The replication machinery of Clostridium difficile:a potential target for novel antimicrobials

Eijk, H.W. van

Chapter 4

The replicative helicase CD3657 of *Clostridium difficile* interacts with the putative loader protein CD3654

Erika van Eijk [1]
Annemieke Friggen [1]
Vasileios Paschalis [2]
Matthew Green [2]
Geoffrey S. Briggs [2]
James Gibson [2]
Panos Soultanas [2]
Wiep Klaas Smits [1]

1 Department of Medical Microbiology, Leiden University Medical Center, Leiden, the Netherlands;

2 School of Chemistry, Center for Biomolecular Sciences, University of Nottingham, United Kingdom.

This chapter is published as part of van Eijk *et al.* Open Biology, 2016

Abstract

Clostridium difficile **is the main cause of health-care associated diarrhoea. Limited treatment options and reports of reduced susceptibility to current treatment emphasize the necessity for the development of novel antimicrobials. DNA replication is an essential and conserved process in all domains of life and may therefore serve as a target for the development of new antimicrobial therapeutics.**

Compared to its well characterized relative, *Bacillus subtilis*, knowledge of the molecular biology and genetics of *Clostridium difficile* is still in its infancy and there is a gap in our knowledge of the replication mechanisms in this organism.

Here, we identified several *C. difficile* genes with homology to *B. subtilis* replication initiation proteins and set out to characterize the replicative helicase (CD3657) and its putative loader protein (CD3654). Intra- and intermolecular protein-protein interactions were assessed by bacterial two-hybrid and analytical gel-filtration experiments. Helicase can form hexamers at high concentrations and interacts with the putative loader protein in an ATP-dependent manner. Binding of ATP to helicase is pivotal for the formation of a stable homo-hexamer and its interaction with the putative loader protein. Despite the formation of a helicase-helicase loader protein complex, no helicase activity could be demonstrated *in vitro*. Our data suggest that the *C. difficile* helicase is of the ring-maker class and that *cd3654* encodes the helicase loader, but critical aspects of helicase loading and activation differ from the Gram-positive model *Bacillus subtilis*.

Background

Extensive research, primarily on the model organisms *Escherichia coli* (Gram-negative) and *Bacillus subtilis* (Gram-positive), has shown that many different proteins are involved in DNA replication. Although the overall mechanism of replication is highly conserved in all domains of life, it is perhaps not surprising that details of the molecular mechanisms can vary substantially as these prokaryotes diverged more than 3 billion years ago [1].

One of the best characterized distinctions between *B. subtilis* and *E. coli* is the mechanism of loading the replicative helicase at the origin of replication (*oriC*), an essential step in the DNA replication process of bacteria [2–4]. Helicase is required to unwind the DNA duplex at the replication fork, and during the loading step a functional helicase multimer is assembled onto the DNA. In the Enterobacteria, Firmicutes and Aquificae, helicase loading is facilitated by a specific loader protein, which is not conserved in bacteria outside these phyla [5]. However, the strategy of helicase loading among bacteria that do code for a loader protein also differs [2,6]. In addition to its role in the helicase-loading process, the loader protein regulates the helicase through interactions and is therefore pivotal in replisome assembly [3,7,8]. For historical reasons, the nomenclature for the replication proteins differs between bacterial species (e.g. *E. coli* helicase; DnaB and *B. subtilis* helicase; DnaC). For clarity, protein names hereafter will be used in conjunction with species and either written in full or abbreviated (e.g. *Ec* and *Bs*). The *E. coli* helicase (*Ec*DnaB) is loaded by a single loader protein (*Ec*DnaC) *in vivo* [9–11], whereas loading of the *B. subtilis* helicase (*Bs*DnaC) requires three accessory proteins (*Bs*DnaD, *Bs*DnaB and *Bs*DnaI) *in vivo* [12–14] in addition to the replication initiator DnaA that is required in both organisms. One possible explanation for the requirement of multiple proteins in *B. subtilis* may lie in the fact that *E. coli* and *B. subtilis* employ different mechanisms to deliver the replicative helicase onto the DNA [2,6]. Alternatively, it may reflect different *oriC* architectures, requiring different mechanisms of origin remodelling [15].

Replicative helicases form hexameric rings and require single-stranded DNA (ssDNA) to be threaded through the central channel of the protein to unwind the DNA duplex [6,16]. To accomplish this, it is thought that either a pre-formed ring is physically opened (ring-breaker) or that the ring is assembled from monomers at *oriC*

(ring-maker) [2]. In *E. coli*, preformed hexamers of the helicase protein are capable of self-loading onto ssDNA. They display *in vitro* translocation and unwinding activities, which are highly induced in the presence of the loader protein [10]. This in contrast with *B. subtilis*, where pre-assembled hexameric helicase is inactive, irrespective of the presence of the loader protein. *In vitro*, *B. subtilis* helicase activity is only observed when the helicase protein is monomeric and loader protein is present [17]. Thus, helicase loading in *E. coli* is an example of the ring-breaker mechanism, whereas the situation in *B. subtilis* exemplifies a ring-maker mechanism.

B. subtilis helicase loading *in vivo* is a hierarchical process [13,14,18]. Initially, the double-stranded DNA (dsDNA) at *oriC* is melted into ssDNA by the initiation protein DnaA, thereby creating a substrate for primosome assembly. The *Bs*DnaD and *Bs*DnaB co-loader proteins, which are structural homologues (PFAM DnaB_2), associate sequentially with the replication origin [14], and possibly contribute to origin remodelling. This ultimately enables the ATPase loader protein to load the helicase [13-15,18-23].

The replication initiation protein (*Bs*DnaA) and helicase-loader protein (*Bs*DnaI) belong to the AAA+ (ATPases associated with various cellular activities) family of ATPases [10,24-26]. These AAA+ enzymes are, in their turn, part of the additional strand catalytic glutamate (ASCE) family [15,25,27-30]. The *Bs*DnaI loader protein consists of a C-terminal AAA+ domain that is necessary for nucleotide and ssDNA binding, and an N-terminal helicase-interacting domain [17].

The *Bs*DnaC helicase is also an ASCE protein and belongs to RecA-type helicase Superfamily 4 (SF4), which is involved in DNA replication [29,31,32]. The SF4 superfamily of helicases is characterized by five sequence motifs; H1, H1a, H2, H3, and H4 [29,33,34]. Motifs H1 and H2 are equivalent to the ATP-coordinating Walker A and B motifs found in many other ATPases [29].

We are interested in DNA replication in the Gram-positive bacterium *C. difficile*, the most common causative agent of antibiotic-associated diarrhoea [35,36]. In this study, we employed an *in silico* analysis to identify homologues of replication initiation proteins from *B. subtilis*. From thereon, we focused on the helicase and a putative helicase loader. Our data show that helicase loading and activation in *C. difficile* may differ critically from the *B. subtilis* model.

Results

In silico identification of putative replication initiation proteins

In *B. subtilis*, replication initiation requires the coordinated action of multiple proteins, DnaA, DnaD, DnaB and DnaI, *in vivo* [14]. BLASTP queries of the genome of *C. difficile* 630 (GenBank AM180355.1) using the amino acid sequence of these proteins from *Bacillus subtilis* subsp. *subtilis* strain 168 (GenBank NC0989.1) allowed the identification of most, but not all, proteins that were found to be essential for replication initiation in *B. subtilis*.

Homologues of the initiation protein DnaA (CD0001; *e*-value = 0.0) and the putative helicase (CD3657; *e*-value = 7×10^{-713}) were identified with high confidence, sharing respectively 62 and 52 percent identity with their *B. subtilis* counterparts across the full length of the protein. Interestingly, no homologue of *Bs*DnaB was found using this strategy. However, BLASTP shows that the genome of *C. difficile* does harbour two homologues of *Bs*DnaD (CD2943: *e*-value = 2×10^{-5}, identity = 29%, query coverage = 32%; CD3653: *e*-value = 4×10^{-5}, identity 29%, query coverage 47%). As *Bs*DnaB and *Bs*DnaD are strictly required for replication initiation in *B. subtilis* [18,37] and are structurally related despite limited amino acid sequence similarity [19], we further examined the *C. difficile* homologues of *Bs*DnaD.

*Bs*DnaD is composed of two domains; DDBH1 and DDBH2, whereas *Bs*DnaB has a DDBH1-DDBH2-DDBH2 structure [19]. The DnaD-like proteins CD3653 and CD2943 of *C. difficile* both consist of three domains (DDBH1-DDBH2-DDBH2) and therefore resemble *Bs*DnaB in domain structure (**Figure 1A**). CD2943 is annotated as a putative phage replication protein and is located in the ~50kb prophage 2 [38]. Also, in *Listeria monocytogenes*, *Staphylococcus aureus* and *Lactobacillus plantarum* DDBH2-containing phage genes have been identified [19,39].

Considering the fact that the prophage is not part of the *C. difficile* core genome, we consider a role for CD2943 in chromosomal DNA replication unlikely, but a role for CD3653 plausible.

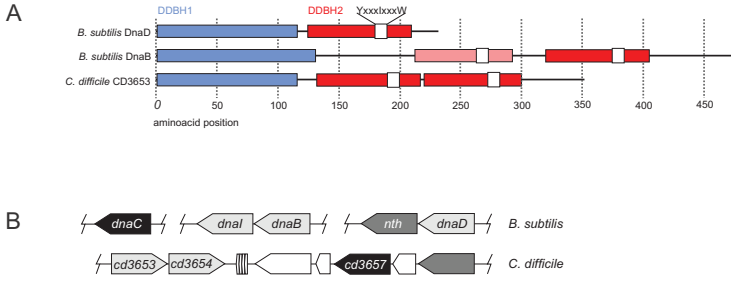


Figure 1. ***In silico*** **analysis of putative replication initiation proteins of** ***C. difficile*****.**

A. Domain structure of BsDnaD, BsDnaD and CD3653. Domain nomenclature according to Marston *et al.* [19]. Note that DDBH2 corresponds to PFAM DnaB_2.
B. Chromosomal organization of the *dnaBI* genomic region of *B. subtilis* and the CD3653–CD3654 genomic region of *C. difficile*.

A similar argument can be made for the putative helicase loader. Two homologues of the *Bs*DnaI protein were identified in *C. difficile* by BLASTP (CD3654: *e*-value = 8×10^{-10}, identity = 26%, query coverage = 46%; CD0410: *e*-value = 1×10^{-18}, identity = 31%, query coverage = 51%). The sequence homology of putative loader proteins with their counterpart in *B. subtilis* is mainly confined to the C-terminal AAA+ domain that contain the Walker A and B motifs [40]. CD0410 is located on the conjugative transposon CTn*2* [41], and therefore not part of the core genome of *C. difficile*. The *dnaD*-like gene CD3653 is located adjacent to the putative loader (CD3654), in the same genomic region as the replicative helicase (CD3657) (**Figure 1B**). Of note, the *dnaB* gene of *B. subtilis* is located next to the gene helicase loader *dnaI* [37], suggesting a functional relationship between the loader ATPase and a DnaB_2 family protein. Indeed, it has been suggested that *Bs*DnaB is a co-loader of the *Bs*DnaC helicase [18]. Taken together, our analyses strongly suggest that CD3654 is the cognate loader protein for the *C. difficile* replicative helicase (CD3657).

Helicase can form hexamers at high concentration

A distinguishing feature of the different modes of helicase loading (ring-maker versus ring-breaker) is the multimeric state of helicase at dilute concentrations of protein [2]. Therefore, we purified recombinant *C. difficile* helicase protein and deter-

mined its multimeric state using analytical gel-filtration. At concentrations below 5 μM we observed predominantly monomeric protein, with a fraction of the protein forming low molecular weight (MW) complexes (probably dimers or trimers) while at 10 μM and above, the helicase formed hexamers (**Figure 2**). Thus, at physiological (nM) concentrations the *C. difficile* helicase is predominantly monomeric, suggesting it is of the ring-maker type, like *B. subtilis*. Multimerization at high concentrations of protein was independent of the presence of ATP (*data not shown*).

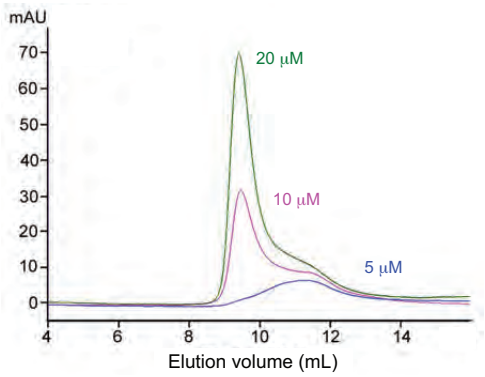


Figure 2. The helicase CD3657 demonstrates concentration dependent hexamerisation.

Analytical gel- filtration was performed in buffer B (see Methods) on a HiLoad 10/300 GL Superdex 200 analytical grade size exclusion column with the indicated concentration of CD3657 protein.

To confirm the gel-filtration data, we investigated the self-interaction of helicase in a bacterial two-hybrid system based on Gateway cloning [42]. This system detects interactions between a protein fused to Zif (Zinc-finger DNA binding domain) and a protein fused to the RNA polymerase ω subunit. Interaction between proteins of interest facilitates transcriptional activation of a Zif-dependent *lacZ* reporter gene in a dedicated reporter strain [42]. In order to quantify the interaction, *E. coli* cells containing the plasmids encoding the fusion proteins were permeabilized and assayed for β-galactosidase activity [43]. We found that in the reporter strain transformed with plasmids harbouring both fusion proteins, β-galactosidase activity was ~3-fold higher than for the reporter strains harbouring the individual plasmids, indicating a clear self-interaction for the *C. difficile* helicase protein (**Figure 3A**).

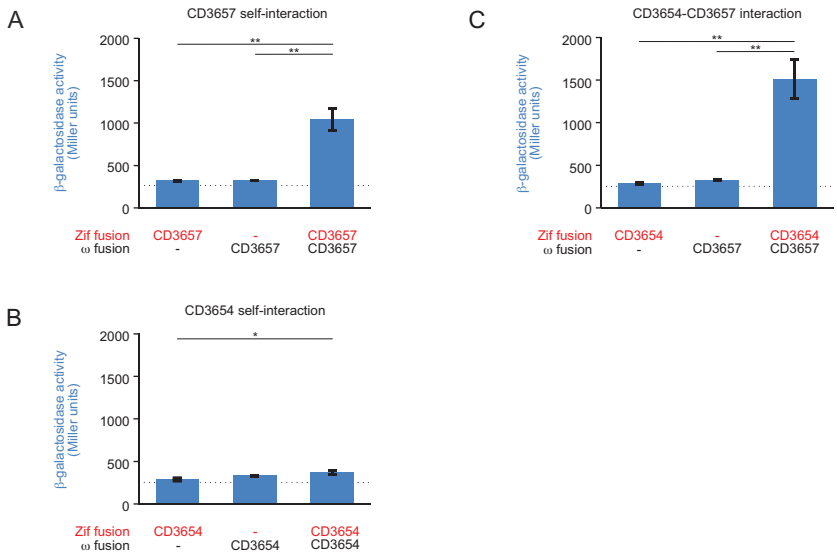


Figure 3. Bacterial two-hybrid analysis of the helicase CD3657 and putative loader protein CD3654.

A. Bacterial two-hybrid analysis of CD3657 self-interaction.

B. Bacterial two-hybrid analysis of CD3654 self-interaction.

C. Bacterial two-hybrid analysis of the CD3657-CD3654 interaction. Bars indicate average values and error-bars indicate standard deviation of the measurements (n=3). Dashed line indicates the maximum background level of β-galactosidase expression observed in our experimental set-up. Significance was determined using the Student's t-test (* $p < 0.05$, ** $p < 0.001$).

Similar experiments were carried out with the putative helicase loader (CD3654). Analytical gel-filtration using purified loader protein (CD3654) showed that it was monomeric at all concentrations tested (*data not shown*). Consistent with this observation, no self- interaction of CD3654 was found in the bacterial two-hybrid system (**Figure 3B**).

We conclude that helicase can form homomultimeric assemblies, whereas the putative loader is monomeric.

Helicase and the putative helicase loader interact in an ATP-dependent manner

If CD3654 is a legitimate loader for the *C. difficile* helicase (CD3657), we expect that the proteins interact *in vivo* and *in vitro*. To determine if this is the case, we performed bacterial two-hybrid and analytical gel-filtration experiments. First, we tested if an interaction between the helicase CD3657 and the putative loader CD3654 could be demonstrated in the bacterial two-hybrid system. CD3657 was fused to the RNA polymerase ω subunit and CD3654 was fused to Zif [42]. The β-galactosidase activity in the reporter strain containing both plasmids was ~5 -fold increased (p <0.01) compared to the reporter strains containing the individual plasmids (background) (**Figure 3C**). The high β-galactosidase activity implies substantial interactions between the *C. difficile* helicase and putative helicase loader. Similar results were obtained when CD3657 was fused to Zif and CD3654 to the RNA polymerase ω subunit (*data not shown*). This suggests that the combination of protein and fusion domain does not influence the results of this assay.

To exclude the possibility for false negative or false positive results as a result of the two-hybrid system, we additionally performed analytical gel-filtration experiments using purified non-tagged CD3657 and CD3654 proteins. In these experiments, *C. difficile* helicase and loader were combined in equimolar concentrations (2.21 μM) in the presence and absence of ATP (1 mM). In the presence of ATP, the elution profile showed a major high molecular weight (MW) peak (~10 mL; ~500 kDa, P1) and a minor low MW peak (~15 mL; ~40 kDa, P2) (**Figure 4**, red profile). In combination with a visual inspection of the fractions collected from both peaks on a Coomassie-stained SDS-PAGE gel, we believe that the major peak can be attributed to a large complex (most probably a dodecameric assembly consisting of six CD3657 monomers and six CD3654 monomers; theoretical MW 522 kDa), whilst the minor peak corresponds to predominantly free monomeric CD3654 (theoretical MW 38 kDa). Similar results were obtained when high concentrations of proteins (~10 μM) were used (**Supplemental Figure 1**), suggesting that pre-formed hexameric helicase retains the ability to interact with the CD3654 protein at the same stoichiometry.

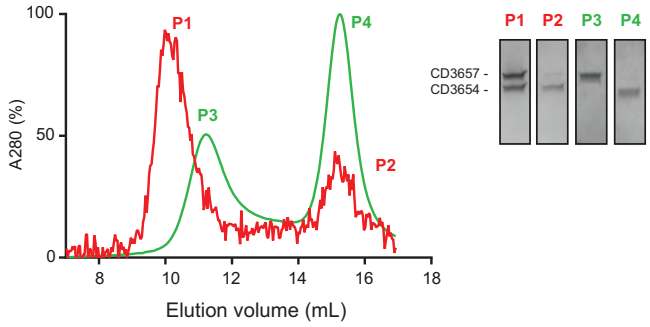


Figure 4. The helicase CD3657 and the putative loader CD3654 interact in an ATP-dependent manner.

Analytical gel filtration was performed on a HiLoad 10/300 GL Superdex analytical grade size exclusion column with 2.21 µM (monomer) of CD3657 and CD3654 in the presence (red) and absence (green) of 1 mM ATP. Inset shows a Coomassie-stained SDS-PAGE gel of the numbered peak fractions.

The elution profile of the same concentration of proteins in the absence of ATP showed a completely different picture (**Figure 4**, green profile). A minor peak was observed at ~11 mL (~300 kDa, P3) and a major second peak eluted at ~15 mL (~40 kDa, P4). Molecular weight estimates, in combination with an evaluation of peak fractions on an SDS-PAGE gel, indicated that the first peak most likely corresponds to a complex of six monomers of CD3657 (theoretical MW 297 kDa), whilst the second peak corresponds to monomeric CD3654 protein. Together, the data shows that the CD3657 helicase and the putative loader CD3654 can form a complex in an ATP-dependent manner.

Mutation of the helicase Walker A motif abrogates protein-protein interactions

To address the question which of the proteins (or whether both) requires ATP to promote the formation of a CD3657- CD3654 complex, mutants in the Walker A motif of both proteins were created. The Walker A motif (GXXXXGK[T/S]) directly and indirectly interacts with ATP and is the principal ATP-binding motif of P-loop ATPases [25]. The motif is highly conserved in both helicase and helicase loader

proteins [6]. The lysine residue (K) forms a direct interaction with the negatively charged nucleotide β or γ phosphate group and mutation of this residue is known to abrogate nucleotide binding and lead to inactivation of P-loop ATPases. The threonine (T) residue in the Walker A motif either directly or indirectly coordinates an Mg^{2+} ion within the ATP-binding site, which in turn coordinates the phosphate groups of ATP. In the *Geobacillus stearothermophilus* replicative helicase, mutation of the threonine residue results in a protein that lacks ATPase and unwinding activities [44]. Based on this knowledge, the equivalent residues were identified in the *C. difficile* helicase protein. Using site-directed mutagenesis, we generated mutant helicase proteins in which the lysine at position 214 was changed into an arginine (K214R) and the threonine at position 215 was changed into an alanine (T215A). We performed filter-binding assays to determine if the mutant helicase proteins demonstrated altered ATP binding. We found that the K214R mutant bound ATP 2-fold less than wild type, whereas the T215A mutant demonstrated 5-fold less ATP binding at a concentration of 40 nM ATP; at lower ATP concentrations, the difference was even more pronounced (**Figure 5**).

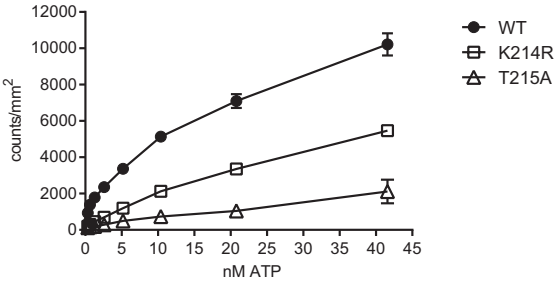


Figure 5. Walker A mutants of the helicase CD3657 show reduced binding of ATP.

2 pmol mM of CD3657 protein was incubated with the indicated amount of α-32P -radiolabelled ATP. The amount of radioactivity that remained associated with the protein was determined by autoradiography. CD3657 K214R and T215A are Walker A mutants.

To determine whether these CD3657 proteins showed altered protein-protein interactions, we performed bacterial two-hybrid experiments. We fused the CD3657 protein to the ω subunit and CD3654 or CD3657 proteins to the Zif subunit and evaluated their ability to drive the expression of a transcriptional reporter.

For both the CD3657 K214R- and CD3657 T215A mutants, interaction with the wild type CD3654 protein were severely reduced or completely lost (**Figure 6C** and **D**). This prompted us to investigate if the mutant helicases still had the capacity to homo-multimeric assemblies, as it has been shown in other bacteria that oligomerization is an important step in the mechanism of action of DNA helicases [45]. We found that the CD3657 K214R also demonstrated reduced (probably absent) self-interactions (**Figure 6A**), whereas the CD3657 T215A mutant had completely lost the ability to self-interact (**Figure 6B**). We conclude that the ability of the CD3657 helicase to coordinate ATP correlates with its ability to interact with the putative loader CD3654 and the ability to self-interact. As the most dramatic effect was observed for CD3657 T215A, we focused our further experiments on this particular mutant.

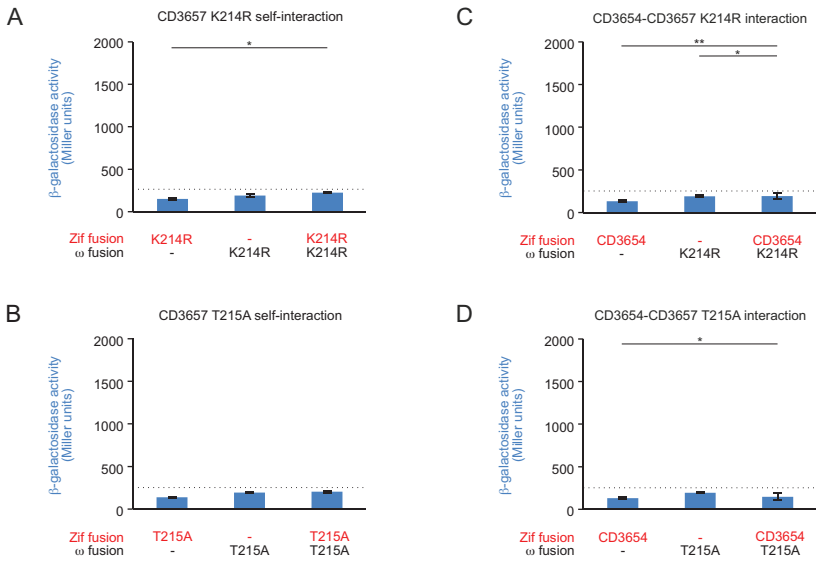


Figure 6. Walker A mutants of CD3657 are defective in protein-protein interactions.

Walker A mutants of CD3657 K214R (**A**);
T215A (**B**) no longer self-interact in a bacterial two-hybrid assay. Walker A mutants of CD3657 K214R (**C**);
T215A (**D**) show no or severely reduced interactions with the putative loader protein CD3654 in a bacterial two-hybrid assay. Bar graphs (A-D) indicate average values and error-bars indicate standard deviation of the measurements (n=3). Dashed line indicates the maximum background level of β-galactosidase expression observed in our experimental set-up. Significance was determined using the Student's t-test (* $p < 0.05$, ** $p < 0.001$).

To confirm the findings from the bacterial two-hybrid experiments, we purified CD3657 T215A protein and subjected it to size exclusion chromatography (**Figure 7**, blue line). *C. difficile* helicase T215A mutant (2.43 µM) was incubated in the presence of ATP (1 mM) and loaded onto a size exclusion column. A major peak was observed at ~15 mL (~40 kDa), probably corresponding to monomeric CD3657 T215A (theoretical MW 49 kDa) (**Figure 7**, P2 SDS-PAGE analysis). We did not observe any high MW complexes under these conditions, in contrast to the wild-type CD3657 protein (**Figure 2** and **4**). Next, we combined the CD3657 T215A mutant and the wild-type CD3654 protein (both 2.43 µM) in the presence of ATP (1 mM) (**Figure 7**, green line). A single peak was observed at ~15 mL (~40 kDa). Analysis of the peak fractions on SDS-PAGE demonstrated that the peak contained both CD3657 T215A and CD3654 protein, and thus corresponds to monomeric forms of both proteins (theoretical MW 49 and 38 kDa, respectively) (**Figure 7**, P1 SDS-PAGE analysis). Also, in these experiments no high MW complexes were found, in contrast to the wild-type CD3657 and CD3654 proteins (**Figure 4**).

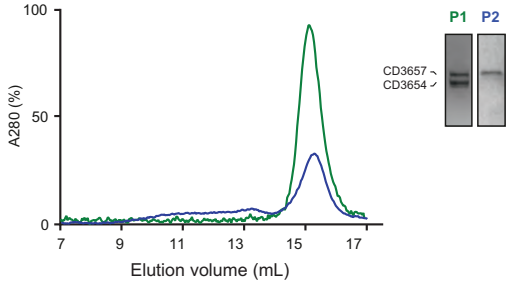


Figure 7. CD3657 T215A no longer has the ability to self-interact or interact with the putative loader CD3654.

Analytical gel filtration was performed in buffer A (see Methods) on a HiLoad 10/300 GL Superdex 200 analytical grade size exclusion column with 2.43 µM of CD3657 T215A in the presence (green) and absence (blue) of 2.43 µM CD3654. Inset shows a Coomassie-stained SDS-PAGE gel of the peak fractions.

We also generated constructs with Walker A mutations in CD3654 (K198R, T199A) and tested these in the bacterial two-hybrid assay and in size exclusion chromatography. The mutant loader proteins retained the ability to interact with the wild-type helicase protein (**Supplemental Figure 2**). Self-interaction was not observed for the

mutant loader proteins in a bacterial two-hybrid assay, concordant with the results for wild-type loader protein (**Supplemental Figure 3**). For this reason, we did not investigate the loader mutants further in this study.

Overall, our data show that the Walker A mutant CD3657 T215A can no longer self-interact and has lost the capacity to interact with the putative loader protein CD3654. We conclude that the ATP requirement for the interaction between the two proteins is most likely the result of ATP binding to CD3657 and not to CD3654.

Helicase loading of *C. difficile* differs from *B. subtilis*

So far, our data shows that the replicative helicase of *C. difficile*, CD3657, interacts in an ATP- dependent manner with the putative helicase-loader protein CD3654 and that loading probably occurs via a ring-maker mechanism, as for *B. subtilis*. In *B. subtilis*, stimulation of activity of the (monomeric) helicase protein by the loader protein was clearly shown using an *in vitro* helicase activity assay [17,46]. Therefore, we set out to investigate if the activity of *C. difficile* helicase could be reconstituted in the presence of CD3654 protein in a similar experiment. DNA-unwinding helicase activity was assayed by monitoring and quantifying the displacement of a radiolabelled oligonucleotide (partially) annealed to single-stranded circular M13 DNA (**Figure 8A**). To enable loading of helicase, the 5'end of the oligonucleotide contained a poly(dCA) tail that produces a forked substrate upon annealing of the complementary region to ssM13. Wild-type CD3657 and CD3654 proteins were mixed in equimolar concentrations (monomers) in the presence of ATP and reaction buffer and displacement of the radiolabelled oligonucleotide was monitored over time. In contrast to the *B. subtilis* proteins, helicase activity was not observed during this time course (**Figure 8B**). This may suggest that another factor is required for *in vitro* loading and/or activation of the *C. difficile* helicase.

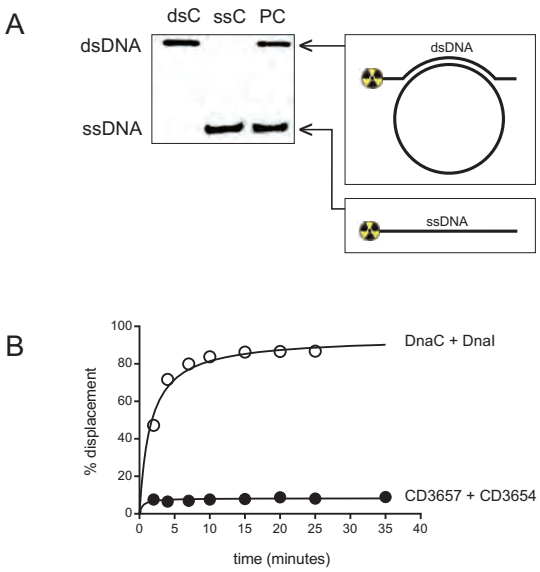


Figure 8. Helicase activity in the presence of helicase and (putative) loader proteins of *B. subtilis* and *C. difficile*.

A. Helicase activity was assayed by quantifying the displacement of a radiolabelled (γ^{32}P–ATP) oligonucleotide partially annealed to the single stranded circular DNA m13mp18. dsC: double-stranded control; ssC: single-stranded control; PC: positive control.

B. Percent displaced signal from the helicase assays in time for the *B. subtilis* proteins or *C. difficile* proteins.

Discussion

In silico analysis of the *C. difficile* genome by BLASTP identified homologues of most proteins that are involved in the DNA replication process of *B. subtilis*, which is generally considered the model for Gram-positive bacteria. However, *C. difficile* does not encode a *B. subtilis* DnaB homologue. This protein, together with *Bs*DnaD and *Bs*DnaI loader, is strictly required for helicase loading in *B. subtilis in vivo* [13,14,18]. A homologue of *Bs*DnaD was identified that may be involved in DNA replication (CD3653; *e*-value = 4×10^{-5}), although query coverage (47 percent) and identity

(29 percent) were low. This situation is reminiscent of that in some Mollicutes, where also only a *DnaD*-like gene was identified [15]. Despite a lack of clear homology at the primary amino acid sequence level, *Bs*DnaB and *Bs*DnaD are structural homologues [19]. Fusions of these proteins are found in phage-related replication proteins and it was suggested that in the absence of DnaB, a single fusion protein may couple or combine both functions [15,39]. Nevertheless, the situation in *C. difficile* differs from those in phage and Mollicutes. Structure predictions reveal that the phage-related and the Mollicutes DnaD-like proteins have a two-domain structure containing one copy of the DDBH1 and DDBH2 domain [19], and the proposed hybrid function of phage proteins is based on limited local amino acid sequence similarity in the DDBH2 domain only [39]. CD3653 on the other hand has a three-domain structure with a single DDBH1 and two DDBH2 domains, like DnaB, despite the lack of sequence similarity to this protein (**Figure 1A**). It is tempting to speculate that CD3653 in *C. difficile* may perform functions similar to both DnaD and DnaB in *B. subtilis*, which include origin remodelling and contributing to the helicase loading process [15,18,47].

DDBH2 domains are characterized by an YxxxIxxxW motif [19]. In *Bs*DnaB, this motif is degenerate in the first DDBH2 domain. By contrast, this motif is readily identified in both DDBH2 domains of CD3653 (**Figure 1A**). Our data are consistent with a model where an ancestral three-domain DnaD-like protein was duplicated and subsequently diverged in certain Firmicutes like *B. subtilis*.

DnaB-like helicases (note that the nomenclature is based on the *E. coli* protein name) belong to the superfamily 4 of DNA helicases (SF4), and the functional unit of this protein is a hexamer [25,28,29,31]. In *E. coli*, the helicase is found to be a stable hexamer over a broad protein concentration range of 0,1 to 10 μM [48] and is active as a pre-formed multimer. Helicases belonging to the ring-maker class, such as *B. subtilis*, can occur in low-oligomeric or monomeric state under dilute conditions [2,18]. Our experiments indicated that CD3657 is monomeric in the low micromolar or nanomolar range (**Figure 2**), which is likely to be reflective of the intracellular concentration of protein [49]. *Clostridium difficile* CD3657 can form hexameric assemblies at higher concentrations (**Figure 2** and **Supplemental Figure 1**), but these pre-formed hexamers are inactive (*data not shown*), in contrast with the situation in *E. coli*. Our data are therefore consistent with the notion that CD3657 belongs to the ring-maker class of helicases [2].

Though the addition of ATP was not strictly required for hexamerization of CD3657 at high concentrations of protein (**Figure 2**), the interaction between putative helicase and loader protein was found to be ATP-dependent (**Figure 4**). Mutations in the conserved Walker A and Walker B motifs of the putative loader protein did not abrogate the interaction with the wild-type helicase (**Supplemental Figure 2**). Similarly, in *E. coli*, nucleotide binding to the helicase loader was not a prerequisite for association with helicase [10,45,49]. Instead, our data indicate that association of ATP with helicase is crucial for the interaction with the loader protein (**Figure 6** and **7**). Notably, there is a correlation between the ability of the helicase interact with loader and to form homohexamers, as a T215A (Walker A) mutant of helicase is defective for both, at least under dilute concentrations of helicase (**Figure 6** and **7**). By contrast, the equivalent mutation in *G. stearothermophilus* helicase (T217A) does not affect its ability to form hexamers [44], and the interaction of this protein with *B. subtilis* DnaI readily occurs in the absence of ATP [40].

Both Walker A mutants of CD3657 demonstrate similar effects on the protein-protein interactions, that we attribute to defects in ATP binding rather than hydrolysis. Both mutants show reduced binding of ATP (**Figure 5**); a Walker B mutant (CD3657 D318A), mirrors our findings with the Walker A mutants in a bacterial two-hybrid assay. But a Walker B mutant that is predicted to be able to bind ATP but not hydrolyse (CD 3657 E239A) does not (*our unpublished observations*).

Binding of the nucleotide to helicase is associated with conformational changes; the N-terminal collar domain constricts upon nucleotide binding in *Aquifex aeolicus* and to a lesser extent *E. coli*. This constricted conformation is believed to favour an interaction with the loader protein [50] and it is therefore conceivable that in the absence of ATP, the *C. difficile* helicase adopts a (dilated) conformation that is incompatible with a functional interaction with the putative loader protein.

Despite substantial bioinformatic- and biochemical evidence that CD3654 is indeed a helicase loader (**Figure 1** and **4**), we did not observe helicase activity in an *in vitro* assay with purified helicase alone (*data not shown*) or in combination with the loader protein (**Figure 8B**). These findings are in contrast with two other Gram-positive replicative helicases; *G. stearothermophilus* helicase demonstrates significant helicase activity by itself [44], and the *B. subtilis* helicase is strongly activated by its cognate loader [46]. However, genes encoding homologues of CD3654 (R20291_3513)

and CD3657 (R20291_3516) were found to be essential in an epidemic strain, supporting their identification as DNA replication initiation proteins [51]. Together, these data strongly suggest that the presence of loader protein alone is not sufficient to activate the helicase in *C. difficile*, and that at least one other factor is needed to reconstitute its activity.

Helicases are complex proteins, and their properties can both alter and be altered by other replication factors [50]. DnaB-like helicases consist of two-tiered homo-hexameric rings, one assembled from six subunits of the C-terminal domain and the other formed by the N-terminal domain. The helicase loader interacts with the C-terminal ATPase domain [49,52,53], and the same domain is required for the interaction with the τ subunit of the clamp loader protein in *E. coli* [54], and *B. subtilis* [55,56]. Strikingly, *B. subtilis* helicase T217A Walker mutant fails to form a complex with τ [55]. This finding is very similar to our observations for the interaction between the *C. difficile* helicase and loader.

The N-terminal domain of helicase forms a platform for the interaction with primase in both *E. coli* and *G. stearothermophilus* [50,53]. Unlike the helicase loader, binding of primase to helicase is promoted by a dilated conformation of the N-terminal domain that exposes the interaction surface [50,57]. The helicase-primase interaction is mutually stimulatory, with distinct but overlapping networks of residues in helicase responsible for the modulation of either helicase or primase activity [58-60]. Primase binding counteracts the binding of the loader protein in *E. coli* [61]. Similarly, helicase loader protein from *B. subtilis* was found to dissociate from the complex when primase and polymerase bind to helicase [46]. It is unknown if *C. difficile* primase can exert a similar function.

In summary, our data show that the mechanisms of loading and activation of the replicative helicase of *C. difficile* are likely critically different from the Gram-positive model *Bacillus subtilis*. It is tempting to speculate that CD3653, primase or other protein factors will allow reconstitution of helicase activity *in vitro*.

Materials and Methods

Plasmid construction

All oligonucleotides and plasmids constructed for this study are listed in **Supplemental Table 1** and **2**. To construct the helicase loader (CD3654) and helicase (CD3657) expression plasmids, the open reading frames were amplified with high fidelity polymerase *Pfu* via PCR from *C. difficile* strain 630Δ*erm* chromosomal DNA [62,63], using primers oEVE-4 and oEVE-6 for CD3654 and oWKS-1185 and oWKS-1367 for CD3657. The reverse primer of both genes introduces a stop codon before the *Xho*I site, thereby ensuring that the protein is in its native form, when expressed (no C-terminal 6x His-tag). The CD3654 PCR product was digested with *Nco*I and *Xho*I restriction nucleases and ligated into vector pET28b (Novagen) to yield pEVE-24. The CD3657 PCR product was digested with *Nde*I and *Xho*I and ligated into vector pET21b (Novagen) to yield pEVE-87. The DNA sequence of the constructs (pEVE-24 and pEVE-87) was verified by sequencing.

Construction of the plasmids for the bacterial two-hybrid system was performed with Gateway cloning technology (Invitrogen), which is based on phage λ site- specific recombination. To construct the CD3654 and CD3657 Entry plasmids, the CD3654/CD3657 open reading frame was amplified with high fidelity polymerase *Pfu* via PCR from *C. difficile* strain 630Δ*erm* chromosomal DNA, using primers oAF-26 and oAF-27 for CD3654 and oAF-28 and oAF-29 for CD3657. This resulted in attB-flanked PCR products (1 µl) that could be recombined into donor vector pDonR™201 (1 µl, 50 ng/ µl) with BP Clonase II enzyme mix (0.5 µl). The reaction was incubated at 25°C for 1.5 hours and transformed into chemically competent *E. coli* DH5α cells by heat shock. After overnight incubation on LB plates at 37˚C, kanamycin resistant colonies were selected. Bacterial two-hybrid constructs were made by sub-cloning the genes of interest from the entry plasmids into the destination plasmids pKEK1286 (Zif fusion plasmid) or pKEK1287 (ω fusion plasmid) in an LR reaction [42]. In brief, the Entry clones (1 µl, 50 ng/ µl) were mixed with one of the pKEK1286 or pKEK1287 (1 µl, 50 ng/ µl) destination vectors and LR Clonase II enzyme mix (0.5 µl). After the reaction was incubated at 25°C for 1.5 hours, the formulation was transformed as a whole into chemically competent *E. coli* DH5α cells by heat shock. Resulting Expression clones were selected with tetracycline for the Zif fusion plasmid

or ampicillin for the ω fusion plasmid. The DNA sequence of all constructs (pEVE-122, pEVE123, pEVE124 and pEVE125) was verified by sequencing.

Site-directed mutagenesis

Walker A mutant constructs of CD3657 and CD3654 mutants were constructed according to the QuikChange protocol (Stratagene). Primers were generated with Primer X, a web-based tool for automatic design of mutagenic primers for site-directed mutagenesis. QuikChange was carried out using Pfu polymerase and plasmids pEVE-24, pEVE-87, pEVE-122, pEVE123, pEVE124 and pEVE125 as templates. All mutant constructs were verified for the correct mutation by DNA-sequencing.

Purification of the helicase CD3657

Overexpression *C. difficile* CD3657 was carried out in *E. coli* BL21 (DE3) from the pEVE-12 plasmid. The growth medium consisting of 2XYT broth (1.2 L), carbenicillin (50 µg/mL) and antifoam 204 (Sigma-Aldrich) was inoculated with a pre-culture (10 mL). The cell culture was incubated at 37°C with mechanical shaking at 180rpm, until an optical density (600 nm) of 0.70-0.85 was reached (after approximately 3 h). Protein expression was induced via the addition of IPTG (1 mM) and the culture was incubated at 30°C for 3 h. The cells were harvested by centrifugation (3000 g, 15 min, 4°C) and resulting cell paste was stored at -80°C. CD3657 cell paste, prepared from 1.2 L cell culture, was re-suspended in 30 mL TED50 buffer (Tris pH 7.5 50 mM, EDTA 1 mM, DTT 1 mM, NaCl 50 mM) with PMSF (1 mM). The bacterial cells were lysed by sonication and crude lysate was clarified by centrifugation (35,000 g, 30 min, 4°C). The resulting supernatant was separated from the cell debris using a 0.22 µm pore filter before a 50% ammonium sulphate precipitation, followed by clarification by centrifugation (35,000 g, 30 min, 4°C). The ammonium sulphate precipitated pellet was suspended in TED50 buffer and loaded onto a 5mL Q sepharose column, equilibrated in TED50 buffer. The protein was eluted using a gradient of 20 to 100% TED1000 buffer (Tris pH 7.5 50 mM, EDTA 1 mM, DTT 1 mM, NaCl 1000 mM) over 15 CV. The fractions containing the protein of interest were pooled (12 mL, 25 mS) and diluted with 33 mL of TED50 buffer (Tris pH 7.5 50 mM, EDTA 1 mM, DTT 1 mM) to give an adjusted volume of 45 mL and conductivity of

10.1 mS. This solution was loaded onto a 5 mL heparin column equilibrated in TED50 buffer; the protein eluted in the flow through. After a further ammonium sulphate precipitation to concentrate the sample, the collected protein was loaded onto a Hiload 26/60 Superdex 200 gel filtration column equilibrated in TED50 buffer, yielding hexameric CD3657 oligomer. A further ammonium sulphate precipitation was used to concentrate the sample in a reduced volume of 4mL in TED50 buffer. Guanidinium chloride solution (8 M) was added portion wise (30 x 44.4 µL) to the protein solution with rapid stirring, to give a final concentration of 2 M guanidinium chloride. The protein was then loaded onto a Hiload 26/60 Superdex 200 gel filtration column equilibrated in TED50-GC buffer (Tris pH 7.5 50 mM, EDTA 1 mM, DTT 1 mM, NaCl 50 mM, guanidinium chloride 2 M), to give the CD3657 monomer. The protein was collected and the buffer was exchanged by dialysis against 2 L storage buffer (Tris pH 7.5 50 mM, EDTA 1 mM, DTT 1 mM, NaCl 50 mM, glycerol 20% v/v) for 18 h at 4°C. A second dialysis step was performed against 1 L storage buffer for 2 h at 4°C, to remove any remaining guanidinium chloride. The protein was quantified by UV spectrophotometry and stored at -80°C. The mutant proteins CD3657 K214R and CD3657 T215A were prepared in an identical manner. Expression of CD3657 E239A required 0.5mM IPTG induction with 16 h expression at 25°C. The purification protocol for CD3657 E239A was the same as described above for the other helicase proteins. Protein purity (all >95%) was estimated by SDS-PAGE electrophoresis and concentration was determined spectrophotometrically using extinction coefficients calculated using the ExPASy ProtParam tool (http://web.expasy.org/protparam).

Protein concentrations mentioned in this manuscript refer to concentration of the monomer of the protein.

Purification of the putative loader protein CD3654

C. difficile CD3654 was expressed from the pEVE-24 plasmid in *E. coli* BL21 (DE3). The growth medium consisting of 2xYT broth (1 L), kanamycin (30 µg/mL) and antifoam 204 (Sigma-Aldrich) was inoculated with a pre-culture (10 mL). The cell culture was incubated at 37°C with mechanical shaking, until an optical density (600 nm) of 0.62-0.65 was reached (after approximately 3 h). Protein expression was induced via the addition of IPTG (1 mM) and the culture was incubated at 30°C for 3 h. The cells were harvested by centrifugation (3000 g, 15 min, 4°C and

the resulting cell paste was stored at -80°C. The bacterial cell paste, prepared from 1L cell culture, was suspended in 25 mL TED50 buffer with PMSF (1 mM) and protease inhibitor cocktail (100 µL). The cells were lysed by sonication, clarified by centrifugation (40,000 g, 30 min, 4°C) and resulting supernatant separated from the cell debris using a 0.22 µm pore filter. Ammonium sulphate (7.32 g) was added slowly to the supernatant (25 mL) with stirring at 4°C, to achieve 49% saturation, the mixture was stirred for 10 min to give a white suspension. The protein pellet was collected by centrifugation (40,000 g, 30 min, 4°C) and washed with TED20 buffer (Tris pH 7.5 50 mM, EDTA 1 mM, DTT 1 mM, NaCl 20 mM) (2 x 4 mL). The precipitate was suspended in TED20 buffer (15 mL) with gentle mechanical shaking (30 min, 4°C). The suspension buffer was exchanged by dialysis against 1 L TED20 buffer for 2 h at 4°C, giving a solution with conductivity of 9.5 mS. The protein solution (~15 mL) was loaded onto combined 5 mL Q sepharose and 5 mL SP sepharose columns connected in series and equilibrated in TED20 buffer, the protein of interest eluted in the flow through. The collected protein was loaded onto a 5 mL heparin sepharose column equilibrated in TED20 buffer and eluted with a step to 15% TED1000 buffer. The collected protein was loaded onto a Hiload 26/60 Superdex 200 gel filtration column equilibrated in storage buffer (Tris pH 7.5 50 mM, EDTA 1 mM, DTT 1 mM, NaCl 50 mM, glycerol 10% v/v). The protein was quantified by UV spectrophotometry and stored at -80°C. The mutant proteins CD3654 K198R and CD3654 T199A were prepared in an identical manner. Protein purity (all >95%) was estimated by SDS-PAGE electrophoresis and concentration was determined spectrophotometrically using extinction coefficients calculated using the ExPASy ProtParam tool (http://web.expasy.org/protparam). Protein concentrations mentioned in this manuscript refer to concentration of the monomer of the protein.

ATP binding assay

The ATP binding assay [64] was performed in a 40 µl reaction containing 2 pmol wild type or mutant CD3657 protein, 0,65-41,6 nM [α-32P]ATP (3000Ci/mmol; Perkin Elmer) and ATP binding buffer (50 mM Tricine-KOH (pH 8.25), 0.5 mM magnesium acetate, 1 µM EDTA, 7 mM DTT, 0,007% Triton X-100 and 5% glycerol). The reaction was incubated for 15 min at room temperature. The samples were transferred to a Bio-Dot apparatus (BioRad) and passed through a 0.45 µm nitrocellulose membrane pre-soaked in wash buffer (50 mM Tricine-KOH (pH 8.25), 0.5 mM

magnesium acetate, 1 µM EDTA, 5 mM DTT, 10 mM ammonium sulphate, 0.005% Triton X-100 and 5% glycerol). The membrane was washed with cold buffer, first in the Bio-Dot apparatus and subsequently after removal from the Bio-Dot apparatus. The air-dried membrane was exposed to a storage phosphor screen for one hour. Binding was quantified by scanning the screen with the Typhoon 9410 imager and using QuantityOne software. Reactions without protein provided a background value that was subtracted.

Gel-filtration experiments

Self-interaction of the CD3657 and CD3654 proteins were studied in the presence and absence of ATP. In brief, purified CD3657 (or mutant) or CD3654 (or mutant) was incubated for 10 min at room temperature with MgCl2 (2 mM) in their storage buffer and ATP (1 mM). The mixture (500 µL) was loaded onto a Hiload 10/300 GL Superdex 200 analytical grade size exclusion column equilibrated in buffer A (with ATP: Tris pH 7.5 50 mM, EDTA 1 mM, DTT 1 mM, glycerol 10% v/v, MgCl2 2 mM, ATP 1 mM) or buffer B (Tris pH 7.5 50 mM, EDTA 1 mM, DTT 1 mM, glycerol 10% v/v, MgCl2 2 mM) at a flow rate of 0.5 mL/min. The elution profiles from each experiment were monitored at 280 nm and plotted as a function of the elution volume. Samples from fractions were analysed by SDS-PAGE and Coomassie Blue staining to verify the identity of the proteins. To assess interactions between CD3657 and CD3654, purified proteins were mixed in a 1:1 stoichiometry in the presence of MgCl2 (2 mM) and ATP (1 mM) and incubated for 10 min at room temperature. The mixture (500 µL) was loaded onto a Hiload 10/300 GL Superdex 200 analytical grade size exclusion column equilibrated in buffer A (Tris pH 7.5 50 mM, EDTA 1 mM, DTT 1 mM, glycerol 10% v/v, MgCl2 2 mM, ATP 1 mM) at a flow rate of 0.5 mL/min. For the experiments without ATP, buffer B was used (Tris pH 7.5 50 mM, EDTA 1 mM, DTT 1 mM, glycerol 10% v/v, MgCl2 2 mM).

Bacterial two-hybrid assays

To determine (self-)interaction, both expression constructs were subsequently transformed in to the *E. coli* reporter strain KDZif1ΔZ[65]. In order to control for background due to (possible) differences in expression of the constructs, single

expression plasmids (Zif- or ω fusion) containing the gene of interest were transformed into the reporter strain. After overnight incubation at 30˚C, three colonies per assay were cultured overnight in LB broth (30°C) in the presence of 1mM IPTG, tetracycline (selects Zif fusion plasmid) and/or ampicillin (selects ω fusion plasmid). Bacterial cells were permeabilized with SDS and chloroform and assayed for β- galactosidase activity according to the method of Miller [43]. In short, cells were diluted in Z buffer (60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$, 50 mM β-mercaptoethanol, pH7) to 1 ml and permeabilized with 50 µl 0,1% w/v SDS and 100 µl chloroform. After 5 minutes of equilibration, 200 µl of o-nitrophenyl-ß-D-galactopyranoside was added to each tube and incubated at room temperature until yellow colour developed. The reaction was stopped with 0,5 ml 1M Na_2CO_3 and measured at OD_{420} and OD_{550} to calculate the β- galactosidase activity in Miller Units. Experiments were performed in triplicate.

Helicase assays

Helicase activity was assayed by monitoring (and quantifying) the displacement of a radiolabelled (γ^{32}P-ATP) oligonucleotide oVP-1 (partially) annealed to the single stranded circular DNA m13mp18 (ssM13; Affymetrix) essentially as previously described [46]. In short, the 105-mer oligonucleotide was radiolabelled at the 5'end using γ^{32}ATP and T4 polynucleotide kinase (New England Biolabs) and subsequently purified through an S-200 mini-spin column (GE Healthcare). All reactions, containing 0.658 nM radiolabelled DNA substrate, were initiated by the addition of 2.5 mM ATP and carried out at 37˚C in buffer containing 20mM HEPES-NaOH (pH 7.5), 50 mM NaCl, 10 mM $MgCl_2$ and 1 mM DTT for various times. The reactions were terminated by adding 5x SDS-STOP buffer (100mM Tris pH8.0, 200mM EDTA, 2.5% (w/v) SDS, 50% (v/v) glycerol, 0.15% (w/v) bromophenol blue).

To investigate the effect of the putative helicase loader (CD3654) on the activity of the helicase (CD3657), the proteins were mixed in equimolar concentrations (1µM) and incubated for 10 minutes at 37˚C prior to adding reaction buffer. The buffer with CD3657 was preincubated for 5 mins before adding CD3654, incubated for 5 more mins after which the reaction was initiated with 2.5mM ATP (final concentration). Stop buffer was added to terminate the reactions (1% v/w SDS, 40 mM EDTA, 8% v/v glycerol, 0.1% w/v bromophenol blue). Reaction samples

(10 µl) were loaded on a 10% non-denaturing polyacrylamide gel, run in 1xTBE (89 mM Tris, 89 mM boric acid, 2 mM EDTA) at 150V, 40mA/gel for 60 mins. The gel was dried, scanned and analysed using a molecular imager and associated software (Biorad). Experiments were carried out in triplicate, and data analysis was performed using Prism 6 (GraphPad Software).

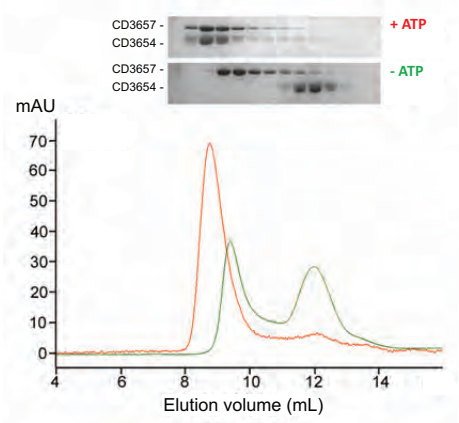
Supplemental information

The replicative helicase CD3657 of *Clostridium difficile* interacts with the putative loader protein CD3654

Erika van Eijk [1], Annemieke Friggen [1], Vasileios Paschalis [2], Matthew Green [2], Geoffrey S. Briggs [2], James Gibson [2], Panos Soultanas [2], Wiep Klaas Smits [1]

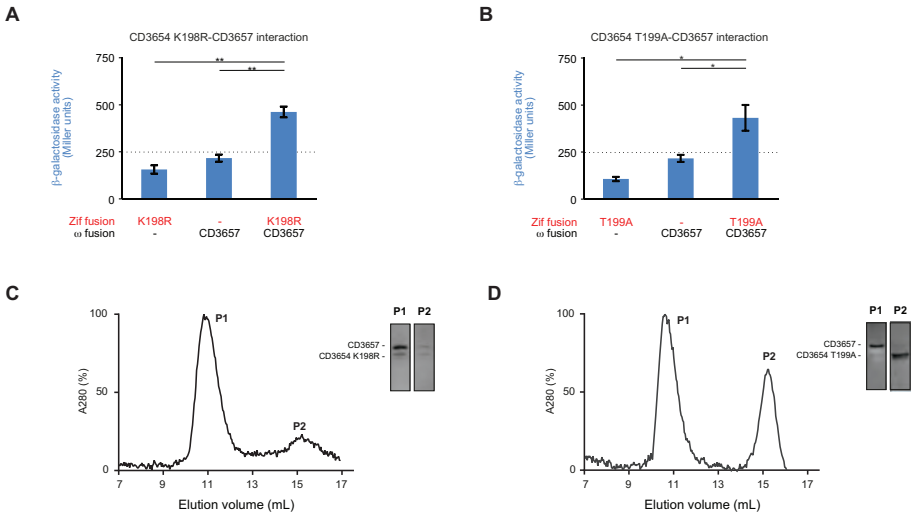
1 Department of Medical Microbiology, Leiden University Medical Center, Leiden, the Netherlands;

2 School of Chemistry, Center for Biomolecular Sciences, University of Nottingham, United Kingdom



Supplemental Figure 1. ATP dependent interaction of the helicase CD3657 and the putative loader CD3654 at high concentrations of proteins.

The helicase CD3657 and the putative loader CD3654 interact in an ATP-dependent manner. Analytical gel filtration was performed in the presence (red) of absence (green) of 1mM ATP on a Hiload 10/300 GL Superdex 200 analytical grade size exclusion column. Inset shows a Coomassie-stained SDS-PAGE gels of of sampled fractions taken during both gel filtration experiments.



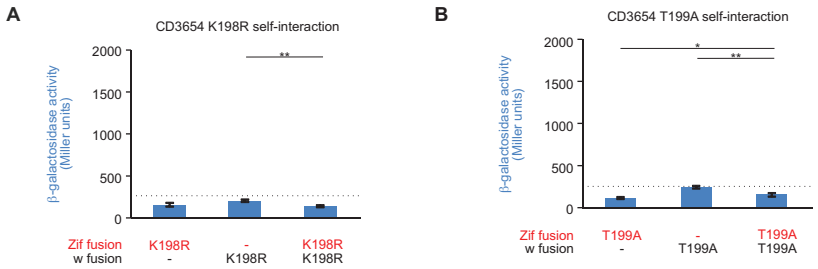
Supplemental Figure 2. Walker A mutants of the putative loader protein CD3654 retain the ability to interact with CD3657.

A. Bacterial two hybrid analysis of the interaction between CD3657 and CD3654 K198R.

B. Bacterial two hybrid analysis of the interaction between CD3657 and CD3654 T199A. Bar graphs in A–B indicate average values and error-bars indicate standard deviation of the measurements (n=3). Dashed line indicates the maximum background level of β-galactosidase expression observed in our experimental set-up. Significance was determined using the Student's t-test (* $p < 0.05$, ** $p < 0.001$).

C. Analytical gel filtration analysis of the interaction between CD3657 and CD3654 K198R.

D. Analytical gel filtration analysis of the interaction between CD3657 and CD3654 T199A. Analytical gel filtration was performed in buffer B (see Methods) with 3.10 uM proteins in the presence of 1mM ATP. Inset in C–D shows Coomassie-stained SDS-PAGE gels of the numbered peak fractions.



Supplemental Figure 3. Walker A mutants of CD3654 are monomeric.

A. Bacterial two-hybrid analysis of the self-interaction of CD3654 K198R.

B . Bacterial two hybrid analysis of the self-interaction of CD3654 T199A.
Bar graphs in A-B indicate average values and error-bars indicate standard deviation of the measurements (n=3). Significance was determined using the Student's t-test (* p <0.05, ** p <0.001).

Supplemental Table 1. Oligonucleotides used in this study

Name	Sequence (5' – 3')	Description
oAF-26	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAATGAATGAGGATAAAATAAGAAAAATAC	Forward CD3654 (B2H) ★
oAF-27	GGGGACCACTTTGTACAAGAAAGCTGGGTCTTTAAATCTTTCCCATCTCAAATCATC	Reverse CD3654 (B2H)
oAF-28	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAATGGAAGATATGACGAGAATTCCTC	Forward CD3657 (B2H)
oAF-29	GGGGACCACTTTGTACAAGAAAGCTGGGTCTAGGTCTCTCAACTTATCCCC	Reverse CD3657 (B2H)
oEVE-4	GCTTCCATGGATGAGGATAAAATAAGAAAAATACTTGC	Forward CD3654 pET
oEVE-6	GACGCTCGAGTTATTTAAATCTTTCCCATCTCAAATCATCTCC	Reverse CD3654 pET
oWKS-1185	GTGATTCATATGGAAGATATGACGAG	Forward CD3657 pET
oWKS-1367	CCGCTCGAGTTATAGGTCTCTCAACTTATCCCC	Reverse CD3657 pET
oWKS-1272	GGTTCTACTGGACTAGGAAGGACCTATATGTGCAATTG	Forward CD3654 K198R (QC) ★★
oWKS-1273	CAATTGCACATATAGGTCCTTCCTAGTCCAGTAGAACC	Reverse CD3654 K198R (QC)
oWKS-1274	GGTTCTACTGGACTAGGAAAGGCCTATATGTGCAATTGTATTG	Forward CD3654 T199A (QC)
oWKS-1275	CAATACAATTGCACATATAGGCCTTTCCTAGTCCAGTAGAACC	Reverse CD3654 T199A (QC)
oEVE-69	CTTGTTTGATTGTGATTTATTAATAATACAAGACCTTGGAAC	Forward CD3654 D258Q (QC)
oEVE-70	GTTCCAAGGTCTTGTATTATTAATAAATCACAATCAAACAAG	Reverse CD3654 D258Q (QC)
oWKS-1276	GCTAGACCAGCAATGGGTAGAACTGCCTTTTCTTTAAAC	Forward CD3657 K214R (QC)
oWKS-1277	GTTTAAAGAAAAGGCAGTTCTACCCATTGCTGGTCTAGC	Reverse CD3657 K214R (QC)
oWKS-1278	CTAGACCAGCAATGGGTAAAGCTGCCTTTTCTTTAAACTTGG	Forward CD3657 T215A (QC)
oWKS-1279	CCAAGTTTAAAGAAAAGGCAGCTTTACCCATTGCTGGTCTAG	Reverse CD3657 T215A (QC)
oVP-1	CACACACACACACACACACACACACACACACACACACACACACACACACACACACACACCCC CTTTAAAAAAAAAAAGCCAAAAGCAGTGCCAAGCTTGCATGCC	Helicase assay

★ Bacterial two-hybrid (B2H); ★★ QuikChange (QC)

Supplemental Table 2. **Plasmids used in this study**

Plasmid	Description	Reference
pEVE24	pET28b-CD3654	This study
pEVE59	pET28b-CD3654 K198R	This study
pEVE60	pET28b-CD3654 T199A	This study
pEVE-203	pET28b-CD3654-D258Q	This study
pEVE87	pET21b-CD3657	This study
pEVE90	pET21b-CD3657 K214R	This study
pEVE92	pET21b-CD3657 T215A	This study
pEVE118	pENTRY-CD3654	This study
pEVE120	pENTRY-CD3657	This study
pEVE122	pKEK1286-CD3654	This study
pEVE123	pKEK1287-CD3654	This study
pEVE124	pKEK1286-CD3657	This study
pEVE125	pKEK1287-CD3657	This study
pEVE167	pKEK1286-CD3654 K198R	This study
pEVE168	pKEK1286-CD3654 T199A	This study
pEVE169	pKEK1287-CD3654 K198R	This study
pEVE170	pKEK1287-CD3654 T199A	This study
pEVE171	pKEK1286-CD3657 K214R	This study
pEVE172	pKEK1286-CD3657 T215A	This study
pEVE173	pKEK1287-CD3657 K214R	This study
pEVE174	pKEK1287-CD3657 T215A	This study

References

1 Battistuzzi, F. U., Feijao, A. & Hedges, S. B. A genomic timescale of prokaryote evolution: insights into the origin of methanogenesis, phototrophy, and the colonization of land. *BMC evolutionary biology* **4**, 44, doi:10.1186/1471-2148-4-44 (2004).

2 Davey, M. J. & O'Donnell, M. Replicative helicase loaders: ring breakers and ring makers. *Current biology : CB* **13**, R594-596 (2003).

3 Li, Y. & Araki, H. Loading and activation of DNA replicative helicases: the key step of initiation of DNA replication. *Genes to cells : devoted to molecular & cellular mechanisms* **18**, 266-277, doi:10.1111/gtc.12040 (2013).

4 Mott, M. L. & Berger, J. M. DNA replication initiation: mechanisms and regulation in bacteria. *Nature reviews. Microbiology* **5**, 343-354, doi:10.1038/nrmicro1640 (2007).

5 Robinson, A., Causer, R. J. & Dixon, N. E. Architecture and conservation of the bacterial DNA replication machinery, an underexploited drug target. *Current drug targets* **13**, 352-372 (2012).

6 Soultanas, P. Loading mechanisms of ring helicases at replication origins. *Molecular microbiology* **84**, 6-16, doi:10.1111/j.1365-2958.2012.08012.x (2012).

7 Kaguni, J. M. Replication initiation at the Escherichia coli chromosomal origin. *Current opinion in chemical biology* **15**, 606-613, doi:10.1016/j.cbpa.2011.07.016 (2011).

8 Remus, D. & Diffley, J. F. Eukaryotic DNA replication control: lock and load, then fire. *Current opinion in cell biology* **21**, 771-777, doi:10.1016/j.ceb.2009.08.002 (2009).

9 Arias-Palomo, E., O'Shea, V. L., Hood, I. V. & Berger, J. M. The bacterial DnaC helicase loader is a DnaB ring breaker. *Cell* **153**, 438-448, doi:10.1016/j.cell.2013.03.006 (2013).

10 Davey, M. J., Fang, L., McInerney, P., Georgescu, R. E. & O'Donnell, M. The DnaC helicase loader is a dual ATP/ADP switch protein. *The EMBO journal* **21**, 3148-3159, doi:10.1093/emboj/cdf308 (2002).

11 Kornberg, A. & Baker, T. A. *DNA Replication*. (University Science Books, 2005).

12 Bruand, C., Farache, M., McGovern, S., Ehrlich, S. D. & Polard, P. DnaB, DnaD and DnaI proteins are components of the Bacillus subtilis replication restart primosome. *Molecular microbiology* **42**, 245-255 (2001).

13 Bruand, C. *et al.* Functional interplay between the Bacillus subtilis DnaD and DnaB proteins essential for initiation and re-initiation of DNA replication. *Molecular microbiology* **55**, 1138-1150, doi:10.1111/j.1365-2958.2004.04451.x (2005).

14 Smits, W. K., Goranov, A. I. & Grossman, A. D. Ordered association of helicase loader proteins with the Bacillus subtilis origin of replication in vivo. *Molecular microbiology* **75**, 452-461, doi:10.1111/j.1365-2958.2009.06999.x (2010).

15 Briggs, G. S., Smits, W. K. & Soultanas, P. Chromosomal replication initiation machinery of low-G+C-content Firmicutes. *Journal of bacteriology* **194**, 5162-5170, doi:10.1128/jb.00865-12 (2012).

16 Bazin, A., Cherrier, M. V., Gutsche, I., Timmins, J. & Terradot, L. Structure and primase-mediated activation of a bacterial dodecameric replicative helicase. *Nucleic acids research* **43**, 8564-8576, doi:10.1093/nar/gkv792 (2015).

17 Ioannou, C., Schaeffer, P. M., Dixon, N. E. & Soultanas, P. Helicase binding to DnaI exposes a cryptic DNA-binding site during helicase loading in Bacillus subtilis. *Nucleic acids research* **34**, 5247-5258, doi:10.1093/nar/gkl690 (2006).

18 Velten, M. *et al.* A two-protein strategy for the functional loading of a cellular replicative DNA helicase. *Molecular cell* **11**, 1009-1020 (2003).

19 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, doi:10.1093/nar/gkq465 (2010).

20 Schneider, S., Zhang, W., Soultanas, P. & Paoli, M. Structure of the N-terminal oligomerization domain of DnaD reveals a unique tetramerization motif and provides insights into scaffold formation. *Journal of molecular biology* **376**, 1237–1250, doi:10.1016/j.jmb.2007.12.045 (2008).

21 Zhang, W., Allen, S., Roberts, C. J. & Soultanas, P. The Bacillus subtilis primosomal protein DnaD untwists supercoiled DNA. *Journal of bacteriology* **188**, 5487–5493, doi:10.1128/jb.00339-06 (2006).

22 Zhang, W. *et al.* The Bacillus subtilis DnaD and DnaB proteins exhibit different DNA remodelling activities. *Journal of molecular biology* **351**, 66–75, doi:10.1016/j.jmb.2005.05.065 (2005).

23 Zhang, W. *et al.* Single-molecule atomic force spectroscopy reveals that DnaD forms scaffolds and enhances duplex melting. *Journal of molecular biology* **377**, 706–714, doi:10.1016/j.jmb.2008.01.067 (2008).

24 Caruthers, J. M. & McKay, D. B. Helicase structure and mechanism. *Current opinion in structural biology* **12**, 123–133 (2002).

25 Hanson, P. I. & Whiteheart, S. W. AAA+ proteins: have engine, will work. *Nature reviews. Molecular cell biology* **6**, 519–529, doi:10.1038/nrm1684 (2005).

26 Mott, M. L., Erzberger, J. P., Coons, M. M. & Berger, J. M. Structural synergy and molecular crosstalk between bacterial helicase loaders and replication initiators. *Cell* **135**, 623–634, doi:10.1016/j.cell.2008.09.058 (2008).

27 Duderstadt, K. E., Chuang, K. & Berger, J. M. DNA stretching by bacterial initiators promotes replication origin opening. *Nature* **478**, 209–213, doi:10.1038/nature10455 (2011).

28 Erzberger, J. P. & Berger, J. M. Evolutionary relationships and structural mechanisms of AAA+ proteins. *Annual review of biophysics and biomolecular structure* **35**, 93–114, doi:10.1146/annurev.biophys.35.040405.101933 (2006).

29 Singleton, M. R., Dillingham, M. S. & Wigley, D. B. Structure and mechanism of helicases and nucleic acid translocases. *Annual review of biochemistry* **76**, 23–50, doi:10.1146/annurev.biochem.76.052305.115300 (2007).

30 Snider, J., Thibault, G. & Houry, W. A. The AAA+ superfamily of functionally diverse proteins. *Genome biology* **9**, 216, doi:10.1186/gb-2008-9-4-216 (2008).

31 Berger, J. M. SnapShot: nucleic acid helicases and translocases. *Cell* **134**, 888–888.e881, doi:10.1016/j.cell.2008.08.027 (2008).

32 Itsathitphaisarn, O., Wing, R. A., Eliason, W. K., Wang, J. & Steitz, T. A. The hexameric helicase DnaB adopts a nonplanar conformation during translocation. *Cell* **151**, 267–277, doi:10.1016/j.cell.2012.09.014 (2012).

33 Ilyina, T. V., Gorbalenya, A. E. & Koonin, E. V. Organization and evolution of bacterial and bacteriophage primase-helicase systems. *Journal of molecular evolution* **34**, 351–357 (1992).

34 Leipe, D. D., Aravind, L., Grishin, N. V. & Koonin, E. V. The bacterial replicative helicase DnaB evolved from a RecA duplication. *Genome research* **10**, 5–16 (2000).

35 Rupnik, M., Wilcox, M. H. & Gerding, D. N. Clostridium difficile infection: new developments in epidemiology and pathogenesis. *Nature reviews. Microbiology* **7**, 526–536, doi:10.1038/nrmicro2164 (2009).

36 Smits, W. K., Lyras, D., Lacy, D. B., Wilcox, M. H. & Kuijper, E. J. Clostridium difficile infection. *Nature reviews. Disease primers* **2**, 16020, doi:10.1038/nrdp.2016.20 (2016).

37 Bruand, C., Ehrlich, S. D. & Janniere, L. Primosome assembly site in Bacillus subtilis. *The EMBO journal* **14**, 2642–2650 (1995).

38 Sebaihia, M. *et al.* The multidrug-resistant human pathogen Clostridium difficile has a highly mobile, mosaic genome. *Nature genetics* **38**, 779–786, doi:10.1038/ng1830 (2006).

39 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, doi:10.1111/j.1365-2958.2004.04091.x (2004).

40 Soultanas, P. A functional interaction between the putative primosomal protein DnaI and the main replicative DNA helicase DnaB in Bacillus. *Nucleic acids research* **30**, 966–974 (2002).

41 Brouwer, M. S., Warburton, P. J., Roberts, A. P., Mullany, P. & Allan, E. Genetic organisation, mobility and predicted functions of genes on integrated, mobile genetic elements in sequenced strains of Clostridium difficile. *PloS one* **6**, e23014, doi:10.1371/journal.pone.0023014 (2011).

42 Karna, S. L. *et al.* A bacterial two-hybrid system that utilizes Gateway cloning for rapid screening of protein-protein interactions. *BioTechniques* **49**, 831–833, doi:10.2144/000113539 (2010).

43 Miller, J. H. in *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory Press, 1972).

44 Soultanas, P. & Wigley, D. B. Site-directed mutagenesis reveals roles for conserved amino acid residues in the hexameric DNA helicase DnaB from Bacillus stearothermophilus. *Nucleic acids research* **30**, 4051–4060 (2002).

45 Biswas, E. E., Chen, P. H. & Biswas, S. B. Modulation of enzymatic activities of Escherichia coli DnaB helicase by single-stranded DNA-binding proteins. *Nucleic acids research* **30**, 2809–2816 (2002).

46 Rannou, O. *et al.* Functional interplay of DnaE polymerase, DnaG primase and DnaC helicase within a ternary complex, and primase to polymerase hand-off during lagging strand DNA replication in Bacillus subtilis. *Nucleic acids research* **41**, 5303–5320, doi:10.1093/nar/gkt207 (2013).

47 Grainger, W. H., Machon, C., Scott, D. J. & Soultanas, P. DnaB proteolysis in vivo regulates oligomerization and its localization at oriC in Bacillus subtilis. *Nucleic acids research* **38**, 2851–2864, doi:10.1093/nar/gkp1236 (2010).

48 Bujalowski, W., Klonowska, M. M. & Jezewska, M. J. Oligomeric structure of Escherichia coli primary replicative helicase DnaB protein. *The Journal of biological chemistry* **269**, 31350–31358 (1994).

49 Galletto, R., Jezewska, M. J. & Bujalowski, W. Interactions of the Escherichia coli DnaB helicase hexamer with the replication factor the DnaC protein. Effect of nucleotide cofactors and the ssDNA on protein-protein interactions and the topology of the complex. *Journal of molecular biology* **329**, 441–465 (2003).

50 Strycharska, M. S. *et al.* Nucleotide and partner-protein control of bacterial replicative helicase structure and function. *Molecular cell* **52**, 844–854, doi:10.1016/j.molcel.2013.11.016 (2013).

51 Dembek, M. *et al.* High-throughput analysis of gene essentiality and sporulation in Clostridium difficile. *mBio* **6**, e02383, doi:10.1128/mBio.02383-14 (2015).

52 Chodavarapu, S. & Kaguni, J. M. Replication Initiation in Bacteria. *The Enzymes* **39**, 1–30, doi:10.1016/bs.enz.2016.03.001 (2016).

53 Liu, B., Eliason, W. K. & Steitz, T. A. Structure of a helicase–helicase loader complex reveals insights into the mechanism of bacterial primosome assembly. *Nature communications* **4**, 2495, doi:10.1038/ncomms3495 (2013).

54 Gao, D. & McHenry, C. S. tau binds and organizes Escherichia coli replication proteins through distinct domains. Domain IV, located within the unique C terminus of tau, binds the replication fork, helicase, DnaB. *The Journal of biological chemistry* **276**, 4441–4446, doi:10.1074/jbc.M009830200 (2001).

55 Afonso, J. P. *et al.* Insights into the structure and assembly of the Bacillus subtilis clamp–loader complex and its interaction with the replicative helicase. *Nucleic acids research* **41**, 5115–5126, doi:10.1093/nar/gkt173 (2013).

56 Haroniti, A. *et al.* The clamp–loader–helicase interaction in Bacillus. Atomic force microscopy reveals the structural organisation of the DnaB–tau complex in Bacillus. *Journal of molecular biology* **336**, 381–393 (2004).

57 Bailey, S., Eliason, W. K. & Steitz, T. A. Structure of hexameric DnaB helicase and its complex with a domain of DnaG primase. *Science (New York, N.Y.)* **318**, 459–463, doi:10.1126/science.1147353 (2007).

58 Chang, P. & Marians, K. J. Identification of a region of Escherichia coli DnaB required for functional interaction with DnaG at the replication fork. *The Journal of biological chemistry* **275**, 26187–26195, doi:10.1074/jbc.M001800200 (2000).

59 Stordal, L. & Maurer, R. Defect in general priming conferred by linker region mutants of Escherichia coli dnaB. *Journal of bacteriology* **178**, 4620–4627 (1996).

60 Thirlway, J. & Soultanas, P. In the Bacillus stearothermophilus DnaB–DnaG complex, the activities of the two proteins are modulated by distinct but overlapping networks of residues. *Journal of bacteriology* **188**, 1534–1539, doi:10.1128/jb.188.4.1534–1539.2006 (2006).

61 Makowska–Grzyska, M. & Kaguni, J. M. Primase directs the release of DnaC from DnaB. *Molecular cell* **37**, 90–101, doi:10.1016/j.molcel.2009.12.031 (2010).

62 Hussain, H. A., Roberts, A. P. & Mullany, P. Generation of an erythromycin–sensitive derivative of Clostridium difficile strain 630 (630Deltaerm) and demonstration that the conjugative transposon Tn916DeltaE enters the genome of this strain at multiple sites. *Journal of medical microbiology* **54**, 137–141, doi:10.1099/jmm.0.45790–0 (2005).

63 van Eijk, E. *et al.* Complete genome sequence of the Clostridium difficile laboratory strain 630Deltaerm reveals differences from strain 630, including translocation of the mobile element CTn5. *BMC genomics* **16**, 31, doi:10.1186/s12864–015–1252–7 (2015).

64 Mizushima, T. *et al.* Site–directed mutational analysis for the ATP binding of DnaA protein. Functions of two conserved amino acids (Lys–178 and Asp–235) located in the ATP–binding domain of DnaA protein in vitro and in vivo. *The Journal of biological chemistry* **273**, 20847–20851 (1998).

65 Vallet–Gely, I., Donovan, K. E., Fang, R., Joung, J. K. & Dove, S. L. Repression of phase–variable cup gene expression by H–NS–like proteins in Pseudomonas aeruginosa. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 11082–11087, doi:10.1073/pnas.0502663102 (2005).