* promoter P_{veg}. Amylase activity was observed both in liquid medium and in a plate-based assay, and although quantitative analysis of the activity was possible, it did not allow *C. difficile* growth with starch as sole carbon source. Additionally, not all *C. difficile* strains have the same metabolic capacity and some strains may have the ability to digest starch, potentially limiting the use of the AmyE^{opt} reporter.

In parallel with the AmyE^{opt} reporter, we fused the PPEP-1 signal sequence to a codon-optimized luciferase, based on the NanoLuc luciferase (Promega) [21]. NanoLuc is a 19 kDa luciferase that requires furimazine for light emission [21-23]. The high sensitivity of the luciferase and the linear dependency on the concentration of the luciferase substrate made this luciferase reporter a promising candidate for gene expression analysis in *C. difficile*. With a synthetic biology approach we were able to engineer a small secreted codon-optimized luciferase, which we named sLuc^{opt}. This reporter was efficiently expressed and secreted into the environment, allowing for easy assessment of gene expression, by sampling the culture medium. The efficient secretion of sLuc^{opt} minimizes artefacts due to the physiological burden imposed by the expression of the reporter, allowing the study of gene expression close to the *in vivo* situation. Despite the high sensitivity of the system, low levels of expression can be diluted beyond the limit of detection in the extracellular environment; in such cases, sampling of an intracellular reporter, such as the non-secreted Luc^{opt}, is advised.

Recently sLuc^{opt} was used in a plate-based-method to study the gene expression of σ^B-dependent promoters [24]. However, this use does not allow for the quantification of the sLuc^{opt} activity and can potentially present artefacts due to exposure to the environmental oxygen (required for the luciferase activity). Nevertheless, it poses a quick approach to assess gene expression on solid media.

sLuc^{opt} is highly stable and can retain activity in the culture medium or extracellular environment for prolonged periods of time. This stability and subsequent accumulation in the

environment may make the analysis of gene expression dynamics more challenging as downregulation of gene expression may be masked by earlier high-level expression.

The high sensitivity and easy detection of the luciferase reporter represented a promising tool that could be adapted for a wide range of other approaches in *C. difficile* (Fig. 1). We, therefore, extended the use of this luciferase reporter as a tool to study *C. difficile* protein localization and protein-protein interactions *in vivo* in the course of our investigations. The entire luciferase reporter toolkit was made available to the *C. difficile* research community through a non-profit repository (Addgene).

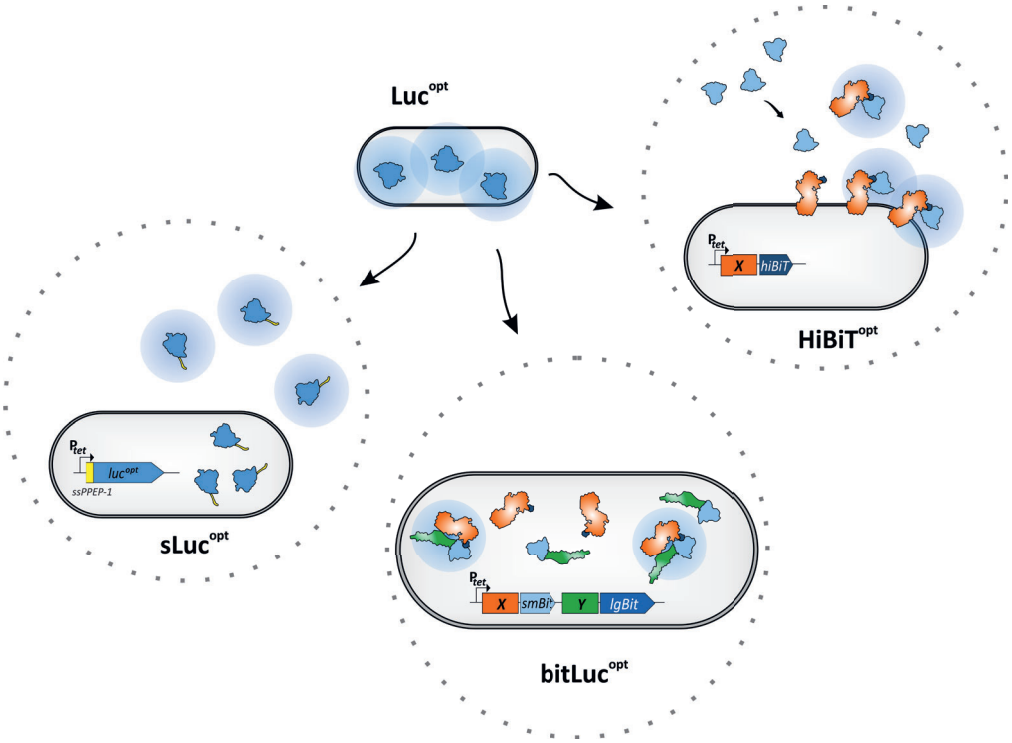

Fig. 1 – Schematic representation of the derived systems from the Luc^{opt} luciferase for *C. difficile* studies. The secreted luciferase reporter system sLuc^{opt} allows for the extracellular detection of the luciferase, enabling quick and easy detection for gene expression studies. The bitLuc^{opt} system provides a valuable tool for protein-protein studies *in vivo* by tagging the proteins of interest (X,Y) with SmBit and LgBit. The HiBiT^{opt} system allows the study of secreted proteins or complexes present at the cell surface.

We implemented a system for the extracellular detection of proteins, HiBiT^{opt} (Fig. 1), which allowed the determination of the localization of the C-terminus of the TcdC protein, in **Chapter 3** (discussed in more detail later on). This complementation-based system allows the detection of extracellularly localized C-termini of proteins or secreted proteins-of-interest, by tagging them with the 11-amino acid tag HiBiT^{opt}. The HiBiT^{opt} will interact with high affinity

to the added luciferase larger subunit, the LgBiT, which can not enter the bacterial cell. When both sub-units interact and in combination with the furimazine substrates it emits light. Likewise, it can also allow the easy quantification of the protein expression by in-gel detection, as demonstrated for TcdC and other proteins (**Chapter 3**).

The small size of the HiBiT^{opt} tag is an advantage, as it might reduce artefacts compared to tagging of proteins with higher molecular weight tags, but as we only explored C-terminal tagging, further studies are required to determine whether the system can also be used for N-terminal fusions. The HiBiT^{opt} system not only presents an interesting tool for topology determination but can potentially also be used for studies on protein secretion, membrane association and inhibitors of these processes. Recently, split-Nanoluc was used to monitor the translocation dynamics of proteins across the cell membrane, such as through the ubiquitous transporter Sec system in *E. coli* [25,26].

Whereas HiBiT^{opt} is engineered to have high affinity for the complementary subunit LgBiT, SmBiT has been engineered to have low affinity for the same protein while retaining the ability to complement activity of the luciferase. We used this characteristic in **Chapter 5** to develop a method for luminescent detection of *in vivo* protein-protein interactions in *C. difficile*, with the bitLuc^{opt} assay. Several complementation assays are used for the *in vivo* detection of interacting partners, such as the bacterial or yeast two-hybrid systems, but these systems generally do not perform very well for *C. difficile* due to distinct codon bias. The split-SNAP complementation assay has successfully been applied in *C. difficile* [27]. The split-SNAPtag subunits (11.1 kDa nSNAP and 11.8 kDa cSNAP) can interact to form a functional SNAPtag, when in close proximity in the cell [28]. This is similar to the bitLuc^{opt} system, where the proteins of interest are tagged with different luciferase subunits (SmBit and LgBiT) that can reconstitute luciferase, resulting in luminescent signal. However, though in total the two systems are comparable in size, the SmBit subunit is much smaller (1.3 kDa) [29] than nSNAP and cSNAP, and may therefore lead to less tag-dependent artefacts in protein function.

The developed luciferase systems contribute substantially to the available *C. difficile* toolkit and allowed us to address several questions related to toxin gene expression and chromosome dynamics in *C. difficile*. We anticipate that similar analyses can not only be performed to characterize the interactions of known *C. difficile* proteins within a complex machinery, such as the replisome, but also to identify unknown interactors.

Despite the usefulness of the luciferase resulting from minimal cell manipulation, technical limitations prevent the use of luciferase in more in-depth approaches, such as the analysis of

the subcellular localization (other than intra- versus extracellular) and dynamics of gene expression at the single-cell level. Other techniques, such as fluorescent microscopy, are better suited for these purposes. This is exemplified by the compartmentalized expression of the *C. difficile* sporulation sigma factors using SNAP^{Cd} fusions, that enabled the characterization of the sporulation regulation pathway [30]. Other fluorophores than SNAP^{Cd} have been used for the labelling of proteins and studying gene expression in *C. difficile*, like CFP^{opt}, mCherry^{Opt} *and* phiLOV2.1. [30-33]. Despite the successful use of the different fluorophores in *C. difficile,* the work described in **Chapter 6** clearly shows that further development of fluorophores is needed.

In **Chapter 5** we applied the HaloTag for the first time in *C. difficile*, in the characterization of the *C. difficile* chromatin protein HupA. Though the HaloTag fluorophore allowed us to successfully determine nucleoid localization of the protein, a more detailed analysis showed that there is significant room for improvement of the fluorophore (**Chapter 6**). The inefficient transcription and/or translation of the cytosolic HaloTag (**Chapter 6**) and the effect observed on the HupA function (**Chapter 5**) might be overcome with codon optimization, as the HaloTag has a high [G+C]-content. Nevertheless, HaloTag, like SNAP^{Cd}, is a substrate-dependent tag and the variety of commercially available substrates is a substantial asset for HaloTag-based visualization. Furthermore, combination of HaloTag with the Luc^{opt} luciferase might allow new applications, such as NanoBRET [34], for analysis of protein interaction and dynamics in *C. difficile in vivo* studies.

In order to understand the limitations of the different fluorophores to study *C. difficile* exponential growing cells, we compared the fluorophores CFP^{opt} [31], mCherry^{Opt} [32], phiLOV2.1 [33], SNAP^{Cd} [30] and HaloTag, under defined conditions, which we present in **Chapter 6**. We were able to point important limitations of existing fluorophores and highlighted important factors required for live-cell imaging of *C. difficile.* Importantly, this is the first comparison of fluorophores under carefully controlled conditions for *C. difficile*. However, we did not explore the potential use of other fluorophores, such as the substrate-dependent fluorophores FAST [35] and UnaG [36,37]. We did perform preliminary experiments to assess some aspects of UnaG for *C. difficile* microscopy, but these demonstrated that further understanding of the dynamics of availability of the intracellular FMN co-factor, as well as optimization of protein expression, is required (data not shown).

Our work has underscored that one of the main practical limitations in *C. difficile* fluorescence microscopy is intrinsic green fluorescence. In **Chapter 6** we explored the influence of *C. difficile* autofluorescence on fluorescence microscopy and the potential limitations for the use of

green fluorophores, such as phiLOV2.1. The results show that autofluorescence in *C. difficile* is growth phase-dependent, and might be a result of oxidation of specific cell components. It would be quite interesting to determine the components conferring the green fluorescence and their relevance for *C. difficile* physiology. Most likely (ribo)flavins and flavoproteins are responsible for the intrinsic green fluorescence, as shown for *E. coli* [38]. Knocking out the *C. difficile* homolog of flavoproteins might provide a route to understand *C. difficile* autofluorescence if these do not encode essential functions. Recently, the role of flavoiiron proteins on the ability of *C. difficile* to cope with low levels of O_2, that can be encountered in the gastrointestinal tract, was studied [39,40]. A more detailed understanding of the autofluorescence may therefore also lead to insights in the response of *C. difficile* to environmental and stress conditions.

In sum, the present work has introduced several tools for *C. difficile* studies: a) two secreted reporters for gene expression analysis, the amylase AmyE^{opt} and the luciferase sLuc^{opt} (Fig. 1); b) the luciferase toolkit for *in vivo* complementation assays for extracellular detection of membrane-associated or secreted proteins (HiBiT^{opt}) and protein-protein interactions (bitLuc^{opt}) (Fig. 1); c) the HaloTag fluorophore for protein tagging and fluorescent microscopy visualization. We conclude our study by analyzing *C. difficile* autofluorescence and benchmarking the previously characterized fluorophores CFP^{opt}, mCherry^{Opt}, phiLOV2.1 and SNAP^{Cd}, and HaloTag in *C. difficile*, exploring the limitations and advantages for live *C. difficile* fluorescence microscopy studies.

TcdC C-terminus is extracellular

A major focus of gene expression analysis over the past years has been the genes encoding the *C. difficile* glucosylating toxins, that are the main virulence factors contributing to CDI. Toxins A and B are encoded within the pathogenicity locus PaLoc by the *tcdA* and *tcdB* genes, together with genes encoding for TcdE, TcdR and TcdC which have been associated with the regulation of the toxins [41]. TcdE has been described as a putative holin-like protein, required for the secretion of the toxins. However, *tcdE*-mediated secretion is still not fully understood and some studies indicate limited to no effects of a *tcdE* deletion on extracellular toxin levels [42-44]. TcdR is a sigma factor, and the role as a positive transcriptional regulator has been shown by direct recognition of the toxin promoters and subsequent activation of toxin gene expression by TcdR *in vitro* as well as *in vivo* [45,46]. However, the role of TcdC in toxin regulation is highly debatable. Some studies have suggested that TcdC acts as an anti-sigma factor, acting as a negative regulator of toxin transcription, while others failed to show a correlation between TcdC and the toxin regulation [47-50].

Over the years, several studies have shown an intricate network of toxin regulation in response to physiological and environmental factors [17,46,51-54]. Expression of TcdA, TcdB, TcdE and TcdR is low during the exponential growth phase but increases strongly upon entering the stationary growth phase [55]. This is also seen in our work, when evaluating *tcdA* expression through sLuc^{opt}, in **Chapter 2**. Thus, toxin gene expression is triggered by features present in the stationary growth phase, such as a high cell density or nutrient limitation.

Indeed, metabolism plays a crucial role in the regulation of toxin expression [46]. The presence of glucose and certain amino acids in the medium, such as cysteine or proline, has been shown to inhibit toxin expression [51,56,57]. In **Chapter 2** we also recapitulated glucose-mediated repression of *tcdA* expression using the novel sLuc^{opt} reporter. To date, several toxin regulators have been identified, such as the CcpA and CodY proteins, that bridge *C. difficile* metabolism with the repression of toxin gene expression [58,59]. σ^H, a sigma factor involved in the regulation of genes that affect sporulation and motility at the onset of stationary phase, is also involved in toxin regulation, as a *sigH* mutant has increased toxin gene expression [60]. Toxin expression is also regulated through quorum-sensing, mediated by the Agr-signaling system, which induces early transcription of *tcdA* and *tcdB* genes [61]. Recently, RstA was identified as a major regulator of toxin expression [17]. RstA represses toxin gene expression by downregulating TcdR expression and direct binding to the toxin promoters. Interestingly, the modulatory effect of RstA on the different genes is strain-dependent [17,62,63].

The role of TcdC is still unclear. The *tcdC* genotype is quite variable between *C. difficile* strains of different PCR ribotypes. It was previously hypothesized that epidemic strains, which contain deletions or frame-shift mutations in the *tcdC* gene, might produce higher levels of toxins [49,64]. However, several studies failed to demonstrate a correlation between *tcdC* genotype and toxin levels [48,50,65]. In order to understand the role of TcdC, previous studies have identified different structural domains and some of their properties [66,67]. The protein contains a (trans)membrane domain, that is believed to mediate its association with membrane, and an OB-fold domain at the C-terminus, that can bind G-quadruplex DNA structures [66,67]. However, no topological data was available. We took advantage of the newly developed HiBiT^{opt} complementation system, combined with additional biochemical experiments, to show that the TcdC C-terminus is localized extracellularly (**Chapter 3**). This raises many questions about the function of this domain and the putative binding partners. Determination of the structure of the TcdC C-terminus, especially in complex with a binding partner, would be a valuable asset to further explore the function of the domain.Previous studies have shown that TcdC might prevent the formation of a TcdR-RNA polymerase complex, thus repressing the toxins expression. Furthermore, *in vitro* binding of TcdC to ssDNA G-quadruplexes was observed [67].

However, both these functions would presumably require a cytosolic location of TcdC's OB-fold domain. The extracellular localization of the OB-fold domain suggests that binding to ssDNA G-quadruplexes *in vitro* could mimic an alternative TcdC binding partner. It is likely that TcdC could bind for instance oligosaccharides present in the environment or even at the host cell surface, contributing to the pathogenesis of the bacterium, as seen for certain OB-fold proteins in other organisms [68-70]. However, extracellular TcdC-binding partners have not yet been identified. Possible pull-down experiments, with the available repertoire of *C. difficile* tags, or direct binding to putative host cell binding partners, including (cell surface-associated) carbohydrates, could elucidate the role – if any - of TcdC in toxin regulation. Additionally, the observed higher-than-expectedmolecular weight of TcdC could suggest post-translational modifications that may be present at the TcdC C-terminus, that determine and/or affect the binding to the extracellular components [71].

An interesting, and so far poorly explored, question is how the expression of TcdC is regulated. Expression of the *tcdC* gene, in contrast to the remaining genes of the PaLoc, was found to be high in early exponential phase, and decreasing upon entry into stationary growth phase, suggesting a possible role as a negative regulator of toxin expression [55]. In contrast, other studies showed a constant expression level during the stationary growth phase [50,58,72]. Further analysis of TcdC expression under different environmental conditions, with for example a P_{tcdC}-Luc^{opt} fusion, might provide valuable insights in the regulation of *tcdC* expression. Additionally, evaluation of P_{tcdC}-Luc^{opt} in a wild-type background as well as in a *tcdC* mutant strain, might elucidate possible TcdC auto-regulation.

To date, no TcdC regulators have been identified. Although regulation of *tcdA*, *tcdB*, and *tcdR* expression appears to be modulated by different proteins, like the σ^H or the Agr quorum-sensing system, no significant differences were observed in *tcdC* expression in these experiments [60,61]. Agr-mediated transcription of TcdA and TcdB was observed in both non-epidemic – that have a functional TcdC - and epidemic strains of *C. difficile* - that lack a functional TcdC - suggesting that Agr-mediated toxin regulation is independent of TcdC [61]. Thus, TcdC regulation appears to be mediated through different mechanisms than toxin expression (if at all).

The possible role of TcdC in toxin regulation is quite unclear. Although previously decreased virulence was observed when the truncated *tcdC* was restored [49], several studies using *tcdC* mutants or restoring truncated *tcdC* showed no significant differences on the toxin expression and did not observe a correlation between *tcdC* and gene expression [48,50,65,72]. If TcdC, in fact, can act as a negative regulator of toxin expression, it likely does so indirectly rather than by

acting as an anti-sigma factor, and due to its extracellular location could play a role in linking environmental cues to toxin gene expression. Due to different *tcdC* alleles, the sensing of environmental factors might also be strain-specific, similar to what has been seen for the RstA mediated regulation of the toxin genes [63]. Importantly, determination of the localization of the N-terminus of TcdC is still necessary to establish the complete TcdC topology, as binding of the TcdC C-terminus to the unknown partner might relay a signal to a potential small intracellular N-terminal domain that might trigger an effect on toxin gene expression.

Together, the advances of the *C. difficile* toolkit have furthered our understanding of TcdC topology and also provide ways to address new and exciting questions regarding TcdC 's function.

C. difficile chromosome dynamics and replication

DNA topology can affect a plethora of cellular processes, most notably transcription and DNA replication [73-75]. DNA replication is a highly conserved process, essential for cell survival [76,77]. In most bacteria, replication of the chromosome starts with the unwinding of the origin of replication (*oriC*) and assembly of the replisome at the unwound region. Usually, bacterial DNA replication is initiated by the DnaA protein, that binds specific sequences in the *oriC* region, subsequently driving the recruitment of other proteins required not only for the formation of a replicative complex, like helicase and primase, but also for its regulation and remodelling, such as the DnaD/DnaB proteins and the nucleoid-associated proteins HU and IHF [74,78,79]. DNA replication is tightly linked to DNA topology, as seen in *E.coli*, where the disruption of the gene encoding HU resulted in perturbations of the coordination between the initiation of DNA replication and the cell growth [79,80].

Only recently, characterization of DNA replication of *C. difficile* has been tackled molecularly, in part motivated by the potential as a target for new therapies [12,81]. Proteins homologous to the *B. subtilis* and *E. coli* replication proteins have been identified and characterized in *C. difficile*, such as the DNA polymerase C (PolC), helicase and helicase loader, and also the primase [82,83]. Interestingly, several differences were observed between *C. difficile* and the gram-positive model organism *B. subtilis* [82], which may provide a handle for species-specific therapeutic approaches.

Little is known about the first steps of initiation of replication and possible regulators of this process in *C. difficile*. The differences in the replisome proteins, made us wonder whether such species-specific differences are also observed for DNA replication initiation. In this thesis, we therefore investigated origin unwinding by the initiator protein DnaA (**Chapter 4**).

In *B. subtilis* a basal system for the DnaA-dependent unwinding has been established [84]. In the proposed model, DnaA initially binds to the double-stranded DNA (dsDNA) DnaA-boxes present in the origin region. Subsequently, several DnaA molecules bind in the origin region, allowing the ATP-dependent formation of a DnaA-filament over the conserved DnaA-trio, essential for the strand separation. The formation of a DnaA-filament imposes a topological alteration, by looping the DNA and stretching the bound region, thus separating the DNA strands and leading to origin unwinding. Additional DnaA molecules bind to the unwound region, binding to the single-stranded DNA, stabilizing the unwound region and recruiting the replisome [84].

In **Chapter 4** we were able to show DnaA-dependent unwinding, in the *dnaA-dnaN* intergenic region named *oriC2.* Analysis of the *oriC* genomic organization of several different clostridia showed a conserved architecture, similar to *B. subtilis*, suggesting a possible conserved mechanism for DnaA-dependent unwinding. However, as differences in later stages of the replication mechanism were observed [82], the steps of origin melting to the helicase loading await biochemical characterization.

One important question in the initiation of replication of *C. difficile* is whether, and if so, how the origin region is remodelled by the action of proteins different from DnaA. In *E. coli*, chromosome remodelling activity at the origin region is dependent on histone-like protein IHF, which aids the formation of a DNA loop during *oriC* unwinding [85]. The *E. coli* HU protein also facilitates the stabilization of the DnaA oligomer, through direct interaction with the DnaA N-terminal domain I [80] but it is unclear whether this mechanism is present in other species. However, *C. difficile* only contains a homologue of histone-like protein HU, the HupA protein, and does not contain a homologue to IHF [86]. Prior to the work described in this thesis, this protein had not been characterized.

In **Chapter 5** we show that HupA binding to DNA induces a conformational change that leads to clear compaction of the chromosome *in vivo* and *in vitro*, suggesting that *C. difficile* HupA can indeed remodel DNA. However, an analysis of the role of HupA in chromosome remodelling for origin unwinding is not explored yet. Therefore, the HupA role in DnaA-dependent unwinding and possible binding to DnaA should be tested, for instance using the P1 nuclease assay implemented in **Chapter 4** and the developed luciferase toolkit. Of note, our work indicates that overexpression and/or C-terminal tagging with HaloTag might exert a modest effect on HupA functionality (**Chapters 5 and 6**), which should be taken into account when interpreting the results from experiments involving this fusion. To overcome this, codon-optimization of the HaloTag should be performed. Furthermore, the effect of tagging

and overexpression on HupA function could be untangled by comparison of the HaloTag-fused protein under conditions of overexpression (as used in this study) or at native levels (as for instance by insertion of the HaloTag-coding sequence through ACE at the *hupA* native locus). Insertion of the tag at the native locus of *hupA* would also allow one to explore native HupA dynamics during *C. difficile* growth in the absence of plasmid selection, and shed light on the involvement of HupA on replication initiation regulation and the reorganization of the nucleoid during sporulation.

Remodelling of the *oriC* region of *B. subtilis* has been shown to involve primosome proteins DnaB and DnaD, which are unique to Firmicutes and are likely derived from phage proteins [82,87,88]. DnaB and DnaD proteins affect the DNA topology, facilitate recruitment of the helicase and are postulated to stabilize the melted region [89-91]. Interestingly, *C. difficile* only contains one homologue of the DnaB-DnaD family of proteins, the CD3653 protein, which might perform similar functions in origin remodelling [12]. If so, CD3653 remodelling of the *oriC* may be mediated through direct interaction with DnaA, as some of the residues in the DnaA interaction surface appear to be conserved in the *C. difficile* proteins [91,92]. Some of the newly developed *C. difficile* tools and other methods used in this thesis could be used to explore the interactions of HupA and/or CD3653 with DnaA or other proteins present in the replisome or at the origin of replication, such as bitLuc^{opt}, size-exclusion chromatography and cross-linking. Additionally, the unwinding assay in combination with atomic force microscopy could be used to assess the influence of the potential interactors on the unwinding of the origin and stabilization of the DnaA filament.

A broader search of proteins interacting with DnaA or other replication proteins is likely to provide valuable information on the replication complex at the origin of replication and/or the replication fork. Multiple methods could be used for this. Tag-mediated pull-down assays in combination with mass-spectrometry identification can be used for the identification of potential binding partners for the proteins studied in this thesis. For instance, the use of the HaloTag Complete Pull-Down System (Promega) could allow the identification of novel interactors, as a HaloTag-fused protein is efficiently captured by binding to the HaloLink resin. Despite the commercial availability of the system, preliminary experiments indicated that further optimization is required for use in *C. difficile* cell cultures (data not shown). Alternatively, purified proteins could be used to raise antibodies and perform pull-down experiments of native, non-tagged, proteins. As a first step, we have already generated antibodies against DnaA and showed that it can recognize non-tagged DnaA protein (data not shown).

Finally, a technique called BioID [93,94] can be used to identify interactors, or proteins in the same complex, even if they associate only transiently. The method is based on proximity-dependent biotinylation of proteins and has successfully been used to study chromatin-associated proteins in eukaryotes [95]. The method employs a tag that is based on a mutant form of a heterologous biotin ligase, BirA, called BirA*. Expression of a BirA*-tagged protein results in the biotinylation of proteins in close proximity to the protein of interest, that can subsequently be characterized by streptavidin-based affinity purification and mass spectrometry [93,94]. Preliminary experiments suggest this strategy is feasible for *C. difficile*, but further optimization is required (data not shown).

Notwithstanding the obtained results so far, the identification of DnaA or other replication proteins interactors will need to be further investigated and validated by *in vivo* and *in vitro* experiments. For instance, using the developed luciferase toolkit to expand on the interaction between the proteins, which can be further confirmed by size exclusion chromatography and crosslinking. The available fluorescent reporters also provide a valuable tool for further characterization of the interactors. We were able to set up different constructs for live-cell microscopy for further analysis of the effect of potential protein regulators (data not shown), but also to evaluate the replication complex *in vivo*. A better understanding of chromosome dynamics and DNA replication, including its regulators, is crucial for our understanding of the coordination between the different cellular pathways and might lead to the identification of components that can be targeted for novel therapies.

In conclusion, the work presented in this thesis has provided a luciferase toolkit for further *C. difficile* studies (Fig. 1), not only for gene expression analysis (sLuc^{opt}), but also for protein-protein interaction studies *in vivo* (bitLuc^{opt}*),* secretion of proteins or even complexes present at the cell surface (HiBiT^{opt}). We investigated the strengths and weaknesses of *C. difficile* fluorescence microscopy, by comparing the fluorophores previously described in C. difficile (CFP^{opt}, mCherry^{Opt}, phiLOV2.1 and SNAP^{Cd}), and introduced the use of the fluorophore HaloTag in *C. difficile*. Additionally, we analyzed *C. difficile* intrinsic fluorescence and report the potential limitations for live *C. difficile* microscopy. The development of the luciferase systems in *C. difficile* enabled the assessment of the localization of the TcdC C-terminus and the characterization of the *in vivo* multimerization of the HupA protein. Finally, we provide new insights into the chromosome remodelling and DNA replication by studying the DnaA-dependent unwinding and characterizing the histone-like protein HupA. In this discussion, we presented the implications of our findings and provided perspectives for further research on TcdC-mediated toxin regulation and chromosome dynamics, specifically during DNA replication.

References

1 Smits, W. K., Lyras, D., Lacy, D. B., Wilcox, M. H. & Kuijper, E. J. *Clostridium difficile* infection. *Nature Reviews Disease Primers* **2**, 16020 (2016).

2 Czepiel, J. *et al. Clostridium difficile* infection: review. *Eur J Clin Microbiol Infect Dis* **38**, 1211-1221 (2019).

3 Hall, I. C. & O'Toole, E. Intestinal Flora In New-Born Infants: With A Description Of A New Pathogenic Anaerobe, *Bacillus Difficilis*. *American Journal of Diseases of Children* **49**, 390-402 (1935).

4 Loo, V. G. *et al.* A predominantly clonal multi-institutional outbreak of *Clostridium difficile*-associated diarrhea with high morbidity and mortality. *N Engl J Med* **353**, 2442-2449 (2005).

5 Kuijper, E. J. *et al.* Emergence of *Clostridium difficile*-associated disease in North America and Europe. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* **12 Suppl 6**, 2-18 (2006).

6 He, M. *et al.* Emergence and global spread of epidemic healthcare-associated *Clostridium difficile*. *Nature genetics* **45**, 109-113 (2013).

7 Rupnik, M., Wilcox, M. H. & Gerding, D. N. *Clostridium difficile* infection: new developments in epidemiology and pathogenesis. *Nat Rev Microbiol* **7**, 526-536 (2009).

8 Dawson, L. F., Valiente, E. & Wren, B. W. *Clostridium difficile*--a continually evolving and problematic pathogen. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases* **9**, 1410-1417 (2009).

9 Robinson, C. D., Auchtung, J. M., Collins, J. & Britton, R. A. Epidemic *Clostridium difficile* strains demonstrate increased competitive fitness compared to nonepidemic isolates. *Infection and immunity* **82**, 2815-2825 (2014).

10 Fossum, S. *et al.* A robust screen for novel antibiotics: specific knockout of the initiator of bacterial DNA replication. *FEMS microbiology letters* **281**, 210-214 (2008).

11 Grimwade, J. E. & Leonard, A. C. Targeting the Bacterial Orisome in the Search for New Antibiotics. *Front Microbiol* **8**, 2352 (2017).

12 van Eijk, E., Wittekoek, B., Kuijper, E. J. & Smits, W. K. DNA replication proteins as potential targets for antimicrobials in drug-resistant bacterial pathogens. *J Antimicrob Chemother* **72**, 1275-1284 (2017).

13 Karlsson, S. *et al.* Expression of *Clostridium difficile* toxins A and B and their sigma factor TcdD is controlled by temperature. *Infection and immunity* **71**, 1784-1793 (2003).

14 Fagan, R. P. & Fairweather, N. F. *Clostridium difficile* has two parallel and essential Sec secretion systems. *J Biol Chem* **286**, 27483-27493 (2011).

15 Edwards, A. N. *et al.* An alkaline phosphatase reporter for use in *Clostridium difficile*. *Anaerobe* **32**, 98-104 (2015).

16 Nawrocki, K. L., Edwards, A. N., Daou, N., Bouillaut, L. & McBride, S. M. CodY-Dependent Regulation of Sporulation in *Clostridium difficile*. *J Bacteriol* **198**, 2113-2130 (2016).

17 Edwards, A. N., Anjuwon-Foster, B. R. & McBride, S. M. RstA Is a Major Regulator of *Clostridioides difficile* Toxin Production and Motility. *mBio* **10** (2019).

18 Boetzkes, A. *et al.* Secretome analysis of *Clostridium difficile* strains. *Arch Microbiol* **194**, 675-687 (2012).

19 Hensbergen, P. J. *et al.* A novel secreted metalloprotease (CD2830) from *Clostridium difficile* cleaves specific proline sequences in LPXTG cell surface proteins. *Mol Cell Proteomics* **13**, 1231-1244 (2014).

20 Hensbergen, P. J. *et al. Clostridium difficile* secreted Pro-Pro endopeptidase PPEP-1 (ZMP1/CD2830) modulates adhesion through cleavage of the collagen binding protein CD2831. *FEBS Lett* **589**, 3952-3958 (2015).

21 Hall, M. P. *et al.* Engineered luciferase reporter from a deep sea shrimp utilizing a novel imidazopyrazinone substrate. *ACS chemical biology* **7**, 1848-1857 (2012).

22 Loh, J. M. & Proft, T. Comparison of firefly luciferase and NanoLuc luciferase for biophotonic labeling of group A *Streptococcus*. *Biotechnology letters* **36**, 829-834 (2014).

23 England, C. G., Ehlerding, E. B. & Cai, W. NanoLuc: A Small Luciferase Is Brightening Up the Field of Bioluminescence. *Bioconjug Chem* **27**, 1175-1187 (2016).

24 Boekhoud, I. M., Michel, A. M., Corver, J., Jahn, D. & Smits, W. K. Redefining the *Clostridioides difficile* sigma(B) Regulon: sigma(B) Activates Genes Involved in Detoxifying Radicals That Can Result from the Exposure to Antimicrobials and Hydrogen Peroxide. *mSphere* **5** (2020).

25 Pereira, G. C. *et al.* A High-Resolution Luminescent Assay for Rapid and Continuous Monitoring of Protein Translocation across Biological Membranes. *Journal of molecular biology* **431**, 1689-1699 (2019).

26 Allen, W. J., Watkins, D. W., Dillingham, M. S. & Collinson, I. Refined measurement of SecA-driven protein secretion reveals that translocation is indirectly coupled to ATP turnover. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 31808-31816 (2020).

27 Serrano, M. *et al.* The SpoIIQ-SpoIIIAH complex of *Clostridium difficile* controls forespore engulfment and late stages of gene expression and spore morphogenesis. *Molecular microbiology* **100**, 204-228 (2016).

28 Mie, M., Naoki, T., Uchida, K. & Kobatake, E. Development of a split SNAP-tag protein complementation assay for visualization of protein-protein interactions in living cells. *Analyst* **137**, 4760-4765 (2012).

29 Dixon, A. S. *et al.* NanoLuc Complementation Reporter Optimized for Accurate Measurement of Protein Interactions in Cells. *ACS chemical biology* **11**, 400-408 (2016).

30 Pereira, F. C. *et al.* The spore differentiation pathway in the enteric pathogen *Clostridium difficile*. *PLoS Genet* **9**, e1003782 (2013).

31 Ransom, E. M., Williams, K. B., Weiss, D. S. & Ellermeier, C. D. Identification and characterization of a gene cluster required for proper rod shape, cell division, and pathogenesis in *Clostridium difficile*. *Journal of Bacteriology* **196**, 2290-2300 (2014).

32 Ransom, E. M., Ellermeier, C. D. & Weiss, D. S. Use of mCherry Red fluorescent protein for studies of protein localization and gene expression in *Clostridium difficile*. *Appl Environ Microbiol* **81**, 1652-1660 (2015).

33 Buckley, A. M. *et al.* Lighting Up *Clostridium Difficile*: Reporting Gene Expression Using Fluorescent Lov Domains. *Sci Rep* **6**, 23463 (2016).

34 Dale, N. C., Johnstone, E. K. M., White, C. W. & Pfleger, K. D. G. NanoBRET: The Bright Future of Proximity-Based Assays. *Front Bioeng Biotechnol* **7**, 56 (2019).

35 Streett, H. E., Kalis, K. M. & Papoutsakis, E. T. A Strongly Fluorescing Anaerobic Reporter and Protein-Tagging System for *Clostridium* Organisms Based on the Fluorescence-Activating and Absorption-Shifting Tag Protein (FAST). *Appl Environ Microbiol* **85** (2019).

36 Chia, H. E., Marsh, E. N. G. & Biteen, J. S. Extending fluorescence microscopy into anaerobic environments. *Current opinion in chemical biology* **51**, 98-104 (2019).

37 Kumagai, A. *et al.* A bilirubin-inducible fluorescent protein from eel muscle. *Cell* **153**, 1602-1611 (2013).

38 Mihalcescu, I., Van-Melle Gateau, M., Chelli, B., Pinel, C. & Ravanat, J. L. Green autofluorescence, a double edged monitoring tool for bacterial growth and activity in micro-plates. *Phys Biol* **12**, 066016 (2015).

39 Martins, M. C. *et al.* How superoxide reductases and flavodiiron proteins combat oxidative stress in anaerobes. *Free Radic Biol Med* **140**, 36-60 (2019).

40 Kint, N. *et al.* How the Anaerobic Enteropathogen *Clostridioides difficile* Tolerates Low O2 Tensions. *mBio* **11** (2020).

41 Aktories, K., Schwan, C. & Jank, T. *Clostridium difficile* Toxin Biology. *Annu Rev Microbiol* **71**, 281-307 (2017).

42 Tan, K. S., Wee, B. Y. & Song, K. P. Evidence for holin function of tcdE gene in the pathogenicity of *Clostridium difficile*. *J Med Microbiol* **50**, 613-619 (2001).

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

44 Govind, R., Fitzwater, L. & Nichols, R. Observations on the Role of TcdE Isoforms in *Clostridium difficile* Toxin Secretion. *J Bacteriol* **197**, 2600-2609 (2015).

45 Mani, N. & Dupuy, B. Regulation of toxin synthesis in *Clostridium difficile* by an alternative RNA polymerase sigma factor. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 5844-5849 (2001).

46 Martin-Verstraete, I., Peltier, J. & Dupuy, B. The Regulatory Networks That Control *Clostridium difficile* Toxin Synthesis. *Toxins* **8** (2016).

47 Matamouros, S., England, P. & Dupuy, B. *Clostridium difficile* toxin expression is inhibited by the novel regulator TcdC. *Molecular microbiology* **64**, 1274-1288 (2007).

48 Murray, R., Boyd, D., Levett, P. N., Mulvey, M. R. & Alfa, M. J. Truncation in the tcdC region of the *Clostridium difficile* PathLoc of clinical isolates does not predict increased biological activity of Toxin B or Toxin A. *BMC Infectious Diseases* **9**, 103 (2009).

49 Carter, G. P. *et al.* The anti-sigma factor TcdC modulates hypervirulence in an epidemic BI/NAP1/027 clinical isolate of *Clostridium difficile*. *PLoS pathogens* **7**, e1002317 (2011).

50 Bakker, D., Smits, W. K., Kuijper, E. J. & Corver, J. TcdC does not significantly repress toxin expression in *Clostridium difficile* 630DeltaErm. *PloS one* **7**, e43247 (2012).

51 Karlsson, S., Burman, L. G. & Akerlund, T. Induction of toxins in *Clostridium difficile* is associated with dramatic changes of its metabolism. *Microbiology* **154**, 3430-3436 (2008).

52 Lyon, S. A., Hutton, M. L., Rood, J. I., Cheung, J. K. & Lyras, D. CdtR Regulates TcdA and TcdB Production in *Clostridium difficile*. *PLoS pathogens* **12**, e1005758 (2016).

53 Purcell, E. B. *et al.* A Nutrient-Regulated Cyclic Diguanylate Phosphodiesterase Controls *Clostridium difficile* Biofilm and Toxin Production during Stationary Phase. *Infection and immunity* **85** (2017).

54 Chen, K. Y. *et al.* The Transcriptional Regulator Lrp Contributes to Toxin Expression, Sporulation, and Swimming Motility in *Clostridium difficile*. *Front Cell Infect Microbiol* **9**, 356 (2019).

55 Hundsberger, T. *et al.* Transcription analysis of the genes tcdA-E of the pathogenicity locus of *Clostridium difficile*. *Eur J Biochem* **244**, 735-742 (1997).

56 Karlsson, S., Burman, L. G. & Akerlund, T. Suppression of toxin production in *Clostridium difficile* VPI 10463 by amino acids. *Microbiology* **145 (Pt 7)**, 1683-1693 (1999).

57 Karlsson, S., Lindberg, A., Norin, E., Burman, L. G. & Akerlund, T. Toxins, butyric acid, and other short-chain fatty acids are coordinately expressed and down-regulated by cysteine in *Clostridium difficile*. *Infection and immunity* **68**, 5881-5888 (2000).

58 Dineen, S. S., Villapakkam, A. C., Nordman, J. T. & Sonenshein, A. L. Repression of *Clostridium difficile* toxin gene expression by CodY. *Molecular microbiology* **66**, 206-219 (2007).

59 Antunes, A. *et al.* Global transcriptional control by glucose and carbon regulator CcpA in *Clostridium difficile*. *Nucleic acids research* **40**, 10701-10718 (2012).

60 Saujet, L., Monot, M., Dupuy, B., Soutourina, O. & Martin-Verstraete, I. The key sigma factor of transition phase, SigH, controls sporulation, metabolism, and virulence factor expression in *Clostridium difficile*. *J Bacteriol* **193**, 3186-3196 (2011).

61 Darkoh, C., DuPont, H. L., Norris, S. J. & Kaplan, H. B. Toxin synthesis by *Clostridium difficile* is regulated through quorum signaling. *MBio* **6**, e02569 (2015).

62 Edwards, A. N., Tamayo, R. & McBride, S. M. A novel regulator controls *Clostridium difficile* sporulation, motility and toxin production. *Molecular microbiology* **100**, 954-971 (2016).

63 Edwards, A. N., Krall, E. G. & McBride, S. M. Strain-Dependent RstA Regulation of *Clostridioides difficile* Toxin Production and Sporulation. *J Bacteriol* **202** (2020).

64 Curry, S. R. *et al.* tcdC Genotypes Associated with Severe TcdC Truncation in an Epidemic Clone and Other Strains of *Clostridium difficile*. (2007).

65 Cartman, S. T., Kelly, M. L., Heeg, D., Heap, J. T. & Minton, N. P. Precise manipulation of the *Clostridium difficile* chromosome reveals a lack of association between the tcdC genotype and toxin production. *Appl Environ Microbiol* **78**, 4683-4690 (2012).

66 Govind, R., Vediyappan, G., Rolfe, R. D. & Fralick, J. A. Evidence that *Clostridium difficile* TcdC is a membrane-associated protein. *J Bacteriol* **188**, 3716-3720 (2006).

67 van Leeuwen, H. C., Bakker, D., Steindel, P., Kuijper, E. J. & Corver, J. *Clostridium difficile* TcdC protein binds four-stranded G-quadruplex structures. *Nucleic acids research* **41**, 2382-2393 (2013).

68 Patel, D., Wines, B. D., Langley, R. J. & Fraser, J. D. Specificity of staphylococcal superantigen-like protein 10 toward the human IgG1 Fc domain. *Journal of immunology* **184**, 6283-6292 (2010).

69 Moreira, C. G. *et al.* Virulence and stress-related periplasmic protein (VisP) in bacterial/host associations. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 1470-1475 (2013).

70 Arunima, A. *et al.* Role of OB-Fold Protein YdeI in Stress Response and Virulence of *Salmonella enterica* Serovar Enteritidis. *J Bacteriol* **203** (2020).

71 Cain, J. A., Solis, N. & Cordwell, S. J. Beyond gene expression: the impact of protein post-translational modifications in bacteria. *J Proteomics* **97**, 265-286 (2014).

72 Merrigan, M. *et al.* Human hypervirulent *Clostridium difficile* strains exhibit increased sporulation as well as robust toxin production. *Journal of Bacteriology* **192**, 4904-4911 (2010).

73 Postow, L., Crisona, N. J., Peter, B. J., Hardy, C. D. & Cozzarelli, N. R. Topological challenges to DNA replication: conformations at the fork. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 8219-8226 (2001).

74 Donczew, R., Zakrzewska-Czerwinska, J. & Zawilak-Pawlik, A. Beyond DnaA: the role of DNA topology and DNA methylation in bacterial replication initiation. *Journal of molecular biology* **426**, 2269-2282 (2014).

75 Dorman, C. J. DNA supercoiling and transcription in bacteria: a two-way street. *BMC Mol Cell Biol* **20**, 26 (2019).

76 O'Donnell, M., Langston, L. & Stillman, B. Principles and concepts of DNA replication in bacteria, archaea, and eukarya. *Cold Spring Harbor perspectives in biology* **5** (2013).

77 Bleichert, F., Botchan, M. R. & Berger, J. M. Mechanisms for initiating cellular DNA replication. *Science* **355** (2017).

78 Chodavarapu, S. & Kaguni, J. M. Replication Initiation in Bacteria. *Enzymes* **39**, 1-30 (2016).

79 Ekundayo, B. & Bleichert, F. Origins of DNA replication. *PLoS Genet* **15**, e1008320 (2019).

80 Chodavarapu, S., Felczak, M. M., Yaniv, J. R. & Kaguni, J. M. *Escherichia coli* DnaA interacts with HU in initiation at the E. coli replication origin. *Molecular microbiology* **67**, 781-792 (2008).

81 Torti, A. *et al. Clostridium difficile* DNA polymerase IIIC: basis for activity of antibacterial compounds. *Current Enzyme Inhibition* **7** (2011).

82 van Eijk, E. *et al.* Primase is required for helicase activity and helicase alters the specificity of primase in the enteropathogen *Clostridium difficile*. *Open Biol* **6** (2016).

83 van Eijk, E. *et al.* Genome Location Dictates the Transcriptional Response to PolC Inhibition in *Clostridium difficile*. *Antimicrob Agents Chemother* **63** (2019).

84 Richardson, T. T. *et al.* Identification of a basal system for unwinding a bacterial chromosome origin. *EMBO J* **38**, e101649 (2019).

85 Cassler, M. R., Grimwade, J. E. & Leonard, A. C. Cell cycle-specific changes in nucleoprotein complexes at a chromosomal replication origin. *EMBO J* **14**, 5833-5841 (1995).

86 Krause, M., Ruckert, B., Lurz, R. & Messer, W. Complexes at the replication origin of *Bacillus subtilis* with homologous and heterologous DnaA protein. *Journal of molecular biology* **274**, 365-380 (1997).

87 Marston, F. Y. *et al.* When simple sequence comparison fails: the cryptic case of the shared domains of the bacterial replication initiation proteins DnaB and DnaD. *Nucleic acids research* **38**, 6930-6942 (2010).

88 Rokop, M. E., Auchtung, J. M. & Grossman, A. D. Control of DNA replication initiation by recruitment of an essential initiation protein to the membrane of *Bacillus subtilis*. *Molecular microbiology* **52**, 1757-1767 (2004).

89 Zhang, W. *et al.* The *Bacillus subtilis* DnaD and DnaB proteins exhibit different DNA remodelling activities. *Journal of molecular biology* **351**, 66-75 (2005).

90 Smits, W. K., Merrikh, H., Bonilla, C. Y. & Grossman, A. D. Primosomal proteins DnaD and DnaB are recruited to chromosomal regions bound by DnaA in *Bacillus subtilis*. *J Bacteriol* **193**, 640-648 (2011).

91 Martin, E. *et al.* DNA replication initiation in *Bacillus subtilis*: structural and functional characterization of the essential DnaA-DnaD interaction. *Nucleic acids research* (2018).

92 Matthews, L. A. & Simmons, L. A. Cryptic protein interactions regulate DNA replication initiation. *Molecular microbiology* **111**, 118-130 (2019).

93 Roux, K. J., Kim, D. I., Raida, M. & Burke, B. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *The Journal of cell biology* **196**, 801-810 (2012).

94 Roux, K. J., Kim, D. I., Burke, B. & May, D. G. BioID: A Screen for Protein-Protein Interactions. *Current protocols in protein science, John E. Coligan et al.* **91**, 19 23 11-19 23 15 (2018).

95 Lambert, J. P., Tucholska, M., Go, C., Knight, J. D. & Gingras, A. C. Proximity biotinylation and affinity purification are complementary approaches for the interactome mapping of chromatin-associated protein complexes. *J Proteomics* **118**, 81-94 (2015).