Bacterial DNA replication initiation:

Structural and functional analysis of the master initiator DnaA

Daniel Stevens

Thesis submitted in partial fulfilment of the requirements of the regulation for the degree of Doctor of Philosophy

Newcastle University

Faculty of Medical Sciences

Institute of Cell and Molecular Biosciences

October 2019

Abstract

The essential process of DNA replication begins with initiator proteins binding to origins of replication and triggering DNA synthesis. The highly conserved bacterial master initiator protein, DnaA, performs several key activities at the bacterial origin (*oriC*) to initiate replication. DnaA binds specifically to *oriC* and assembles into a filament that engages and stretches a single DNA strand to induce duplex unwinding. Subsequently, DnaA recruits a loading complex that deposits the replicative helicases around single DNA strands.

In this thesis I have investigated the molecular mechanisms underpinning some of the essential activities of DnaA in the model organism *Bacillus subtilis*.

Using a chimeric DnaA system I was able to identify several activities required for origin unwinding by DnaA bound to a specific DnaA-box located upstream of the site of unwinding. This result suggested that the protein binding here is directly involved in unwinding the DNA duplex, and the likely role of the upstream region is to increase the local DnaA concentration at the site of unwinding.

To unwind *oriC*, DnaA engages and stretches a specific DNA strand with a recently identified repeating tri-nucleotide motif, termed the DnaA-trios, providing the specific sequence. Utilising an inducible heterologous replication initiation system I determined which DnaA residues from a region implicated in ssDNA binding were essential *in vivo*. Using recombinant DnaA protein variants, two isoleucine residues were determined to be required for forming filaments on ssDNA and unwinding the DNA duplex *in vitro*. Further work is required to determine if these residues are required for the specific interaction with DnaA-trios or more generally for DNA binding/unwinding.

A range of essential residues required for the interaction between DnaA and the firmicute specific initiation accessory protein DnaD, the first step in helicase recruitment, were identified. The DnaA residues overlap with a binding site for the developmental regulator, SirA, a developmentally expressed inhibitor of DNA replication initiation. This suggested that SirA functions by blocking the interaction between DnaA and DnaD, preventing helicase loading. I found that SirA inhibits the interaction of DnaA with DnaD, providing a molecular mechanism for this SirA activity and revealing, for the first time, an endogenous system for regulating helicase recruitment in bacteria.

Acknowledgments

I would like to thank Professor Heath Murray for giving me the opportunity to work on this project I have thoroughly enjoyed. I could not have asked for better help, guidance and encouragement over the last 3 years. I've learnt a lot, not just about biology and techniques but about how to be a better scientist.

I also wish to thank all members of the Murray lab past and present as I could not have wished for a better group of people to work alongside. Each and every one of them has helped me in ways big and small over the last few years and I thank them for the memorable times spent outside of work also. A big thank you also goes to all my friends and colleagues on the 4th floor of the CBCB. They have made these last few years incredibly enjoyable and thank you especially to Grace and Jess for providing plentiful opportunities to forget about the PhDs for a while.

A special thank you to Fran, Tom, Julie, Sean and the other support staff who have helped make my time here easier and enabled me to focus on my science.

Lastly thank you to all my friends and family back home. Thank you to Liam, Carla, Maisie, Freddie and Toby for their continuous love and support. Thank you to my boyfriend Mark for pushing me to be my best, and without whom I wouldn't have made it to Newcastle. Finally thank you to my amazing parents, Darren and Julie, whose support, encouragement and confidence in me throughout my whole life have bought me to where I am today.

Publications

Martin, E., Williams, H.E.L., Pitoulias, M., Stevens, D., Winterhalter, C., Craggs, T.D., Murray, H., Searle, M.S. and Soultanas, P. (2019) 'DNA replication initiation in Bacillus subtilis: structural and functional characterization of the essential DnaA-DnaD interaction', *Nucleic Acids Res*, 47(4), pp. 2101-2112.

Richardson, T.T., Stevens, D., Pelliciari, S., Harran, O., Sperlea, T. and Murray, H. (2019) 'Identification of a basal system for unwinding a bacterial chromosome origin', *Embo j*, p. e101649.

Contents

List of Tables

List of Figures

List of Abbreviations

3C – Chromatin Conformation Capture

AAA+ – ATPase Associated with various cellular Activities

ACS – ARS Consensus Sequence

ARS – Autonomously Replicating Sequences

ADP – Adenosine triphosphate

ATP – Adenosine diphosphate

bp – Base Pair(s)

BHQ – Black Hole Quencher

BMOE – Bismaleimido ethane

BTH – Bacterial Two-hybrid

cAMP – cyclic AMP

CAP – Catabolite activator protein

Cdc6 – Cell division cycle 6

Cdc45 – Cell division cycle 45

Cdt1 – Chromosome licensing and DNA replication factor 1

Chi – Chimera/Chimeric

ChIP – Chromatin Immunoprecipitation

CMG – (Cdc45-MCM-GINS)

CTD – C-terminal Domain

Ctf4 – Chromosome transmission fidelity

CtrA – Cell cycle Transcriptional Regulator A

Dam – DNA Adenine Methylase

DBC – DnaA-box Cluster

DBD – DNA Binding Domain

DBR – DnaA Binding Region

DiaA – DnaA initiator-associating protein

DMSO – Dimethyl sulfoxide

DNA – Deoxyribonucleic acid

DNAP – DNA polymerase

dsDNA – double stranded DNA

DUE – DNA Unwinding Element

EDTA – Ethylenediaminetetraacetic acid

EM – Electron Microscopy

EMSA – Electrophoretic shift assay

FP – Fluorescence polarisation

FtsZ – Filament temperature-sensitive mutant Z

HTH – Helix-turn-helix

IHF – Integration Host Factor

IPTG – Isopropyl β-D-1-thiogalactopyranoside

ISM – Initiator Specific Motif

LB – Luria-Bertani

MCM – Minichromosome Maintenance

MD – Middle Domain

NA – Nutrient agar

NMR – Nuclear Magnetic Resonance

NTD – N-terminal Domain

OBP – Origin Binding Proteins

OGRE – Origin G-rich repeat element

ORB – Origin Recognition Box

ORC – Origin Recognition Complex

PBS – Phosphate buffer saline

PBST – PBSTween

PCNA – Proliferating Cell Nuclear Antigen

PCR – Polymerase Chain Reaction

Pol – Polymerase

pre-IC – pre-Initiation Complex

pre-RC – pre-Replicative Complex

RFC – Replication Factor C

RIDA – Regulatory Inactivation of DnaA

RNA – Ribonucleic acid

RPA – Replication Protein A

RTP – Replication Termination Protein

SDS-PAGE – Sodium dodecyl sulphate-polyacrylamide gel electrophoresis

SMM – Spizizen Minimal Medium

SPR – Surface Plasmon Resonance

SSB – Single-strand Binding

SSC – Saline sodium citrate

ssDNA – single stranded DNA

TBE – Tris-borate-EDTA

Tus – Terminus Utilisation Substance

WH – Winged Helix

Chapter 1

Introduction

The accurate transmission of genetic material is essential for the viability and proliferation of cells. This transmission is dependent upon the accurate replication of DNA, a fundamental process across all three domains of life.

The process of accurate DNA replication can be broken down into three distinct subprocesses; initiation, where the AAA+ (ATPase Associated with various cellular Activities) initiator proteins assemble at the origin and load the helicases, elongation, where the replication machinery synthesises new DNA strands and termination, where DNA synthesis is halted. A simplified overview of this, and subsequent processes, for a bacterium possessing a single chromosome with a single origin is outlined in Figure 1.1 (O'Donnell *et al.*, 2013).

Although there is a variety of molecular mechanisms utilised to coordinate and regulate genome duplication across domains the basic machinery is similar. Throughout all life conserved proteins possessing AAA+ domains assemble into multiprotein complexes at specialised genomic sites. At these sites, known as origins of replication, these proteins direct the loading of the replicative helicase and the initial unwinding of the DNA duplex. Helicase loading and subsequent activation promotes the assembly of the replication machinery at two distinct replication forks from where DNA synthesis proceeds. Finally DNA replication is terminated and the newly synthesised DNA is arranged for accurate segregation prior to cell division (O'Donnell *et al.*, 2013; Jha *et al.*, 2016).

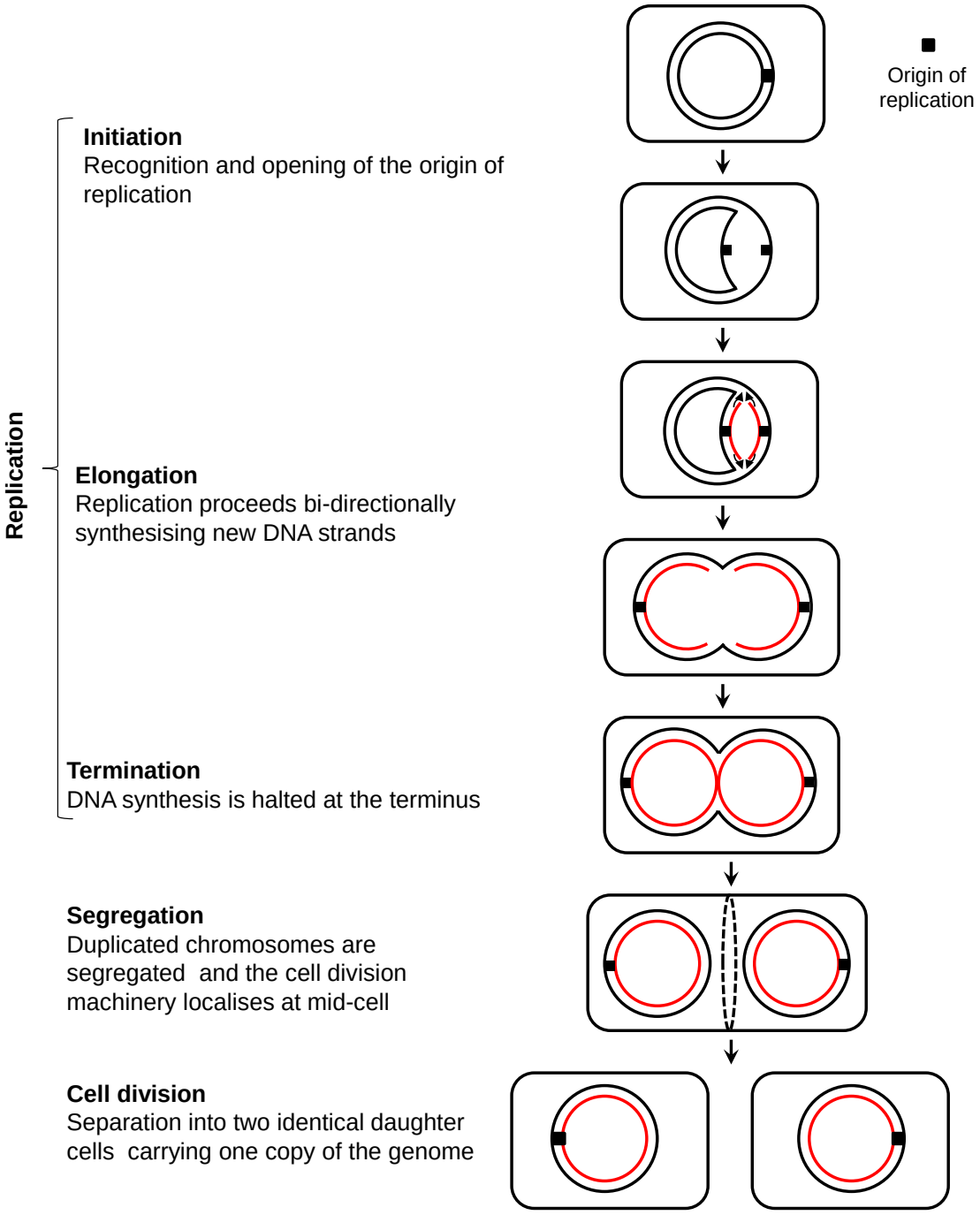


Figure 1.1. The bacterial cell cycle. A simplified overview of the cell cycle for a bacterium possessing a single chromosome with a single origin. The steps of chromosome replication, segregation and cell division are highlighted. The original chromosome is shown in black with newly synthesised DNA shown in red and the origin of replication represented as a black box.

1.1. DNA Replication in Bacteria, Archaea and Eukaryotes

1.1.1. Initiation of DNA Replication

The first step in replicating DNA is the separation of the duplex DNA into single strands. During the initiation of replication the stable double helix is actively opened at specific sites for a sufficient duration and to a sufficient length to allow for the loading of the replicative helicase. Following loading the helicase extends the open complex and mediates the loading of primase and the replication machinery required for DNA duplication (Jha *et al.*, 2016).

The specific chromosomal sites where initiator proteins bind prior to opening are termed origins of replication. Bacteria commonly contain a single circular chromosome with a single origin at which two replication forks assemble before replication proceeds bi-directionally (Figure 1.2.A). There are a minority of bacterial genomes, however, that are multipartite containing large secondary chromosomes or chromids comparable in size to the main chromosome and carrying essential genes (Pinto *et al.*, 2012; diCenzo and Finan, 2017). For example the 4 Mb genome of *Vibrio cholera* is split between two chromosmes of 2.9 Mb and 1.1 Mb respectively (Heidelberg *et al.*, 2000).

Similar to bacteria, archaea possess circular chromosomes, with some species having extrachromosomal elements. Many of the archaea characterised have multiple origins, although some possess just a single one (Figure 1.2.A) (Creager *et al.*, 2015; Ausiannikava *et al.*, 2018). Interestingly it has been demonstrated that archaea can live without origins, initiating replication via homologous recombination at dispersed sites throughout the chromosome. Deleting the origins in *Haloferax volcanii* appeared to give a fitness advantage with cells growing significantly faster without origins (Hawkins *et al.*, 2013).

Eukaryotes generally have larger more complicated genomes ranging from 12 Mbp in *Saccharomyces cerevisiae* up to 3 Gbp in humans. The much larger genomes necessitate that replication starts from several hundred to thousands of origins in parallel to ensure complete duplication during a single cell cycle (Figure 1.2.A) (Ekundayo and Bleichert, 2019). For example human chromosome 1 is 250 Mb and would take ~50 days to replicate from a single origin, compared to the 24 hour eukaryotic cell cycle (O'Donnell *et al.*, 2013).

The origins of replication in bacteria, termed *oriC*, are defined by a series of repeating motifs that form binding sites for the master initiator protein, DnaA. These sequences are termed DnaA-boxes, and vary in number, arrangement and even sequence throughout the bacterial domain. The origins of several bacterial species also contain further binding sites for accessory or regulatory proteins (such as IHF and SeqA in *Escherichia coli*, CtrA in *Caulobacter crescentus* and Spo0A in *Bacillus subtilis*). Proximal to these is a region of AT-rich repeats, the site of DNA strand separation often termed the DNA unwinding element (DUE). These features are highlighted in Figure 1.2.B for the *E. coli* origin (Wolański *et al.*, 2015).

As mentioned some bacterial species possess large secondary genomic elements. These elements do not utilise the DnaA/*oriC* system for initiating replication and instead use plasmid-type mechanisms (Pinto *et al.*, 2012). Most of our current understanding of multichromosome bacteria comes from *V. cholera*. The secondary chromosome origin (*oriC2*) of *V. cholera* is laid out in Figure 1.2.B. *oriC2* possess an origin containing iterons of various lengths that are binding sites for the initiator protein RctB. The origin also contains a binding site for DnaA, IHF and an AT-rich DUE (Fournes *et al.*, 2018).

Archaeal origins vary in number between species and even within the same organism. A series of conserved repeats close to the genes of the archeal initiator protein ORC (origin recognition complex) are considered to form origins. Some of these repeats are longer and located either side of an AT-rich region, expected to be the site of unwinding, and as such are referred to as a DUE. These extended repeat sequences are termed origin recognition boxes (ORB). This layout has been found among several archaeal species, although the number of ORB sequences vary (Wigley, 2009; Wu *et al.*, 2014). The layout of the archaeal origin of *Aeropyrum pernix* is highlighted in Figure 1.2.B.

Eukaryotic origins are not typically defined by the DNA sequence, appearing to be more defined by chromatin organisation instead, with the eukaryotic origin recognition complex (ORC) generally not recognising a specific sequence (Leonard and Méchali, 2013). DNA accessibility is suggested to be a major determinant for eukaryotic initiation sites, with ORC having a preference for nucleosome-free chromatin (Sequeira-Mendes and Gómez, 2012). As such many origins are associated with genomic regions where activities that would allow initiator proteins access to the DNA

are occurring. For example active promoter regions, as AT-rich or G-rich regions do not favour nucleosome formation (Leonard and Méchali, 2013).

It is suggested that for some eukaryotes a DNA structure formed by specific base positioning may be more important for specifying initiation start sites rather than a specific DNA sequence. For example the majority of mouse and human cell origins contain an origin G-rich repeat element (OGRE) that can form G quadruplexes in which guanine hydrogen bonds into a four-stranded DNA structure (Cayrou *et al.*, 2012) .

Unlike most eukaryotes, *S. cerevisiae* (and related yeast species) possess identifiable origins formed of specific sequences. These elements within the replication origin regions are called autonomously replicating sequences (ARS) and are formed of three domains (A, B and C). The A domain contains an ARS consensus sequence (ACS) required for initiator binding. The B domains are numerous short motifs that may serve as a DUE and contain binding sites for the initiator protein and helicase. The C domain is a site for transcription and regulatory factor binding (Figure 1.2.B) (Leonard and Méchali, 2013; Dao *et al.*, 2018).

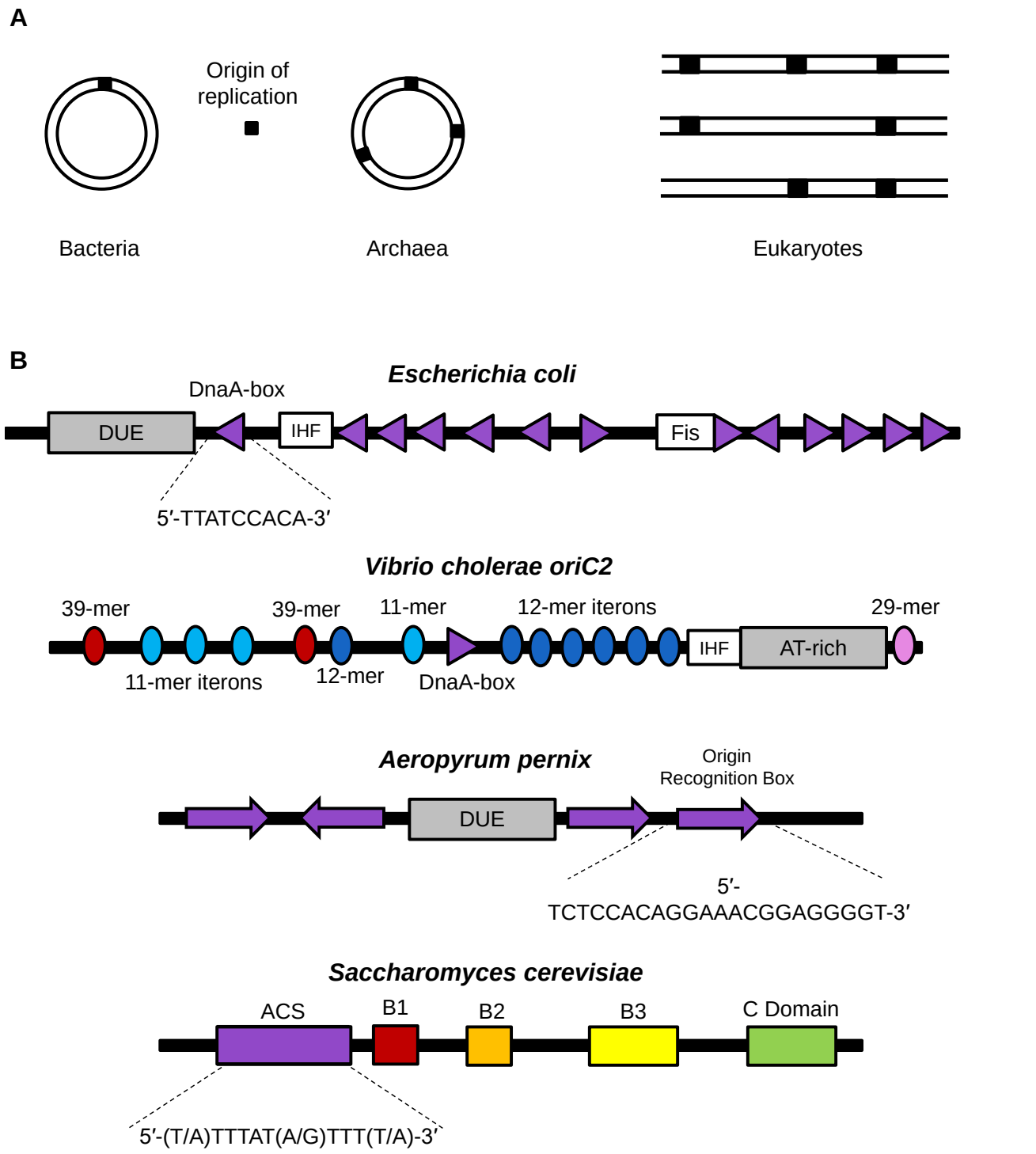


Figure 1.2. Origins of replication across the domains of life. (A) Simplified comparison of the number of chromosomes and origins in Bacteria, Archaea and Eukaryotes. **(B)** Simplified schematic comparing the origin of replication for the bacteria *E. coli* to that of the secondary chromosome of the bacteria *V. cholera,* the origin of replication for the archaea *A. pernix* and the ARS of the eukaryote *S. cerevisiae*. DnaA-boxes/origin recognition boxes (purple triangles/arrows respectively) are shown in their respective orientations with a representative sequence highlighted. The ARS consensus sequence (ACS) is highlighted.

All three domains of life utilise origin-binding proteins composed of a related AAA+ motif to recognise the origin and recruit the replicative helicase to initiate replication. In bacteria the master initiator protein is DnaA which actively unwinds the origin before recruiting the helicase (Hansen and Atlung, 2018). In eukaryotes the six subunit origin recognition complex (ORC), formed of 5 subunits related to AAA+ proteins, form a ring-shaped hexamer that together with the AAA+ protein Cdc6, binds DNA but does not unwind it. For archaeal initiation AAA+ proteins related to Cdc6 and the largest ORC subunit, Orc1, are utilised. These archaeal initiator proteins are often referred to as ORC/Cdc6 or simply ORC (Dueber *et al.*, 2007; Creager *et al.*, 2015). To exemplify the level of structural similarity between these proteins the AAA+ and DNA binding domains of DnaA from *Aquifex aeolicus* is compared to the ORC1 protein from *A. pernix* in Figure 1.3.A.

One of the key activities of these initiator proteins is the coordinated loading of the two replicative helicases onto the origin for unwinding the DNA and assembling the replication machinery for each of the resulting replication forks. In bacteria DnaA bound by ATP forms a filament upon the DnaA-boxes which unwinds the AT-rich DUE producing an open complex of ssDNA, onto which the homohexameric helicase is loaded directly by DnaA with the assistance of a helicase loader protein (Figure 1.3.B) (Katayama *et al.*, 2010). After helicase loading, primase is recruited and synthesises a short RNA primer required by DNA polymerase to begin replicating DNA. Binding of the primase to the helicase may help to release the helicase loader by stimulating its ATP hydrolysis activity and inactivating the protein. The ejection of the helicase loader results in an active helicase which is capable of unwinding DNA (Bell and Kaguni, 2013; O'Donnell *et al.*, 2013).

While DnaA is highly conserved, DnaA-independent initiation of replication is possible. In cyanobacteria, for example, the deletion of *dnaA* in certain genus (*Anabaena* and *Synechocystis*) does not affect DNA replication (Ohbayashi *et al.*, 2015). In *E. coli* mechanisms for *oriC*-independent initiation have been discovered involving R-loops. During transcription the newly synthesised RNA polymer can hybridise to the complimentary DNA strand, displacing the second strand leaving it single-stranded and producing the nucleic acid structure termed an R-loop. These structures are normally resolved, but a stable R-loop can function as a primer providing a recognition site for the helicase/polymerase to initiate replication from. RNase HI is an enzyme involved in the resolution of R-loops by degrading the RNA. Deletion of RNase HI leads to stable

R-loop formation and *oriC*-independent initiation (Kogoma and von Meyenburg, 1983; Lombrana *et al.*, 2015). During transcription negative supercoiling is generated behind the RNA polymerase which can result in R-loop formation. Topoisomerase I relaxes this supercoiling inhibiting R-loop formation. Evidence suggests deletion of this enzyme leads to extensive R-loop formation and *oriC*-independent initiation (Usongo and Drolet, 2014).

For bacteria possessing a secondary chromosome, initiation at the main chromosome is controlled by DnaA, as described above, while at the other it is controlled by another specific initiator protein. At the second chromosome in *V. cholera* RctB binds to the iterons and actively unwinds the DNA. Interestingly, RctB is actually inhibited by ATP binding (Duigou *et al.*, 2008). How the helicase is loaded at *oriC2* is unknown as no interaction between it and RctB has been shown (Fournes *et al.*, 2018).

The eukaryotic helicase is a heterohexamer known as the MCM (minichromosome maintenance) complex, with each of the subunits being AAA+ ATPases. A double hexamer MCM is recruited to origin bound ORC-Cdc6 by Cdt1 and is initially loaded onto double-stranded DNA forming a pre-replicative complex (pre-RC) (Figure 1.3.B) (O'Donnell *et al.*, 2013). The isomerisation of the duplex DNA bound by the pre-RC to encircling ssDNA involves Cdc45 and GINS, which together with MCM and various other initiation factors (including a DNA polymerase) form a pre-initiation complex (pre-IC). An active helicase or CMG (**C**dc45-**M**CM-**G**INS) complex is formed from the pre-IC through a series of regulated molecular events. The key steps are the MCM double hexamer disassociating into two single hexamers, and these hexamers opening the DNA duplex and expelling one DNA strand (Yuan *et al.*, 2019). Once the CMG is activated the polymerases are recruited, including the Pol α-primase (potentially recruited by Mcm10) which synthesises the ~37 nt RNA/DNA hybrid primer (Figure 1.3.B) (Dhingra and Kaplan, 2016).

The helicase of archaea is also recruited as a double hexamer, but it is a homohexamer of a single MCM protein which is related to the eukaryotic MCM and has homology to Cdc6. Archaeal MCM does not appear to require accessory proteins for its loading or activation, suggesting it may bind the archaeal ORC directly (O'Donnell *et al.*, 2013).

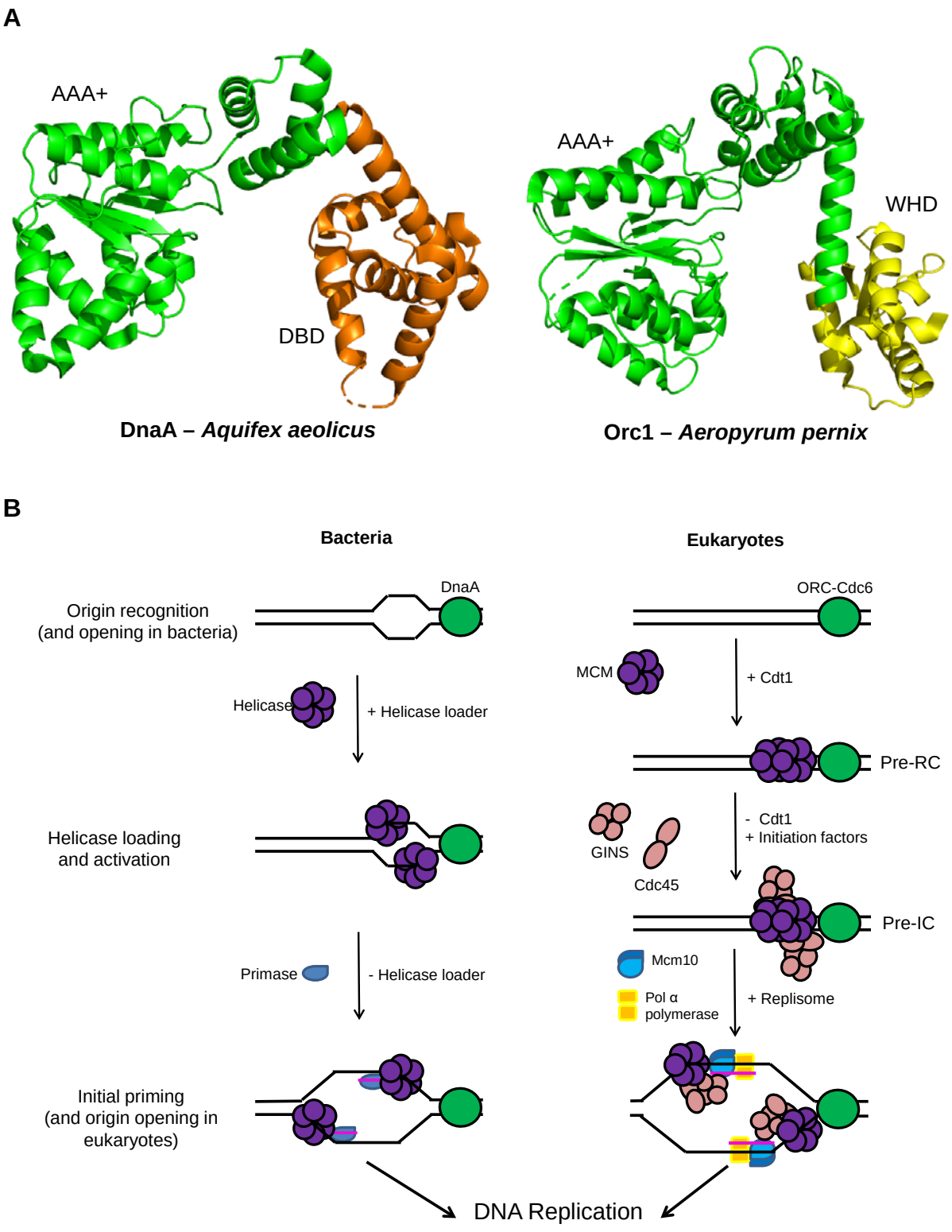


Figure 1.3. Origin recognition proteins and the initiation of DNA replication. (A) The AAA+ and DNA binding domain (DBD) of DnaA from *A. aeolicus* (PDB ID 1L8Q) and ORC1 from *A. pernix* (PDB ID 2V1U) which is formed of a AAA+ and winged helix domain (WHD). **(B)** Origin recognition, helicase loading/activation and initial priming stages of DNA replication initiation in bacteria and eukaryotes. Only the key proteins involved in each stage are shown for simplicity, adapted from O'Donnell *et al*., 2013.

1.1.2. Synthesis of DNA

Following the opening of the origin, loading and activating of the helicase and priming with RNA, the multiprotein replication machinery assembles to synthesise both new strands of DNA simultaneously. The replication machinery, or replisome, has the same core components across all domains of life. This machinery, highlighted in Table 1.1, is composed of a helicase, a primase, DNA polymerases, circular sliding clamps, a clamp loader and a single-strand binding protein (O'Donnell *et al.*, 2013).

Component	**Bacteria**	**Eukaryotes**	**Archaea**
Polymerase (Family)	Pol III (C-family)	Pol α/δ/ε (B-family)	Pol (B/D-family)
Clamp	β	PCNA complex	PCNA
Clamp Loader	γ/τ complex	RFC	RFC
SSB	SSB	RPA	RPA
Primase	DnaG	Pol α-primase	PriSL
Helicase (Family)	DnaB (RecA)	MCM/CMG (AAA+)	MCM (AAA+)
Other	-	Ctf4, Mcm10	-

Table 1.1. Replisome components across the domains of life.

Figure 1.4.A illustrates the bacterial replisome which is proposed to be organised around the replicative helicase. The helicase translocates 5′→3′ along the lagging strand separating the DNA duplex. Single-strand binding (SSB) protein forms tetramers and binds the unwound strands to remove secondary structures. DNA polymerase III (Pol III) then uses the ssDNA as a template to synthesise the complimentary strand. A circular β sliding clamp (DnaN) tethers Pol III to the template. A clamp loader complex composed of three τ subunits binds to the polymerase and helicase to organise the replisome (Yao and O'Donnell, 2010).

The anti-parallel DNA strand structure means that one strand is synthesised more continuously (the leading strand) while the other discontinuously (the lagging strand) in short 1-2 kb pieces (Okazaki fragments) that are later joined together by a DNA ligase. At the lagging strand Pol III translocates in the opposite direction to that of the leading strand Pol III and the direction of the replisome as a whole (Figure 1.4.A). As such, the lagging strand polymerases must regularly disassociate and re-associate to

synthesise each new Okazaki fragment. Here primase (DnaG) synthesises short RNA primers and the clamp loader repeatedly loads new β clamps to re-associate new Pol IIIs. The loader may bind three Pol IIIs which aids in the efficient replication of new Okazaki fragments. SSB protects the ssDNA from nucleases and enables the loader to dislodge primase from the primed site to facilitate replication (O'Donnell *et al.*, 2013).

The replisomes of eukaryotes and archaea are composed of a similar set of machinery (Table 1.1), but the connections between these differ to those of the bacterial replisome as can be seen from the replisomes outlined in Figures 1.4.B and 1.4.C. The eukaryotic CMG helicase functions analogously to the bacterial helicase, but interestingly it translocates 3′→5′ encircling the leading strand (Dhingra and Kaplan, 2016). The eukaryotic replisome, outlined in Figure 1.4.B, contains two polymerases which synthesise either the leading (Pol ε) or lagging (Pol δ) strand. Both polymerases function with a trimeric processivity factor PCNA (proliferating cell nuclear antigen) which performs a similar role to the dimeric bacterial β clamp. The Pol α-primase creates a short RNA primer that itself extends with DNA. The lagging strand is also bound by the trimeric replication protein A (RPA) that is structurally and functionally similar to tetrameric bacterial SSB. The clamp loader RFC (replication factor C) is formed of subunits with a sequence and structural homology to the bacterial loader (Yao and O'Donnell, 2010; O'Donnell *et al.*, 2013).

In addition to these components the eukaryotic replisome possess additional factors. One such factor is Ctf4 (chromosome transmission fidelity) which tethers Pol α to the helicase. Ctf4 forms a trimer which binds a conserved motif within the GINS component of the CMG helicase and can contact two Pol α (Simon *et al.*, 2014). Cft4 also acts a hub for further proteins connecting the helicase to multiple accessory factors such as Dna2, which is involved in Okazaki fragment processing (Villa *et al.*, 2016). Mcm10 is another additional factor that is required for the association of various replication proteins with chromatin. One of these proteins is the Pol α primase which Mcm10 interacts with and may recruit and stabilise on DNA (Warren *et al.*, 2009).

The archaeal replisome is outlined in Figure 1.4.C and is eukaryotic in nature. The archaeal MCM helicase, related to eukaryotic MCM and Cdc6, translocates 3′→5′ encircling the leading strand identically to the eukaryotic helicase. The archaeal replisome usually possess two polymerases which synthesis either the leading strand (PolD/B) or lagging strand (PolB). Both polymerases function with PCNA loaded by

RFC both of which are homologues to their eukaryotic counterparts, as is the lagging strand binding protein RPA. The heterodimeric archaeal primase, PriSL functions similar to eukaryotic Pol α-primase. Both the primase and helicase interact with Gins through the unique Gins-associated nuclease (Gan), which aids in maintaining the structural integrity of the replisome (Lindås and Bernander, 2013).

As discussed the eukaryotic and archaeal replisomes are composed of a homologues core machinery. The key replication enzymes, DNA polymerase (DNAP), however differ between these two domains, as well as being distinct from the bacterial enzymes. Replicative polymerases are from four families; A, B, C and D (Yao and O'Donnell, 2016). The bacterial polymerase (Pol III) is a C-family DNAP with two distinct copies required for replicating the leading and lagging strands. Removal of the RNA primer and Okazaki fragment maturation requires the monomeric A-family polymerase Pol-I (Raia *et al.*, 2019). Eukaryotic replication is performed by B-family DNAPs α, δ and ε (O'Donnell *et al.*, 2013). Mitochondrial DNA replication is performed by Polγ, an A-family polymerase (Lee *et al.*, 2009). The polymerases responsible for replicating the genomes of archaea are not known to the same degree as for eukaryotes and bacteria. What is known is that all archaea possess B-family DNAPs homologues to the catalytic subunits of the eukaryotic DNAPs. Archaea also possess D-family polymerases present in all phyla except Crenarchaea. It has been proposed that replication may occur differently between archaeal species with PolD responsible for the replication of both strands in some species or only being responsible for lagging strand synthesis in others with PolB replicating the leading strand (Raia *et al.*, 2019). DNAPs are also involved in the process of DNA repair. Y family polymerases repair DNA lesions by bypassing damaged bases that would block replication fork progression (Yang, 2014). The eukaryotic specific X family polymerases specialise in template independent repair mechanisms such as base excision repair and non- homologous end joining (Yamtich and Sweasy, 2010).

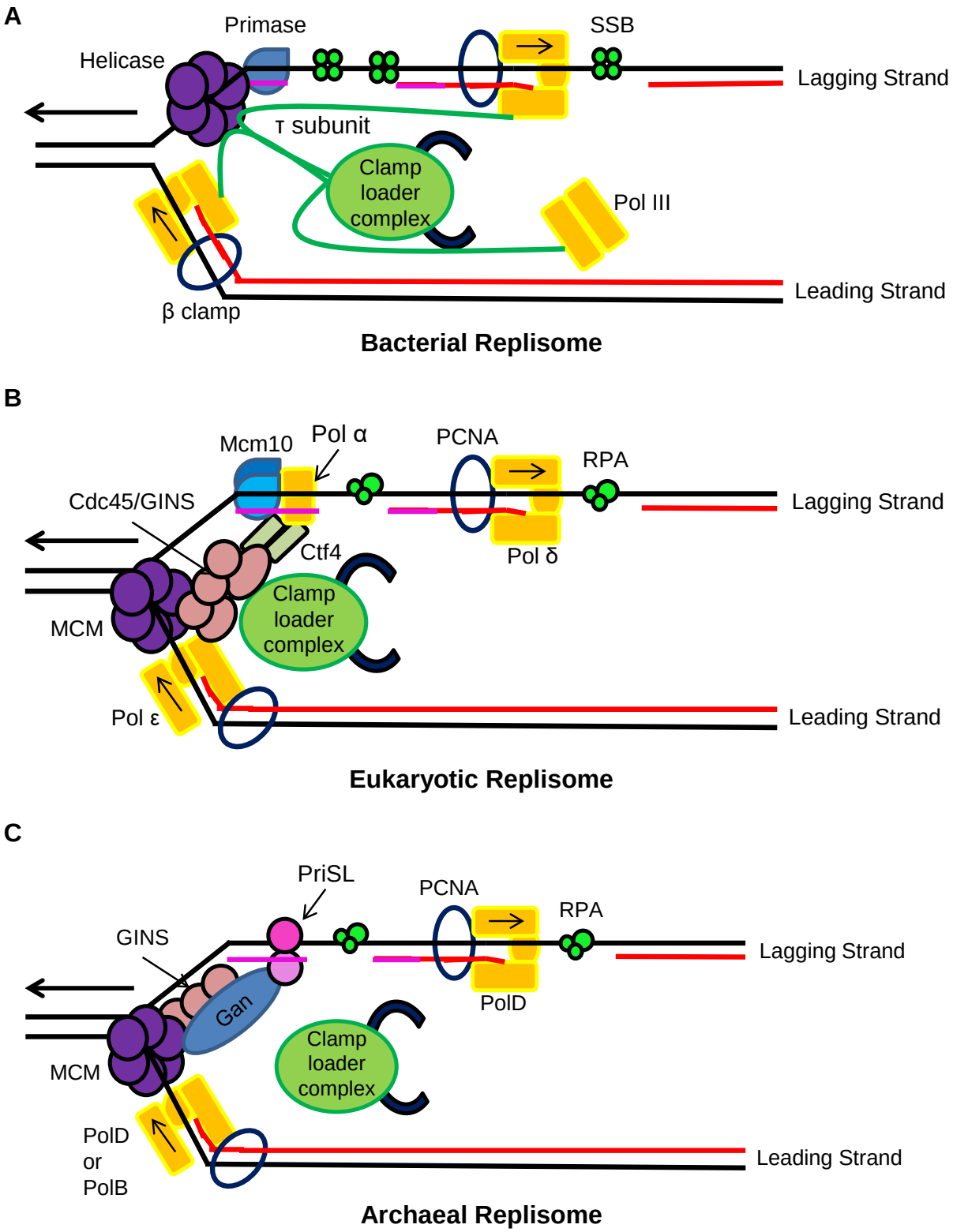


Figure 1.4. The replication machinery of Bacteria and Eukaryotes. The replisomes of **(A)** Bacteria, **(B)** Eukaryotes and **(C)** Archaea highlighting the key components of each. The original DNA template is highlighted in black with newly synthesised DNA strands shown in red and RNA primers shown in magenta. Arrows indicate the direction of polymerase translocation or the direction the replication fork is moving. Figure adapted from O'Donnell *et al*., 2013 and Lindås and Bernander, 2013.

1.1.3. Termination of DNA Synthesis

The final step in the DNA replication process is the termination of DNA synthesis which requires the arrest and/or disassembly of the replication fork. This usually occurs when replisomes meet and terminate proximal to one another. Termination involves replication forks converging, followed by the dissociation of the replisome and gap filling to complete DNA synthesis (Gowrishankar, 2015).

In eukaryotes, replication is terminated by the merger of opposing replication forks, with the vast majority of termination events being non-specific to the DNA sequence. A current model for termination is that as the replication forks converge the CMG helicases bypass each other and continue translocating until reaching an Okazaki fragment. The ssDNA encircled helicases move onto the dsDNA and the leading strand is synthesised up to the downstream Okazaki fragment which is then processed completing replication. CMG encircling dsDNA is ubiquitylated and removed from the chromatin (Dewar and Walter, 2017).

In bacteria termination, similarly to initiation, will occur at a specific locus containing a termination sequence and proteins which halt replication. Many bacterial chromosomes contain a terminus region which arrests replication fork progression. This region contains *Ter* sequences which are bound by a site-specific DNA binding termination protein (e.g. Tus in *E. coli* and RTP in *B. subtilis*). In *E. coli* the terminus contains at least ten *Ter* sites, five trap the clockwise fork and five the counter clockwise, while in *B. subtilis* there are six *Ter* sequences, with three for each replication fork (Figure 1.5.A) (Mirkin and Mirkin, 2007).

In *E. coli* Tus blocks the DnaB helicases in an orientation-dependent manner and does not prevent progression of the other fork. This depends on which *Ter* sites are bound as they are orientated so that the fork can pass through the first five *Ter* sites they encounter but stall at the last five. This allows forks to enter the termination zone but prevents them leaving thereby creating a trap for the first arriving fork to await the arrival of the slower moving fork (Figure 1.5.A) (Dewar and Walter, 2017).

Different models have been proposed to explain how Tus-*Ter* prevents orientation dependent replication fork progression (Figure 1.5.B) (Berghuis *et al.*, 2018). The helicase interaction model proposes that DnaB interacts specifically with Tus orientated non-permissively, with amino-acid residues on one side of the protein

directly interacting with the helicase (Figure 1.5.B) (Mulugu *et al.*, 2001). The dynamic clamping model proposes that Tus binds with differing strength along *Ter* leading to differences in blocking efficiency depending from which direction the helicase approaches (Figure 1.5.B) (Neylon *et al.*, 2005). Finally the mousetrap model proposes that strand separation from the permissive end of *Ter* leads to Tus dissociation. In contrast separation from the non-permissive end results in a conserved cytosine becoming bound by a cytosine-specific pocket on the Tus surface forming a stable locked complex preventing Tus dissociation (Figure 1.5.B) (Mulcair *et al.*, 2006).

In *B. subtilis* the RTP (replication termination protein) binds as pairs of dimers to the core and auxiliary elements of the bipartite Ter sites (so each Ter is bound by four RTP) (Figure 1.5.A). These two elements are bound with differing affinity and it is this difference in binding strength that determines which helicase will be blocked. If the replication fork reaches the core site first the helicase is arrested but if the auxiliary site is reached first the RTP dimers are displaced and the replisome can progress (Manna *et al.*, 1996; Mirkin and Mirkin, 2007).

A possible mechanism for replication termination from *E. coli* is that the forks converge between the *Ter* sites, the lagging strand encircled helicases pass each other and collide with the downstream leading strand. The helicase then disassociates, gaps are filled and the final Okazaki fragment processed completing DNA synthesis (Dewar and Walter, 2017).

The role of this replication fork trap is not fully understood as there are no obvious phenotypes associated when it is inactivated (neither *Tus* nor *RTP* are essential (Iismaa and Wake, 1987; Roecklein *et al.*, 1991)). It has been demonstrated in *E. coli* that head-on fusion of replication forks can result in the formation of intermediates which can trigger unwanted reactions such as over-replication or recombination events (Dimude *et al.*, 2016). For example during fork fusion the leading strand polymerase may be dislodged by the helicase resulting in a 3' flap. This flap can provide a substrate for replisome assembly (mediated by PriA) triggering initiation (Rudolph *et al.*, 2013). These and other intermediates are processed by enzymes such as 3' exonucleases and RecG helicase (which processes recombination intermediates) preventing the pathological effects of fork fusion events. If these processing enzymes are absent, fork fusion is safely contained within the Ter region (Dimude *et al.*, 2016).

Another possible suggestion for the role of the fork trap is for coordinating replication and transcription. Replication and transcription use the same DNA strand template but transcription is ~10 times slower leading to potential collisions and possible replication fork collapse. Trapping of the replication fork enforces directionality on replication with each half of the genome being replicated in a defined direction. Highly transcribed genes are preferentially located on the leading strand template (93% of such genes in *E. coli*) allowing both process to move co-directionally avoiding problems associated with head-on collisions (Dimude *et al.*, 2016).

A final possible function for the Ter locus is that immediately following termination the chromosomes are separated for cell division and these separation events occur at a specific chromosomal site, *dif* (Figure 1.5.A). The cell cycle processes of chromosome replication, segregation and cell division overlap leading to the suggestion that the replication forks are arrested at a specific locus to coordinate replication with the mechanisms of post-replicative segregation (Duggin *et al.*, 2008).

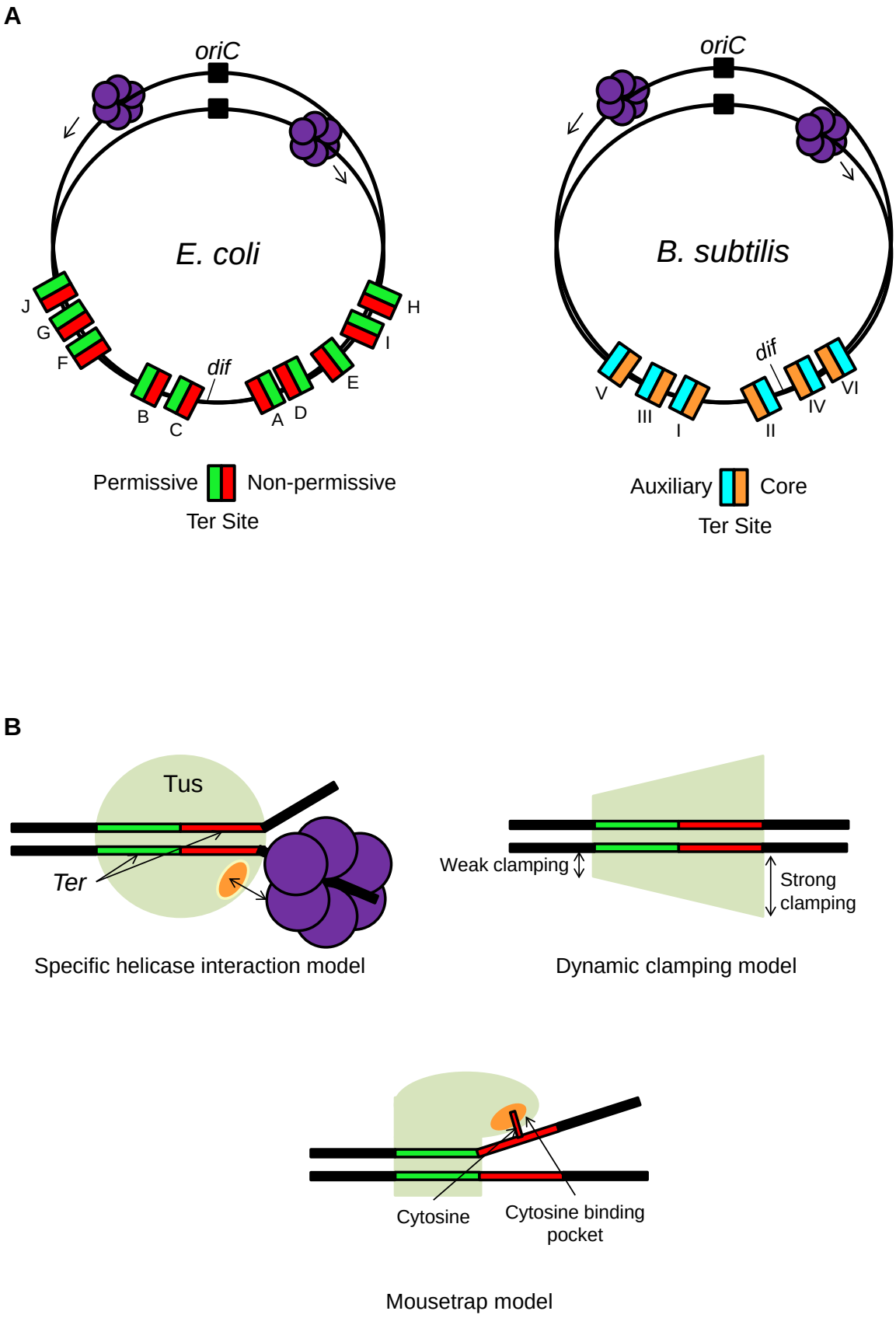


Figure 1.5. Termination of replication. (A) Organisation of the terminus region of *E. coli* and *B. subtilis*. **(B)** Models for the mechanism of replication fork arrest by Tus-*Ter* in *E. coli*. Figure adapted from Mirkin and Mirkin 2007 and Berghuis *et al*., 2018.

1.2. The bacterial origin of replication

So far this chapter has given an overview of DNA replication across the three domains of life. Now I will focus on the process of initiating DNA replication in bacteria.

As described previously, the process of initiating DNA replication begins at specific sites on the chromosome called the origin of replication. Many bacterial replisomes assemble at a single origin, *oriC*, which is composed of multiple repeated DNA sequences that mediate both origin opening and replication initiation as briefly touched upon in Section 1.1.1 (O'Donnell *et al.*, 2013). The bacterial domain contains a vast array of species, and while there is some conservation of the *oriC* architecture, there are also significant differences (Wolański *et al.*, 2015).

Both the DnaA protein and the DnaA-boxes are highly conserved; however there is great diversity in terms of number, arrangement and occasionally sequence, of DnaA-boxes. Different DnaA proteins also show variation in affinity towards specific DnaA-boxes, likely leading to the diversity in origin architecture which has been optimised for the requirements of individual species (Mott and Berger, 2007).

1.2.1. The E. coli origin of replication

The most extensively studied origin of replication is that of *E. coli*. The *E. coli oriC* is approximately 250 base pairs (bp) long and formed of several important motifs depicted in Figure 1.6.A. *E. coli oriC* can be divided into the DUE and the DnaA binding region (DBR). The unwinding element is formed of several AT rich repeats discussed further in Section 1.2.3 while the DBR is formed of multiple DnaA boxes. *E. coli* DnaA binds the consensus DnaA-box sequence, 5′–TTATCCACA–3′ with the highest affinity and independent of its nucleotide bound state (ADP vs ATP), while only in the presence of ATP can DnaA bind lower affinity sites that are distinct from the DnaA-box (consensus sequence 5′–AGATCT–3′). A final set of “DnaA-ATP-boxes” are located within the DUE and are selectively bound by DnaA-ATP (Mott and Berger, 2007). Beyond binding sites for DnaA, *oriC* also contains recognition sequences for the architectural and regulatory proteins IHF, Fis and Seq-A, which will be discussed later in relation to their regulatory roles (Wolański *et al.*, 2015).

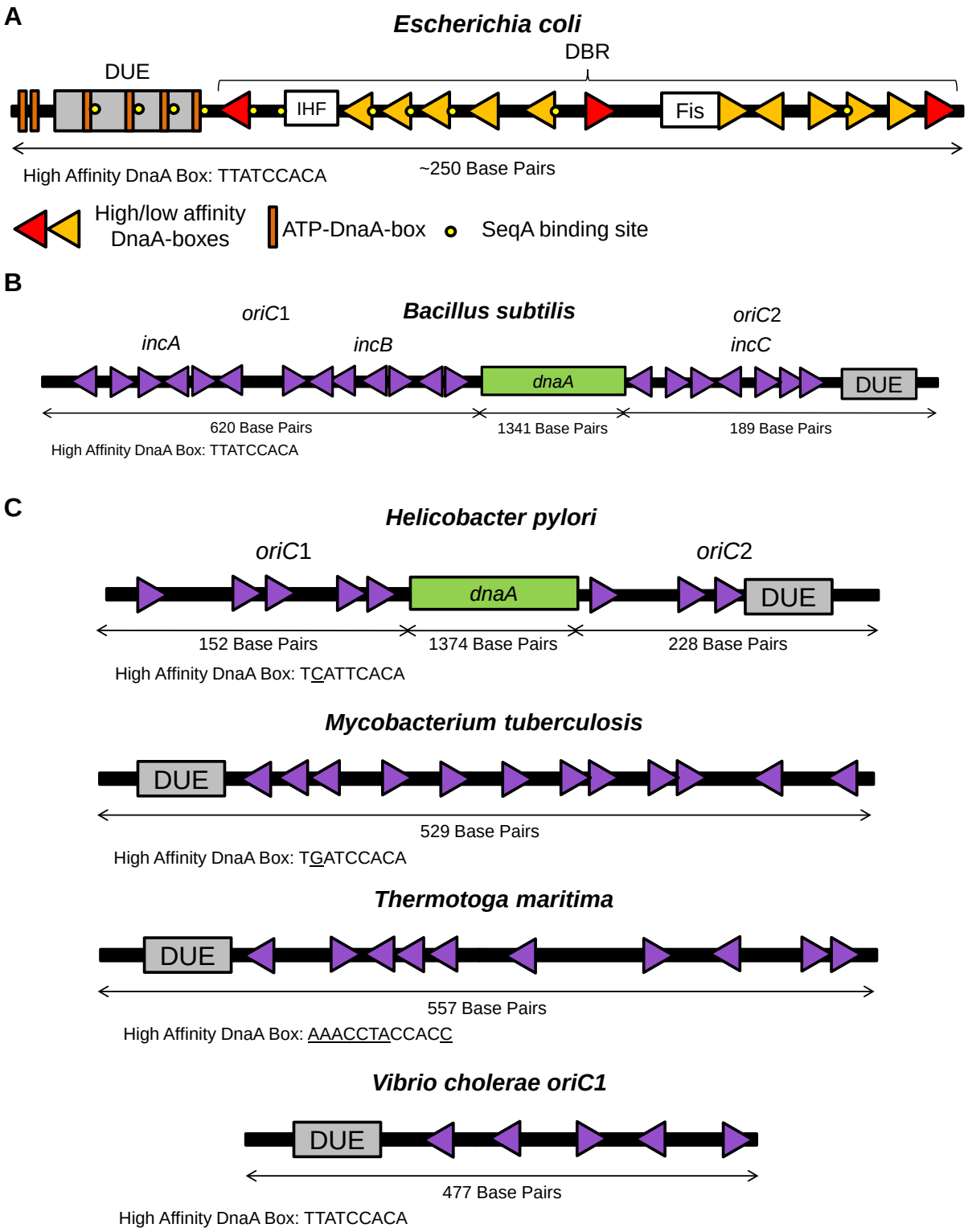


Figure 1.6. The bacterial origin of replication. (A) The origin of replication of *E. coli* showing the arrangement of key DNA motifs and protein binding sites. **(B)** The *B. subtilis oriC* highlighting DnaA-boxes as purple triangles, the DUE in grey and the *dnaA* gene in green. DnaA-boxes are shown in their relative orientation. **(C)** The structure of the origin of replication for select bacterial species, origin architecture is coloured as per B. The sequence of the highest affinity DnaA box is highlighted for each origin with positions that differ from the *E. coli* consensus sequence underlined.

1.2.2. Origins of replication in other bacterial species

While many of the features of the origin of replication are conserved, origins themselves can be quite diverse. For example the *B. subtilis oriC*, shown in Figure 1.6.B, is roughly 2150 bp in length due to being bipartite. Unlike the continuous *E. coli* origin, the *B. subtilis* origin is formed of two distinct subregions, *oriC*1 (formed of *incA* and *incB*) and *oriC*2 (*incC*) of 620 and 189 base pairs respectively, separated by the *dnaA* gene (1341 bp). The DnaA-boxes bound by *B. subtilis* DnaA with the highest affinity share the *E. coli* consensus sequence, 5′–TTATCCAC–3′ (Krause *et al.*, 1997; Wolański *et al.*, 2015). *oriC*2 alone is sufficient for origin unwinding and the role of *oriC*1, which is essential, is unknown (Murray Lab, unpublished).

Figure 1.6.C highlights further the observed diversity of bacterial replication origins. They can be continuous or bipartite, containing only a few DnaA-boxes (e.g. 5 in *V. cholera*) or many (e.g. 20 in *B. subtilis*). Moreover, the high-affinity DnaA-box sequence also may vary from the *E. coli* consensus sequence (Figure 1.6.C), either slightly as for *Mycobacterium tuberculosis* (5′–TGATCCAC–3′) or more significantly as for *Thermatoga maritima* (5′–AAACCTACCACC–3′) (Wolański *et al.*, 2015).

1.2.3. The DNA unwinding element and the DnaA-trios

As discussed above *oriC* is formed of specific DNA sequences that act as the binding sites for initiation proteins and regulatory factors. The most abundant of these sequences is the DnaA-box, the binding site for the master initiator protein DnaA, which constitute the DnaA binding region (DBR). Generally the DBRs of *oriC* neighbour regions featuring an AT-rich stretch of nucleotides termed the DNA unwinding element or DUE (Rajewska *et al.*, 2012).

The AT-rich sequences of the DUE have been observed as repeating motifs separated by short non AT-rich insertions (Figure 1.7.A). For example, the DUE of *E. coli* features three repeating AT-rich sequences of 13 nucleotides each (13-mer) carrying the consensus sequence 5′-GATCTnTTnnTTT-3′ separated by insertions of 2-3 bp. Three 13-mers are also found in the DUE of *Pseudomonas putida* and *V. cholera* except the 13-mers are separated by longer (11-12 bp) or shorter insertions, respectively. In the DUE of *T. maritima* the three repeating sequences are not separated and are 9 nucleotides in length carrying the sequence 5′-TATnATTnn-3′ (Rajewska *et al.*, 2012; Wolański *et al.*, 2015).

AT-rich regions have lower thermodynamic stability which facilitates destabilisation of the double helix, supporting the role of the DUE in initiation (Rajewska *et al.*, 2012). In *E. coli* it is proposed that the binding of DnaA to the DnaA-boxes induces unwinding of the proximal R (right) and M (middle) 13-mers (Figure 1.7.A) creating an open complex upon which the helicases are loaded. The helicases encircle the R and M sites and force the L (left) site to unwind, creating the replication bubble within which the enzymes are positioned for the bidirectional unwinding of the chromosome and assembly of the replisome (Coman and Russu, 2005).

The chromosome origin region initially unwound by DnaA ranges in size from ~20-60 bp (~50 bp in *E. coli*) which appears to be sufficient to provide space to accommodate loading of the replicative helicases and the other pre-replication proteins (Rajewska *et al.*, 2012). Unwinding of the DUE has been experimentally confirmed *in vitro* for a range of organisms including *B. subtilis*, *E. coli*, *Helicobacter pylori*, *M. tuberculosis* and *T. maritima* (Wolański *et al.*, 2015).

The AT-rich stretches of *oriC* are not located directly adjacent to the final DnaA-box of the DBR but instead the two regions are normally separated by a sequence of varying size. The DNA in this location has recently been identified as playing a key role in the unwinding of the replication origin. It was initially identified in *B. subtilis* that this region contains another essential DNA motif, a repeating trinucleotide sequence termed the DnaA-trio carrying the consensus sequence 3′-GAT-5′ (Richardson *et al.*, 2016). The DnaA-trios specify DnaA binding to a specific single strand of DNA where they promote unwinding of the DNA duplex (details discussed further in later sections). Figure 1.7.B shows the unwinding region of the *B. subtilis oriC* highlighting the DNA sequence motifs located after the DnaA-boxes within *incC*, these being the GC-rich region, the DnaA-trios and the AT-rich region. Bioinformatics identified DnaA-trio motifs present within the unwinding elements of well characterised origins and could also be identified within the predicted origins of many diverse bacterial species (Figure 1.7.C). These results suggest that the DnaA-trio is a key element within the bacterial origin of replication (Richardson *et al.*, 2016)..

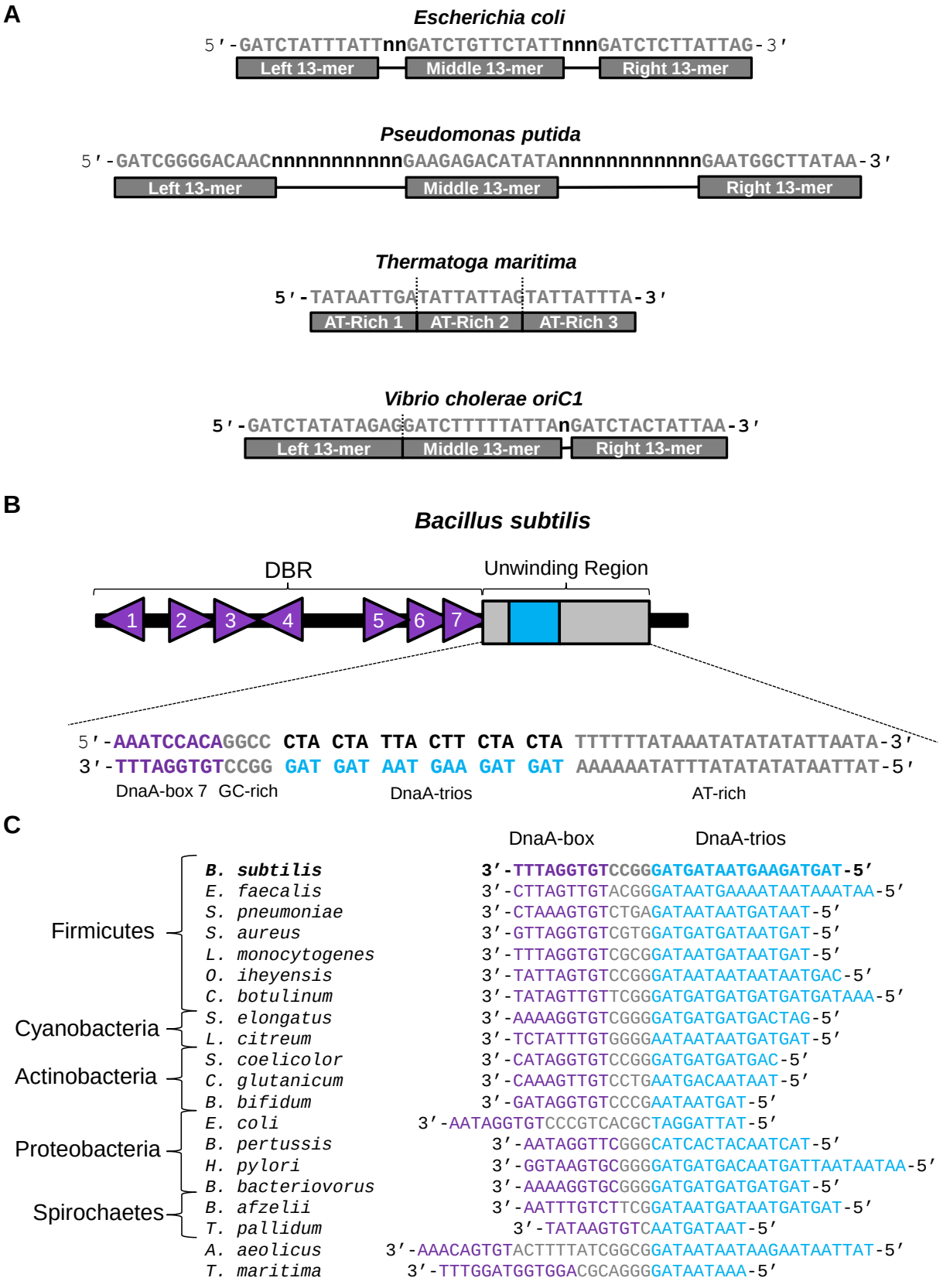


Figure 1.7. Sites of bacterial origin unwinding. (A) Sequence and layout of the AT-rich regions from the DUE of selected bacterial chromosomal origins of replication. **(B)** The sequence and layout of the unwinding region of *Bacillus subtilis incC*. **(C)** Bioinformatics (from Richardson *et al.,* 2016) identifying DnaA-trio motifs adjacent to a DnaA box throughout the bacterial domain.

1.3. DnaA: The bacterial master initiator protein

The major replication initiator of bacteria is the highly conserved DnaA which has been extensively studied in many bacterial species including the model organisms *E. coli* and *B. subtilis,* as well as *A. aeolicus*, *H. pylori, M. tuberculosis* and *T. maritima* (Shimizu *et al*., 2016).

1.3.1. The structure and function of DnaA

DnaA is formed of four functional domains (I-IV, Figure 1.8.A) with each domain performing specific functions. The domains of DnaA and the functions associated with each are briefly outlined below and in Figure 1.8.

Domain I: The N-terminal domain of DnaA is the protein interaction hub. In *E. coli*, as highlighted in Figure 1.8.B, this domain contains a helicase-binding site (involving a glutamate residue) and a hydrophobic surface surrounding a key tryptophan residue involved in low affinity domain I/domain I interactions promoting dimerization (Shimizu *et al.*, 2016). In several bacterial species this domain is a binding site for regulatory proteins.

Domain II: Domain II is a poorly conserved, flexible linker region that covalently tethers domain I to the rest of the protein (Mott and Berger, 2007).

Domain III: The conserved AAA+ domain is shown in Figure 1.8.C. This domain can bind and hydrolyse ATP. Domain III guides oligomerisation, using a conserved arginine, termed the arginine finger, and is also responsible for binding to single-stranded DNA (Duderstadt *et al.*, 2010; Duderstadt *et al.*, 2011). The domain also contains a unique α-helical insertion (highlighted in cyan in Figure 1.8.C) which separates DNA replication initiation proteins from the rest of the protein family termed the initiator specific motif or ISM. The final 3 helices of domain III form the domain III/IV boundary referred to as the domain III/IV junction (Figure 1.8.E) (Erzberger *et al.*, 2002).

Domain IV: The carboxyl-terminal domain of the protein is highlighted in Figure 1.8.D. This domain is responsible for binding to double-stranded DNA and guiding origin recognition through specifically binding to DnaA-boxes, using a helix-turn-helix motif and a basic loop carrying a conserved arginine (Fujikawa *et al.*, 2003).

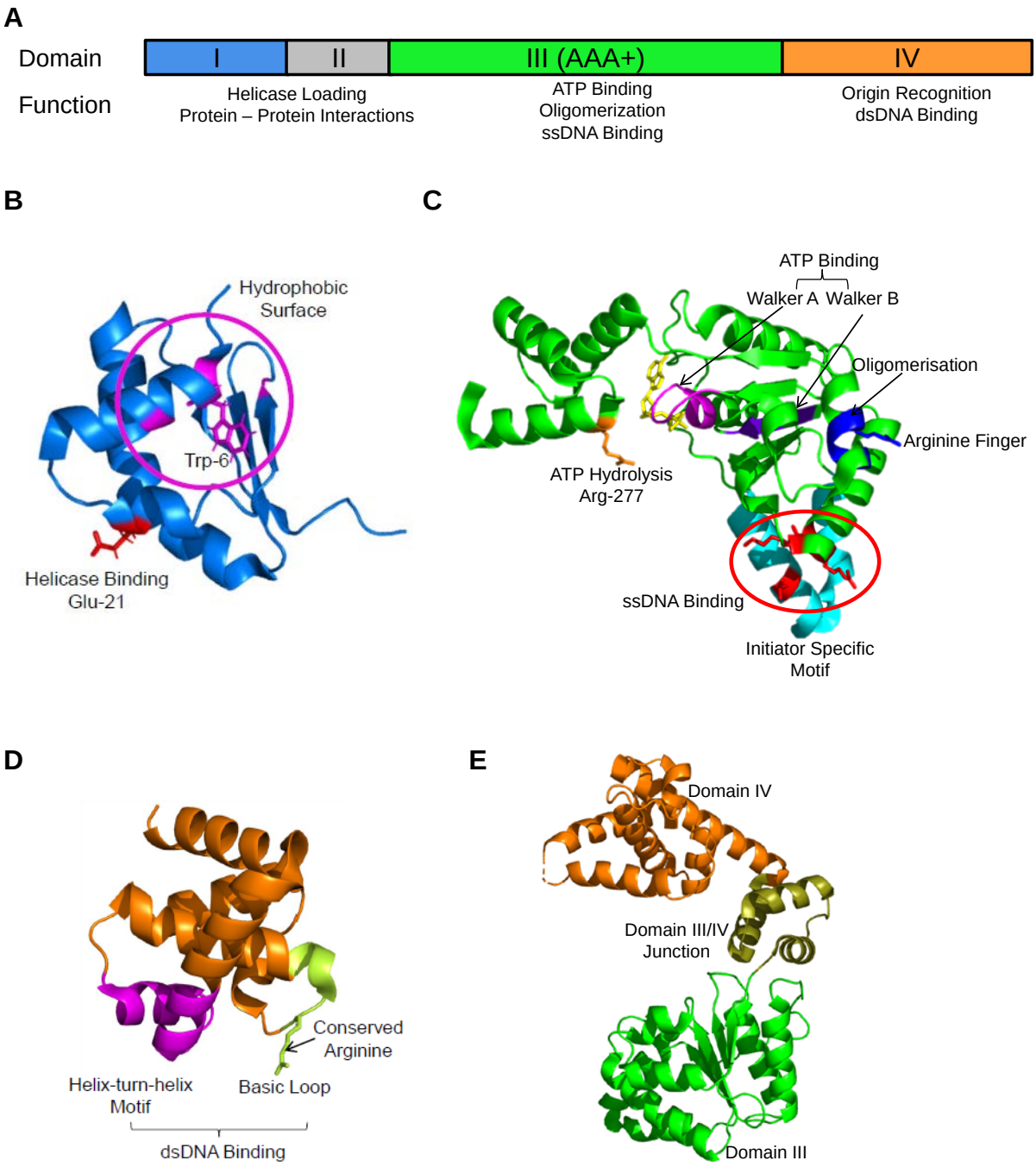


Figure 1.8. Domain organisation and key activities of DnaA. (A) Schematic of the domain organisation of DnaA and the functions these domains perform. **(B)** Crystal structure of DnaA domain I from *E. coli* (PDB ID 2E0G) **(C)** domain III from *A. aeolicus* (PDB ID 1L8Q) and **(D)** domain IV from *E. coli* (PDB ID 1J1V). Key residues and motifs are highlighted along with the associated DnaA activity. **(E)** Crystal structure of DnaA from *A. aeolicus* lacking domains I and II (PDB ID 1L8Q). The domain III/IV junction is highlighted.

1.3.2. Origin recognition and binding by DnaA

DnaA-boxes are specific DnaA binding sites within *oriC*, with the general consensus sequence being 9 base pairs long (5′-TTATCCACA-3′). The binding of DnaA to DnaA-boxes is the first step of DNA replication initiation. As mentioned in section 1.3.1 domain IV of DnaA binds a DnaA-box through base-specific interactions involving several residues of this DNA binding domain, or DBD, that form two key motifs highlighted in Figure 1.9.A (Fujikawa *et al.*, 2003). DnaA binds to both grooves of the DNA within the DnaA-box sequence, with one helix and a loop of the helix-turn-helix (HTH) motif of the DBD inserting into the major groove and the conserved arginine of the basic loop making contacts within the minor groove.

In *E. coli* it has been identified from structural studies that two histidine, a threonine, a proline and an aspartic acid residue within the HTH domain (P423 D433, H434, T435, H439) make hydrogen bonds and van der Waals contacts with the base pairs of the DnaA-box sequence in the major groove (Figure 1.9.B). Mutations to D433, H434 and T435 residues, which form part of the DnaA signature motif (RDHTTVL in *E. coli*), cause a defect in DnaA-box specific recognition but do not affect non-specific DNA binding (Fujikawa *et al.*, 2003). In the minor groove the conserved arginine (R399), which is required for double-stranded DNA binding, forms direct or water mediated hydrogen bonds with the 3rd, 4th and 5th base pairs (Figure 1.9.B). These base-specific interactions help explain DnaA affinity for the consensus DnaA-box sequence and how alterations to this sequence leads to weaker affinity. Outside of the base-specific interactions the residues of the DBD also form many electrostatic interactions with the phosphate groups of the DNA backbone of both grooves (Fujikawa *et al.*, 2003).

The specific major groove base-pair interactions have been found to be required for DnaA-box specificity in *M. tuberculosis* (Tsodikov and Biswas, 2011) and the key residues of the signature sequence have also been found to be highly conserved. This indicates the molecular mechanism for specific DnaA-box recognition to the consensus DnaA-box sequence is highly conserved, (Hansen and Atlung, 2018).

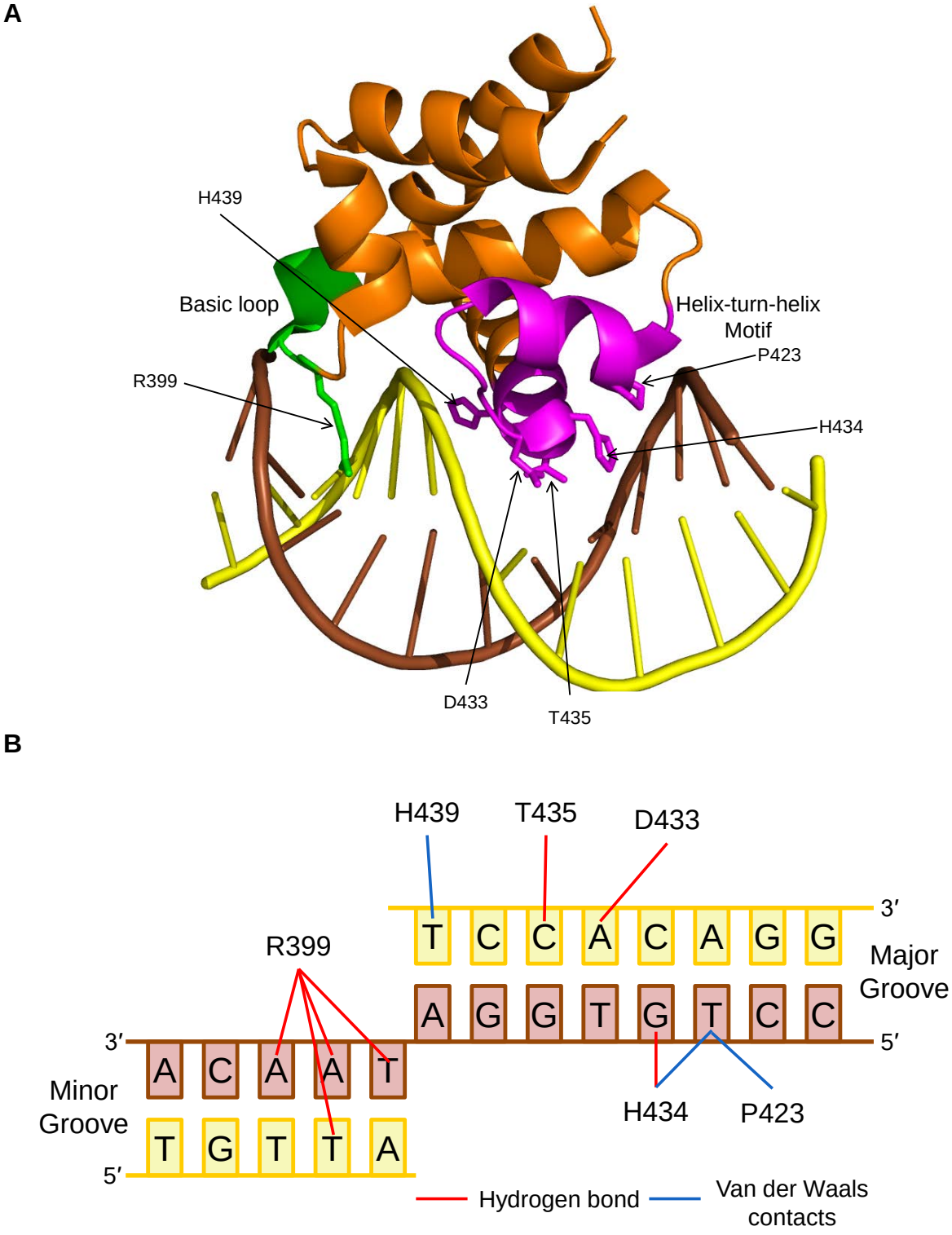


Figure 1.9. DnaA-box binding by DnaA domain IV. (A) Crystal structure of *E. coli* DnaA bound to a consensus DnaA-box sequence (PDB ID 1J1V). The Helix-turn-helix motif is highlighted magenta and the basic loop green. Key residues involved in base-specific interactions are highlighted. **(B)** Schematic highlighting the direct DNA base contacts by residues within DnaA domain IV adapted from Fujikawa *et al.,* 2003.

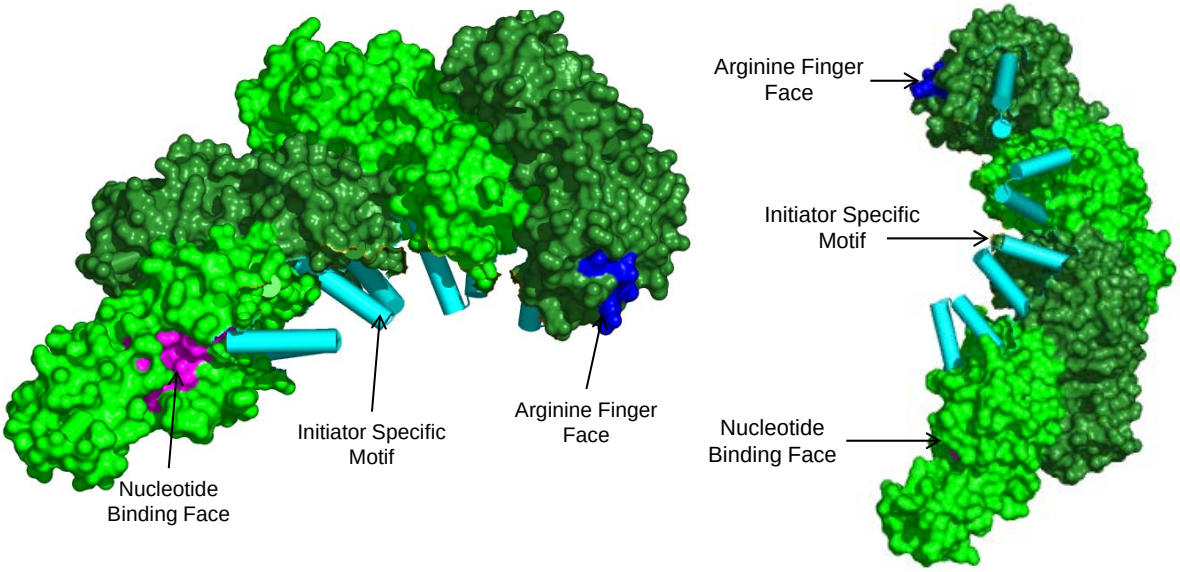
1.3.3. DnaA filament formation

DNA replication begins with the sequence specific recognition of *oriC* by DnaA. In *E. coli* evidence suggests that DnaA monomers remain stably bound to the high affinity DnaA-boxes (Figure 1.6.A) throughout most of the cell cycle. During initiation additional ATP-bound DnaA protomers bind to the lower affinity sites leading to the ATP-dependent formation of a large nucleoprotein complex (Miller *et al.*, 2009). Genetic and biochemical assays support the model that the *E. coli* origin guides the formation of DnaA into filaments bound to dsDNA (Fujikawa *et al.*, 2003). DnaA bound to the high affinity DnaA-boxes are required to promote the binding of DnaA-ATP to neighbouring low-affinity sites. The placement of the lower affinity binding sites between these high affinity boxes is believed to stabilise this large protein filament (Miller *et al.*, 2009). The DnaA filament, as will be discussed later, promotes DNA unwinding. DnaA filaments have been shown in many bacteria to be required for origin opening, making oligomerisation a key DnaA activity (Duderstadt *et al.*, 2010; Ozaki *et al.*, 2012; Richardson *et al.*, 2019).

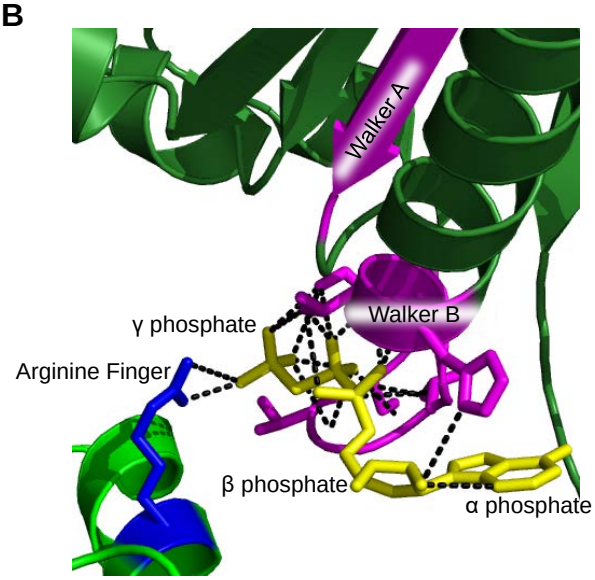
The molecular mechanism underpinning DnaA filament assembly has been investigated through structural and biochemical studies. Oligomerisation is dependent on ATP, which forms a bridge between the AAA+ domains of neighbouring protomers. ATP makes contacts with the nucleotide binding pocket (involving Walker A and B motifs) of one protomer and a conserved arginine residue (the arginine finger) of the adjacent protein (Figures 1.10.A and B) (Jha *et al.*, 2016). Several residues on either side of the AAA+/AAA+ interaction have been proposed as being important for guiding and stabilising DnaA filament assembly. This AAA+/AAA+ interface is highlighted in Figure 1.10.C (Duderstadt *et al.*, 2010; Ozaki *et al.*, 2012).

The majority of AAA+ proteins assemble into hexameric, closed-ring structures. The DnaA oligomer, however, is a right-handed helix assembling one monomer at a time (Figure 1.10.A) (Erzberger and Berger, 2006; Cheng *et al.*, 2015). This open-ended conformation is a consequence of an α-helical insertion in the core of the AAA+ domain that nudges the next monomer out of plane, preventing the oligomer forming a closed-ring. This insertion, along with a neighbouring α-helix, forms a V-shaped ‘steric wedge’ that has been termed the Initiator Specific Motif (ISM) (highlighted cyan in Figure 1.10.C) as it distinguishes initiator proteins from other AAA+ family members (Erzberger *et al.*, 2006).

A


Nucleotide
Binding Face
Initiator Specific
Motif
Arginine Finger
Face
Arginine Finger
Face
Initiator Specific
Motif
Nucleotide
Binding Face

B


Walker A
Walker B
γ phosphate
Arginine Finger
β phosphate
α phosphate

C

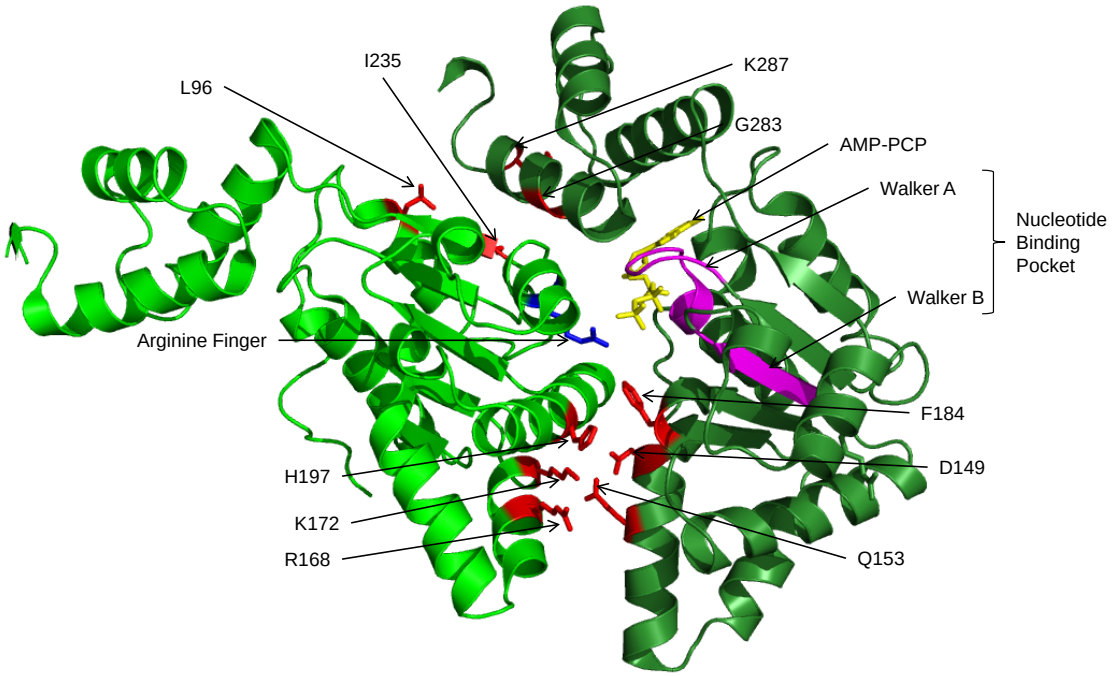


Figure 1.10. DnaA filament formation. (A) Crystal structure of an *A. aeolicus* DnaA oligomer of four AAA+ domains (PDB ID 3R8F). Oligomerisation occurs in a head to tail manner, with the ATP-bound side of one protomer contacting the arginine finger of another. **(B)** ATP (yellow) forms a bridge between neighbouring DnaA protomers. The nucleotide binding pocket is highlighted in magenta with the arginine finger shown in blue and polar contacts indicated by a dashed black line. **(C)** Crystal structure of an *A. aeolicus* DnaA oligomer (PDB ID 2HCB). Only the AAA+ domains of two protomers are shown for simplicity. The nucleotide binding pocket, arginine finger and ATP are shown as per B. Residues forming the interaction interfaces on either protomer are highlighted in red.

DnaA filaments are known to assemble on single-stranded DNA in the 3′→5′ direction and a co-crystal structure of DnaA bound to a ssDNA substrate uncovered a potential role the DBD of DnaA plays in the ability of the protein to self-assemble and bind ssDNA (Duderstadt *et al.*, 2010; Cheng *et al.*, 2015). This finding has led to the proposal of there being a second filament interaction interface between the DBD of one DnaA protein and the AAA+ domain of the one adjacent. A consequence of this interaction is that the DBD is reoriented such that it would be unable to contact dsDNA, suggesting that DnaA-filaments utilises distinct conformations for engaging either single- or double-stranded DNA (Duderstadt *et al.*, 2010).

Figure 1.11.A displays the co-crystal structure of a DnaA-filament bound to a single stranded substrate, as the structure shows the filament has assumed a conformation in which the DBD of one DnaA has docked against an adjacent AAA+ domain. Also highlighted in magenta is the helix-turn-helix domain, required for binding DnaA-boxes (Section 1.3.2). The HTH is buried within the filament, a conformation that would make binding to DnaA-boxes unfeasible. Figure 1.11.B highlights (red) the residues proposed as forming the interfaces on either side of the interaction (Duderstadt *et al.*, 2010).

These findings, coupled with biochemical and genetic approaches, have led to the proposal that DnaA-filaments assume two distinct assembly states, where the DBD is either extended away from the protein (extended state, capable of dsDNA binding) or docked against the AAA+ domain (compact state, incapable of dsDNA binding) (Figure 1.11.C) (Duderstadt and Berger, 2013). The transition between the two states involves a flexible junction between domains III and IV (Figure 1.8.E). Sequestering the HTH motif to stabilise the compact filament has been suggested to render it competent for binding ssDNA (Duderstadt *et al.*, 2010).

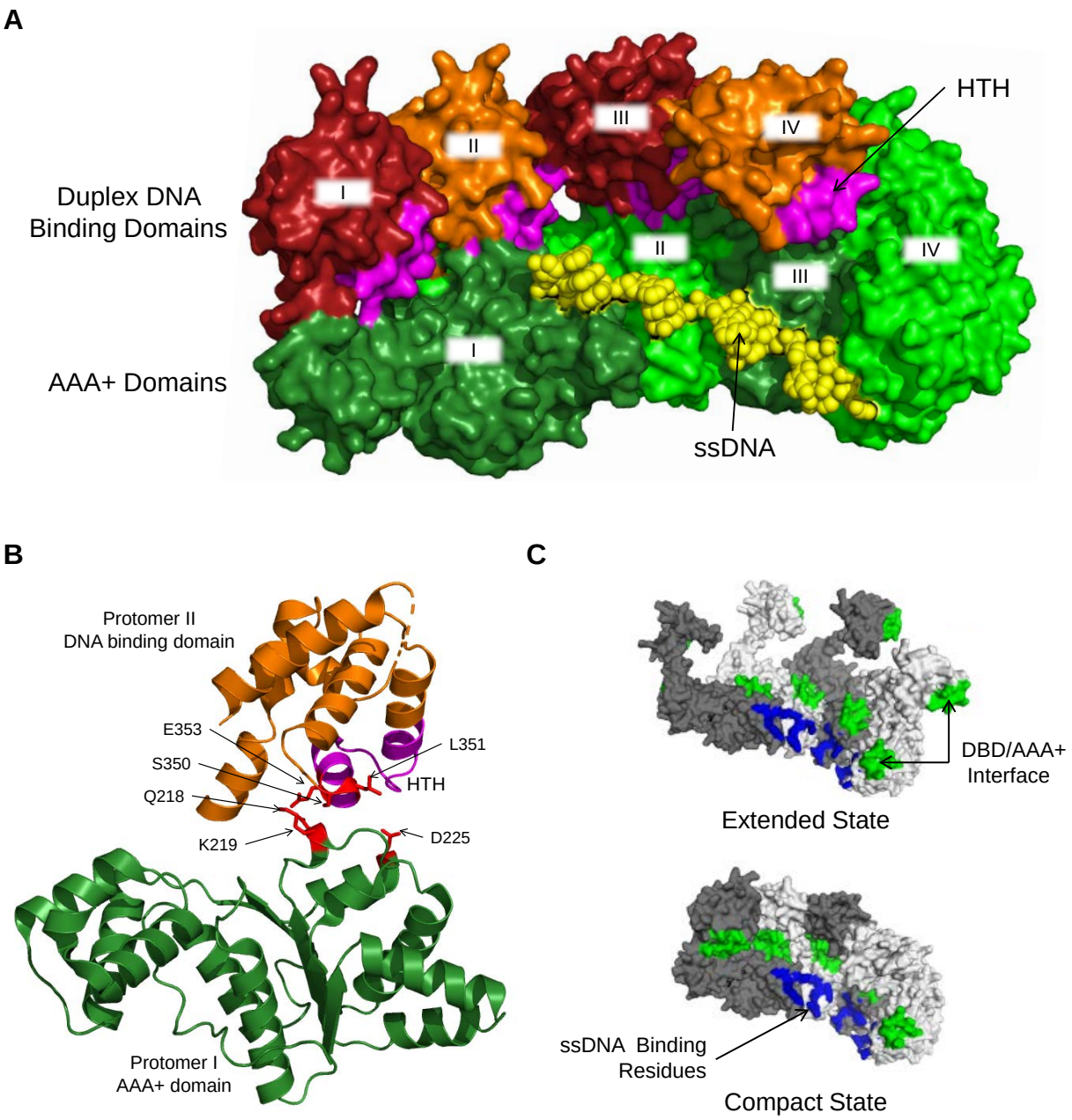


Figure 1.11. DnaA 'compact state' filament. (A) Crystal structure of a DnaA filament bound to ssDNA (PDB ID 3R8F) from *A. aeolicus*. Only the AAA+ and DNA binding domains for four DnaA protomers are shown and numbered sequentially. The AAA+ domains are highlighted in alternating shades of green while the DBD is highlighted in alternating red and orange. The Helix-turn-helix (HTH) motif is highlighted in magenta and ssDNA is shown as yellow spheres. **(B)** The AAA+ domain of DnaA protomer I and the DBD domain of promoter II from A highlighting the residues that form the DBD/AAA+ interaction interface in red. **(C)** Models of extended or compact state DnaA (Duderstadt *et al.*, 2010).

1.3.4. Single-stranded DNA binding by DnaA

While DnaA has been shown to bind to dsDNA using domain IV (section 1.3.2), it binds to ssDNA using the AAA+ domain (domain III). A co-crystal structure of a DnaA filament bound to a single stranded synthetic DNA substrate (poly-A_{12}) has led to a proposed mechanism for how DnaA binds to ssDNA (Figure 1.12.A). DnaA uses two pairs of helices: α3/α4 (the ISM) and α5/α6, with each DnaA protein within a filament binding three nucleotides (Figure 1.12.A-B). The ISM forms a shelf for a set of trinucleotides and a conserved hydrophobic residue, V156 in the *A. aeolicus* structure shown, contacts the sugar and base of the first nucleotide. The central phosphate of the second nucleotide is hydrogen bonded by Thr 191 and an electropositive, amino-terminal dipole of the α6 helix. The positively charged residues, Arg 190 and Lys 188, make salt-bridge interactions with the phosphates of the 1st and 3rd nucleotide, respectively (Figure 1.12.B) (Duderstadt *et al.*, 2011). These residues or those homologous to these, have been shown to be required for ssDNA binding by DnaA from *A. aeolicus* (Duderstadt *et al.,* 2011), *E. coli* (Ozaki *et al.*, 2008) and *B. subtilis* (Scholefield *et al.,* 2012).

An interesting observation from the crystal structure is that each trinucleotide segment bound by DnaA is separated by large (~10Å) gaps. These gaps extend the DNA substrate by ~50%. This is highlighted in Figure 1.12.C where the ssDNA substrate bound by DnaA can be seen compared to a single strand of B-DNA (Duderstadt *et al.*, 2011).

The act of the DnaA filament stretching ssDNA is strikingly similar to the behaviour of the homologous recombination protein RecA. The RecA filament binds to a single DNA strand through a similar mechanism of engaging three nucleotides and introducing gaps between each tri-nucleotide segment, extending the substrate. The two mechanisms differ however in the orientation of the nucleotide segments. The DNA bound by RecA is orientated so that the tri-nucleotide segments form a smooth spiral to enable complementary base pairing, whereas the DnaA bound tri-nucleotides are offset in such a way that prevents annealing of the DNA strands (Duderstadt and Berger, 2013). These observations suggest that both proteins stretch the DNA to promote unwinding of the duplex but that DnaA alters the bound strand further to prevent base-pairing and stabilise the open complex, providing a potential mechanism for origin opening by DnaA.

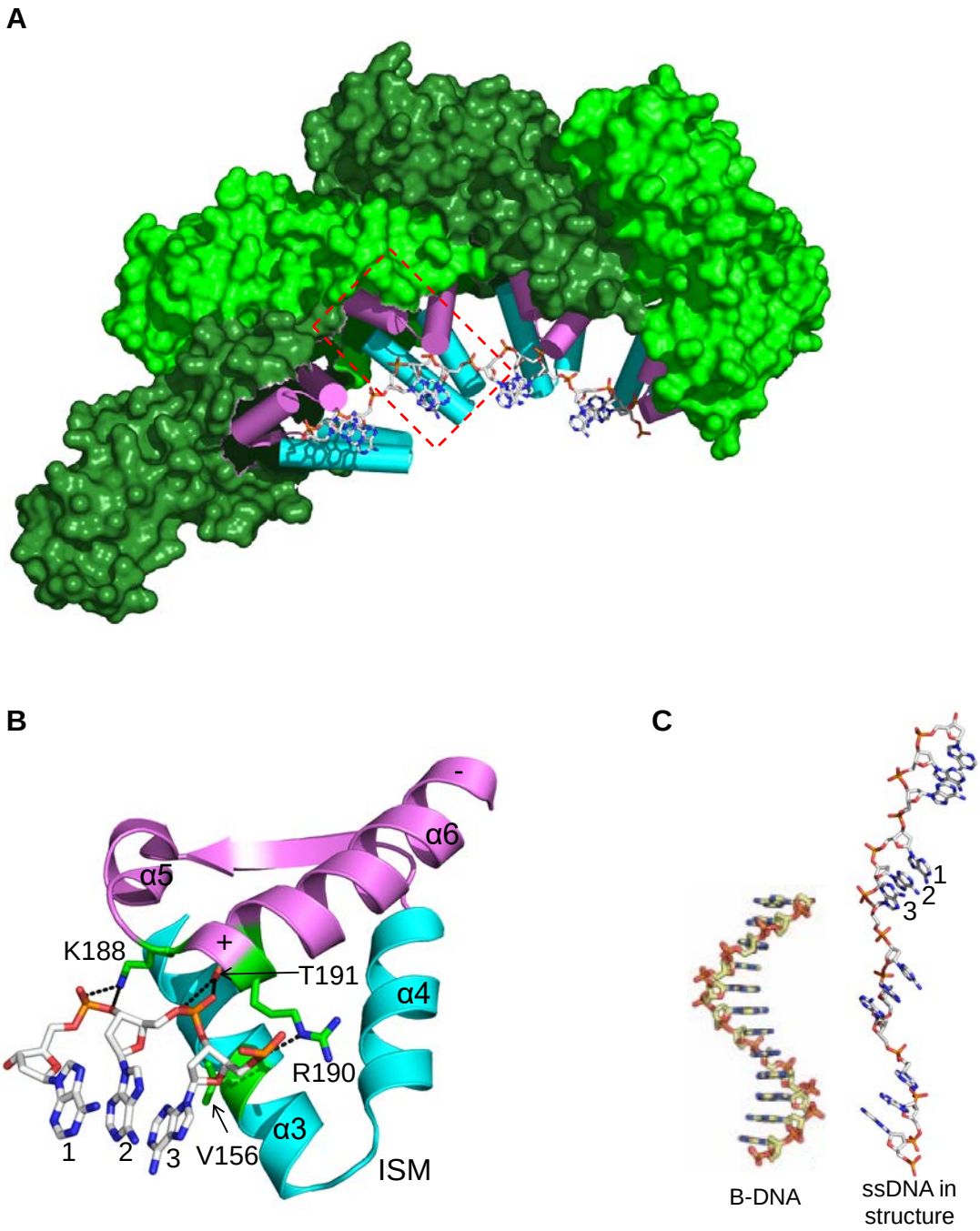


Figure 1.12. DnaA single-stranded DNA binding. (A) Crystal structure of an *A. aeolicus* DnaA filament bound to ssDNA (PDB ID 3R8F). For simplicity only the AAA+ domains are shown with individual protomers coloured alternate shades of green. The ISM and the neighbouring pair of helices are highlighted cyan and magenta respectively. Single-stranded DNA is coloured by element (carbon white, hydrogen grey, oxygen red, nitrogen blue, sulphur orange). The area indicated by a red box is displayed in B. **(B)** Contacts between DnaA and ssDNA. The helix numbers and the dipole of α6 are indicated. Colouring same as A. Residues making contacts are highlighted by element as per A (except carbon is green). Polar contacts are indicated by dashed black lines. **(C)** Comparison of the structure of single stranded B-DNA (Duderstadt *et al.*, 2011) with the ssDNA from the structure in A.

1.3.5. Protein-protein interactions of DnaA

The previous sections have shown how DnaA is able to interact with itself and DNA. Beyond these interactions, DnaA also interacts with a range of other proteins to help perform or to regulate a diverse range of functions. As briefly touched upon in Section 1.4.1, the N-terminal domain of the protein (domain I) is the site for interactions with other proteins involved in initiation across the bacterial domain (Shimizu *et al.*, 2016).

Following the unwinding of the chromosome origin in *E. coli,* DnaA loads two replicative helicase-helicase loader complexes onto the single-strands of DNA. Two regions of DnaA appear to interact with the replicative helicase DnaB, one is near the N-terminus of domain III while the other lies within Domain I (Kaguni, 2011). Within Domain I two residues have been identified through structural and biochemical studies as being required for direct DnaB binding: Glu-21 (Abe *et al.*, 2007) and Phe-46 (Kaguni, 2011) (Figure 1.13.A). Domain I of *E. coli* DnaA has also been proposed as being involved in DnaA homodimerization with a hydrophobic interaction occurring between DnaA monomers (Weigel *et al.*, 1999; Simmons *et al.*, 2003). An exposed hydrophobic patch has been identified at the N-terminus of domain I, and a tryptophan residue within this patch (Trp-6) has been shown as being required for weak DnaA-DnaA interactions (Figure 1.13.A) (Abe *et al.*, 2007).

In *B. subtilis* loading of the replicative helicase appears to be a more complicated process involving several initiator proteins recruited sequentially to *oriC* (Smits *et al.*, 2010). A key initial step in this process is the interaction between DnaA and DnaD (a *Firmicute* specific initiation protein). Domain I of DnaA has been identified as the site of the DnaD interaction surface, with several amino acids shown to be required through two-hybrid studies (Figure 1.13.B) (Martin *et al.*, 2019).

Studies looking at the helicase loader, DnaC, from *A. aeolicus* have suggested that the protein is capable of directly interacting with DnaA. This interaction is proposed to occur between the AAA+ domains of the respective proteins as it requires both ATP-bound DnaA and an intact DnaC oligomerisation interface. The proposed mode of interaction is that DnaC uses the AAA+ domain of DnaA-ATP as a docking site, binding to the end of a DnaA filament following origin unwinding and directs the accurate recruitment and deposition of the helicase onto one strand of *oriC* (Figure 1.13.C) (Mott *et al.*, 2008).

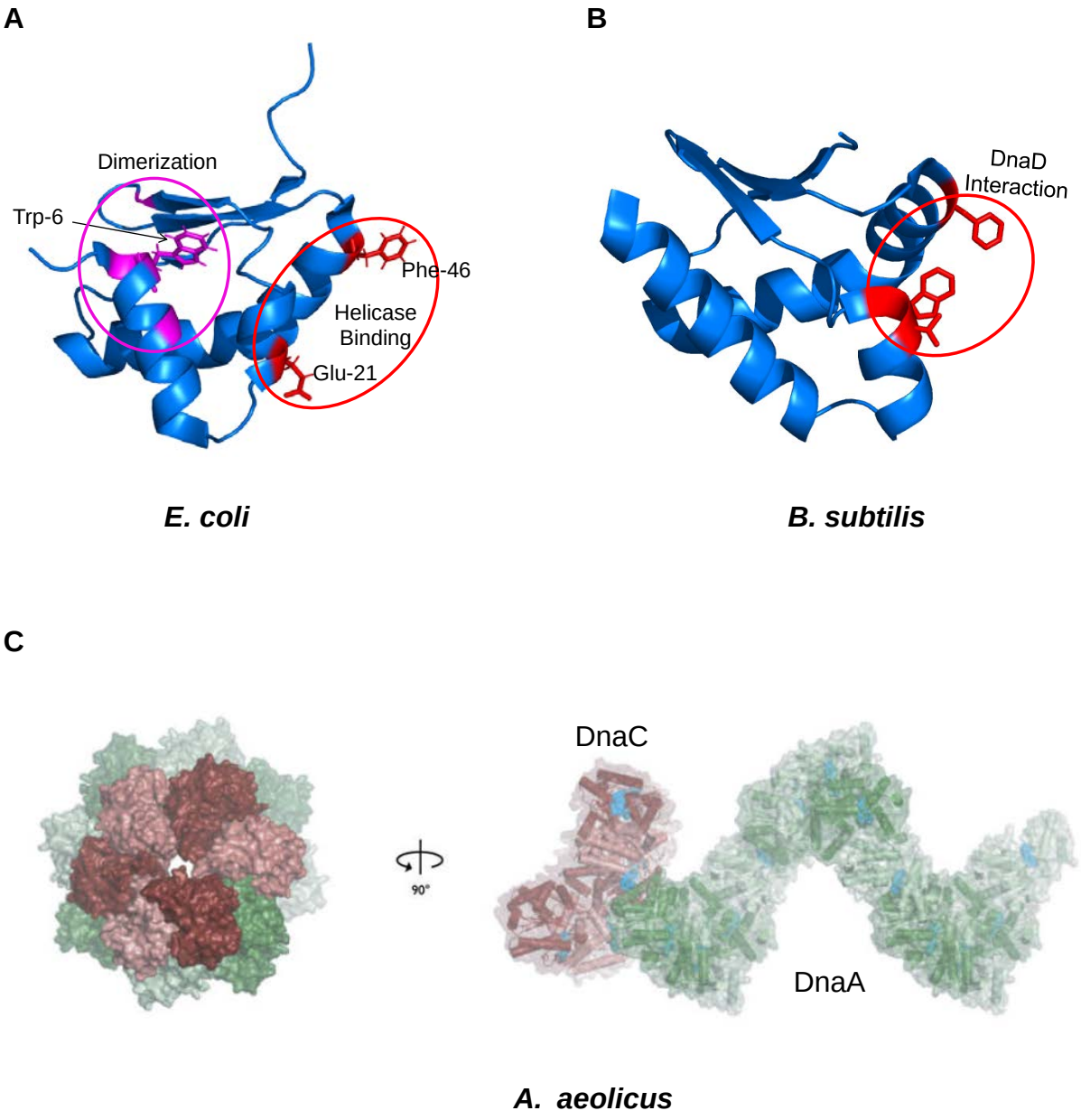


Figure 1.13. Protein-protein interactions of DnaA. Crystal structures for domain I of DnaA from **(A)** *E. coli* (PDB ID 2E0G) and **(B)** *B. subtilis* (4TPS), residues and binding patches involved in various protein-protein interactions are indicated. **(C)** Structural model for oligomeric DnaA-DnaC interactions showing the axial and side views (Mott *et al.*, 2008). Blue spheres represent bound nucleotides.

1.4. Remodelling and opening bacterial replication origins

1.4.1. Models for bacterial replication origin unwinding by DnaA

There have been several models proposed, based mostly on *in vitro* observations, which attempt to explain the mechanism by which DnaA opens the origin of replication. These models, outlined below and in Figure 1.14, generally agree that the opening of the origin occurs in three stages with origin opening being dependent upon the formation of the DnaA oligomer; I) DnaA recognises the origin and remains bound to high affinity binding sites throughout the cell cycle, II) low affinity binding sites are bound in a cell-cycle dependent manner coupled with the formation of an ATP-dependent filament, III) the unwinding region is unwound.

Super-helical strain model – As mentioned earlier, DnaA forms a helical filament at the origin. The simplest model for origin opening is that DNA wrapping around the filament could be similar to a positive supercoil generating a compensatory negative supercoil in the neighbouring unwinding region, which coupled with the lower stability of the AT-rich repeats leads to the melting of the DNA duplex (Figure 1.14.A). Subsequent binding of the DnaA filament to a single strand of the unwound DNA could maintain the open conformation (Duderstadt *et al.*, 2011).

Direct DnaA interactions model – DNA wrapped around the DnaA filament could allow protomers of the oligomer to directly engage with and melt the unwinding region. DnaA bound to ssDNA has been shown to alter the DNA conformation, stretching it so that each trinucleotide segment is separated by large gaps extending the DNA ~50% (Section 1.3.4) (Duderstadt *et al.*, 2011). DNA in this conformation is unable to base pair with the complementary strand, thereby leading to melting. Therefore, it is possible that the DnaA oligomers are binding to the site of unwinding and distorting the region to generate single strands (Figure 1.14.A) (Duderstadt *et al.*, 2011; Jha *et al.*, 2016).

Super-helical strain and direct DnaA interactions model – A third model is a combination of the two models above with negative supercoiling and direct DnaA interactions both involved in opening the origin. If this is the case a likely scenario would be the supercoiling partially melting the dsDNA, followed by engagement of ssDNA by DnaA to extend the open complex enough for helicase loading (Duderstadt and Berger, 2013).

Two-state DnaA assembly model – Duderstadt *et al.,* 2010 proposed that DnaA-filaments assume two distinct conformations for binding duplex or ssDNA (Section 1.3.3). The same investigation also proposed a model for how the two conformations work together to promote origin opening, highlighted in Figure 1.14.B. For the first two stages (binding high and low affinity binding sites) the extended state allows binding to the dsDNA DnaA-boxes. The high affinity sites would act as anchor points for promoting additional binding to lower affinity sites through ATP mediated interactions. The extended state is believed to be a less stable oligomer, but the proximity of the DnaA-boxes would increase the local concentration of protein assisting in protomer association and filament stability. The melted DUE is then engaged by a compact state filament. Without DnaA-box binding stabilising the filament, the protein assembly is stabilised through the DBD/AAA+ interaction (Duderstadt *et al.*, 2010). Whether the origin is being directly opened through DnaA interactions or if the filament is capturing the unwound DNA after spontaneous opening through super helical strain are both possibilities in this model.

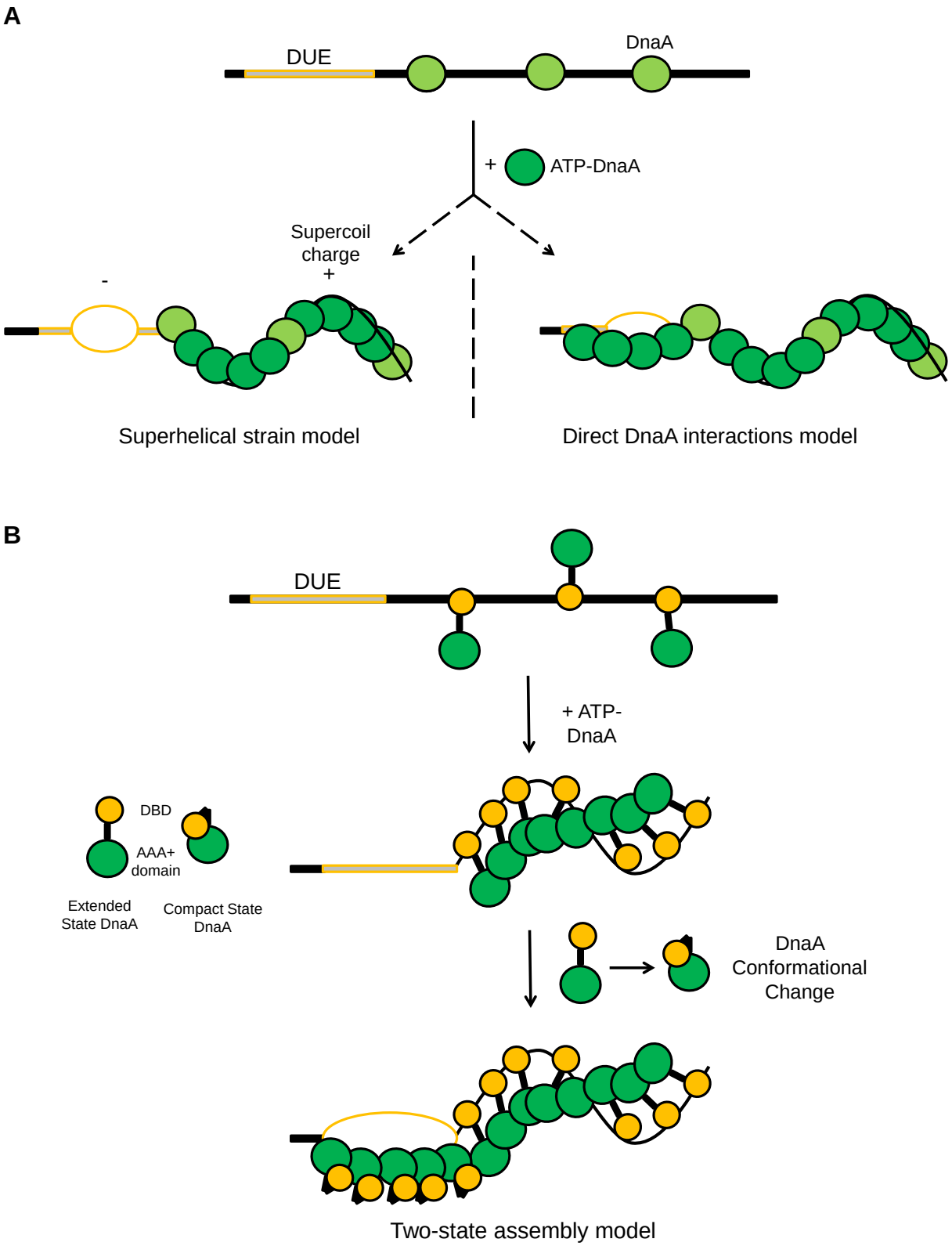


Figure 1.14. Models for opening the origin of replication. (A) Models for how DnaA unwinds the DUE either through superhelical strain or direct DnaA interactions. **(B)** The two-state assembly model. DnaA in an extended state binds the origin using the DBD and forms an ATP-dependent filament before undergoing a conformational change into a compact state for binding a single strand of the DUE via the AAA+ domain. Figure adapted from Duderstadt *et al*., 2010 and Duderstadt *et al*., 2011.

1.4.2. Opening of the E. coli origin of replication

The *oriC* of *E. coli* was discussed in Section 1.3 and is outlined in greater detail in Figure 1.15.A. The *E. coli* origin is located some distance away from the *dnaA-dnaN* operon, unlike the origins of other bacteria which are located proximal to, or surrounding the *dnaA* gene. The origin is formed of high and low affinity DnaA-boxes, an AT-rich DUE and binding sites for regulatory and accessory origin binding proteins (OBP). The layout of the origin and the presence of these species specific OBP have led to the proposal of alternate models to those outlined above for how DnaA opens the *E. coli* origin of replication.

The models for *E. coli oriC* opening by DnaA agree that DnaA remains bound to high-affinity DnaA-boxes for much of the cell cycle and that binding to the low affinity sites requires DnaA-ATP. The accessory protein Fis prevents the extension of the DnaA filament, regulating the timing of initiation. As DnaA-ATP concentration increases, oligomerisation of DnaA occurs from DnaA-box R4 to C3 causing the disassociation of Fis (Jha *et al.*, 2016). Fis displacement is believed to remove a steric barrier that allows IHF to bind the origin. IHF binding and bending leads to more DnaA-ATP associating with lower affinity binding sites and extension of the DnaA filament (Figure 1.15.B) (Wolański *et al.*, 2015).

The DnaA molecules, along with the DNA bending protein IHF, introduce a bend in the DNA helix which gradually wraps around the DnaA filament. This induces super-helical tension which focuses on the AT-rich repeats of the DUE. This leads to the initial unwinding of the DNA duplex generating single strands of DNA. The ssDNA-recruitment model (1.15.B) proposes that one of the single strands is engaged by the AAA+ domains of the DnaA-filament to prevent re-annealing and maintain the open complex. A speculative alternative to this model is that following origin remodelling, further DnaA molecules form a distinct second filament that either binds to the open complex or unwinds the DNA through direct interactions. As shown in Figure 1.15.B this proposal can incorporate the two-state assembly model where DnaA adopts either an extended or compact conformation (Ozaki *et al.*, 2012; Sakiyama *et al.*, 2017).

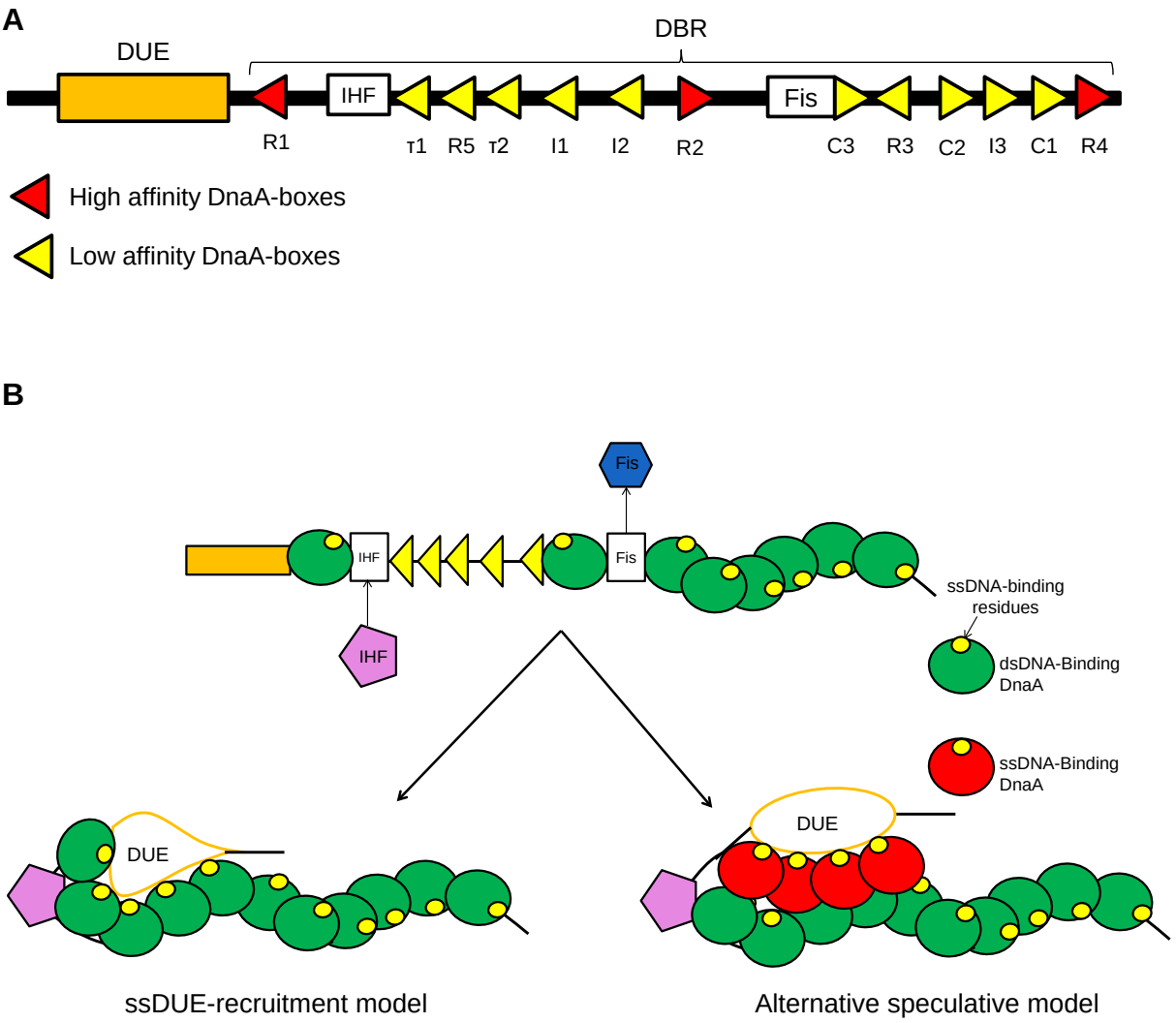


Figure 1.15. Models for the opening of the *E. coli* origin of replication. (A) The origin of replication of *E. coli* showing the arrangement of key protein binding sites and the numbering and relative affinity of DnaA-boxes. **(B)** Proposed models for the unwinding of the *E. coli* DUE for simplicity only the ssDNA-binding residues are highlighted. Figure adapted from Sakiyama *et al*., 2017.

1.4.3. ssDNA recruitment or continuous DnaA filaments?

The various models presented above outline the possible mechanisms by which DnaA could open the origin of replication. As already alluded, there are two alternate proposals for the mechanism by which DnaA engages the DNA of the unwinding region.

One proposal is that the unwinding region is engaged or captured by the DnaA filament bound to the DnaA-boxes, termed ssDNA recruitment. This proposal is formed from the appreciation that DnaA utilises different domains for binding double or single stranded DNA. As such DnaA would bind to the dsDnaA-boxes using the DBD leaving the AAA+ domain free to either capture a single strand of an already opened unwinding region, or directly engage the DNA in this location and melt it via a direct interaction (as described in Section 1.4.2 and Figure 1.15.B).

The second proposal is that DnaA forms a filament upon the DNA of the unwinding region. This filament is an extension of the filament that has formed on the DnaA-boxes and as such this proposal is termed the continuous filament. The filament could form on a strand of an already unwound origin to maintain or extend the open complex. Alternatively the filament could engage a single strand of the unwinding region and melt the duplex by stretching. Continuous filaments were shown in the models presented in Figure 1.14.

1.4.4. The DnaA-trios and unwinding of the origin

A key step in the models outlined above is that during the unwinding reaction DnaA engages a single DNA strand. Whether this is for directly unwinding the duplex through DNA stretching or for maintaining/extending an open complex is still debated.

As discussed in Section 1.2.3, the recently identified DnaA-trios are an essential repeating trinucleotide motif located within *oriC* that were shown to stabilise DnaA filaments on a single strand of DNA, thus providing specificity to the ssDNA binding at the origin (Richardson *et al.*, 2016). The DnaA-trios appear to be widespread throughout the bacterial kingdom, suggesting they are a key component of *oriC* architecture. The models outlined above were proposed before the identification of the DnaA-trio motif, but the element could quite easily fit into any of them, in that the DnaA-trio sequence could be either engaged and stretched to melt the origin or captured to maintain an open complex.

During the course of identifying the DnaA-trio element, genetic analysis was performed on the wider unwinding region of the *B. subtilis* origin (*incC*) (Figure 1.16.A). Through this it was established that the two DnaA-boxes proximal to the site of unwinding were critical for origin activity. This led to the proposal of an updated model for origin opening, outlined in Figure 1.16.B. This model proposes that the proximal DnaA-boxes direct the DnaA-filament onto the DnaA-trios, with a single DnaA protein binding dsDNA via the DBD before engaging a DnaA-trio via its AAA+ motif. It is further proposed that the method for origin unwinding is through direct DnaA interactions stretching the DnaA-trios. The model proposed is consistent with the two-state DnaA assembly and direct DnaA interactions continuous-filament models (Richardson *et al.*, 2016).

Interestingly, in *B. subtilis* where the DnaA-trios were first identified, there are 6 repeating trinucleotide motifs representing a potential initial unwinding region of 18 base pairs, which is in good agreement with the 20-60 bp initially unwound in the unwinding regions previously identified in bacteria (Section 1.2.3). In *B. subtilis,* directly downstream of the DnaA-trios is an AT-cluster 27 base pairs in length (Figure 1.16.A). This region has been shown to be intrinsically unstable (Krause *et al.*, 1997) and so could be opened by DnaA bound to the DnaA-trios to extend the open complex, making more space to accommodate helicase loading.

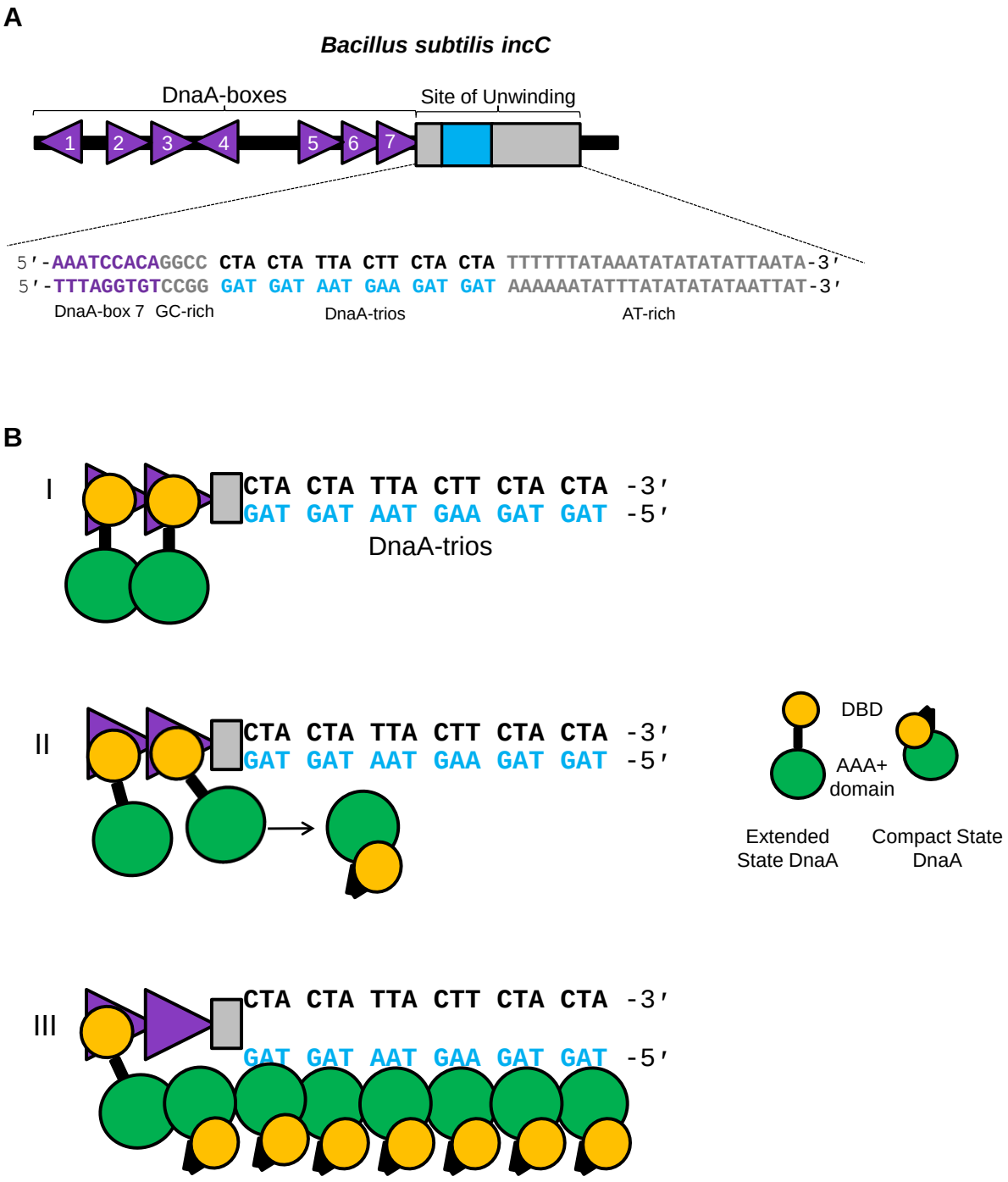


Figure 1.16. Model for DnaA filament formation on DnaA-trios. (A) The sequence and layout of the unwinding region of the *Bacillus subtilis* origin, *incC*. **(B)** Model showing how the DnaA filament is loaded from the double-stranded DnaA-boxes (purple triangle) onto a single DNA strand containing the DnaA-trios (light blue) via a conformational change in the state of DnaA. Figure adapted from Richardson *et al*., 2016.

1.5. The replication initiation machinery of *Bacillus subtilis*

Once the origin of replication has been recognised, remodelled and opened, the final step of initiation is the coordinated loading of two replicative helicases onto the opposite strands of DNA to begin assembly of the replisome. In *E. coli* the replicative helicase is loaded directly onto the single-stranded DNA by DnaA and a helicase loader protein (Section 1.1.1).

The process of helicase loading in *B. subtilis* involves not just DnaA and the helicase loader (DnaI) but two unique accessory proteins, DnaD and DnaB (Section 1.3.5). Together these proteins form the initiation machinery in *B. subtilis* (and related bacteria). As there are no known homologues for DnaD or DnaB outside the Firmicutes, this helicase loading complex operates in a way distinct from other bacteria (Briggs *et al.*, 2012). Table 1.2 compares the proteins involved in replication initiation in *B. subtilis* and *E. coli*. The helicase loaders for both bacterial species are AAA+ proteins.

	Protein Name	
Function	***B. subtilis***	***E. Coli***
Master Initiator	DnaA	DnaA
Accessory/Remodelling Protein	DnaB, DnaD	IHF, Fis
Replicative Helicase	DnaC	DnaB
Helicase Loader	DnaI	DnaC

Table 1.2. Comparisons of the proteins involved in replication initiation in *B. subtilis* and *E. coli*

During initiation in *B. subtilis* it has been determined through chromatin immunoprecipitation (ChIP) assays that the initiator proteins are recruited to the origin in a hierarchical manner, beginning with DnaA and then followed by, in sequence, DnaD, DnaB and the DnaI:DnaC complex (Smits *et al.*, 2010).

1.5.1. DnaD and DnaB: The B. subtilis accessory and remodelling proteins

DnaD and DnaB are structurally related, essential proteins sharing a similar architecture formed of two and three domains respectively as mapped in Figure 1.17.A. DnaD is a formed of an N-terminal winged-helix domain and a C-terminal domain, which will be referred to as the DnaD NTD and CTD, respectively (Briggs *et al.*, 2012).

DnaB is composed of an NTD and a CTD sharing homology to those of DnaD. DnaB is also formed of a third degenerate domain between the two, referred to as the middle domain (MD) that is structurally and functionally similar to the CTD (Figure 1.16.A) (Li *et al.*, 2017).

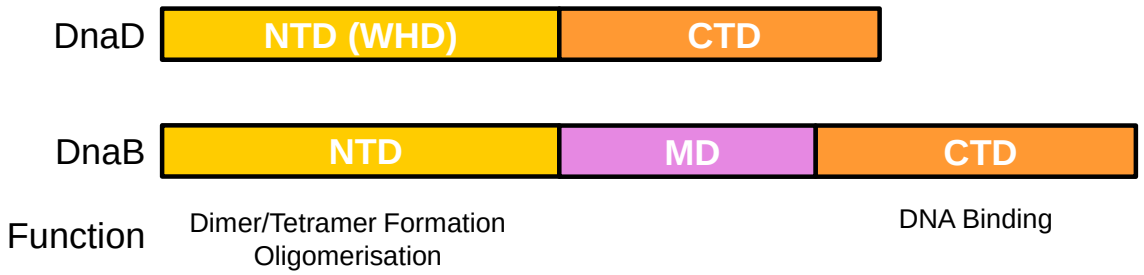
Crystal structures for both domains of DnaD from *B. subtilis* are shown in Figures 1.17.B-C, while the first two domains of DnaB from *Geobacillus stearothermophilus* are shown as a tetramer in Figure 1.17.D.

The NTD of DnaD (Figure 1.17.B) has a winged-helix with two extensions, a helix-strand-helix at the N-terminus and a single helix at the C-terminus. The helix-strand-helix is involved in the formation of DnaD dimers and tetramers. $DnaD^{NTD}$ has been shown structurally and biochemically to be involved in higher-order oligomerisation and capable of forming large scaffolds (Briggs *et al.*, 2012). The NTD of DnaB has also been shown to be involved in the formation of tetramers and a linker region attaches it to the middle domain (Figure 1.17.D) (Li *et al.*, 2017).

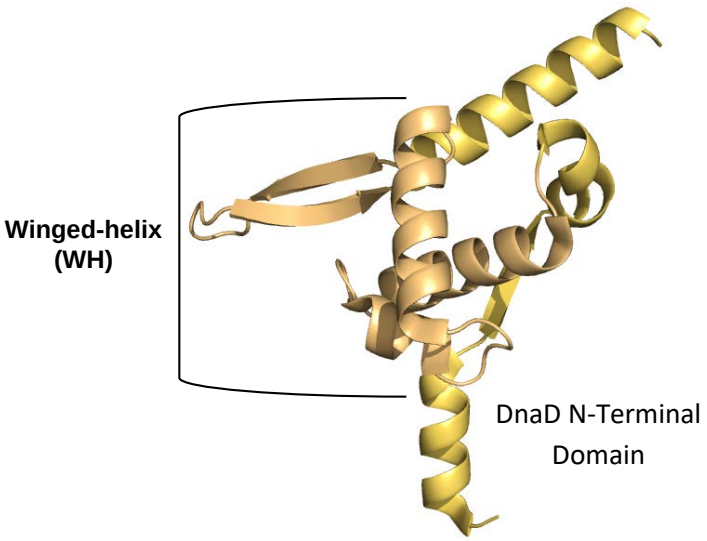
The CTD (Figure 1.17.C) of both DnaD and DnaB contains double and single-stranded DNA binding activity. The proposed dsDNA binding site is located near the C-terminus and encompasses a highly conserved motif YxxxIxxxW (Marston *et al.*, 2010).

Both DnaD and DnaB proteins are capable of performing several different functions. Both proteins show DNA remodelling activity, with DnaD shown to be able to bend and unwind DNA while DnaB can laterally compact it (Briggs *et al.*, 2012). DnaD has been shown to interact with DnaA. DnaB on the other hand interacts with the replicative helicase (DnaC) and so is believed to cooperate with DnaI as a helicase co-loader (Marston *et al.*, 2010). Both DnaD and DnaB interact with one another, leading to speculation that DnaB may act as a bridge between an *oriC*-DnaA-DnaD complex and a DnaC-DnaI complex (Briggs *et al.*, 2012). Finally, DNA replication in bacteria is membrane-associated and this membrane association in *B. subtilis* is DnaB dependent (Sueoka, 1998). Therefore, it is proposed that membrane-bound DnaB recruits DnaD to the membrane. DnaD in turn recruits DnaA bound to *oriC* thereby recruiting the initiation complex to the membrane. The purpose of membrane attachment during DNA replication initiation is unclear although may be linked to arranging and anchoring the chromosomes ready for accurate segregation (Badrinarayanan *et al.*, 2015).

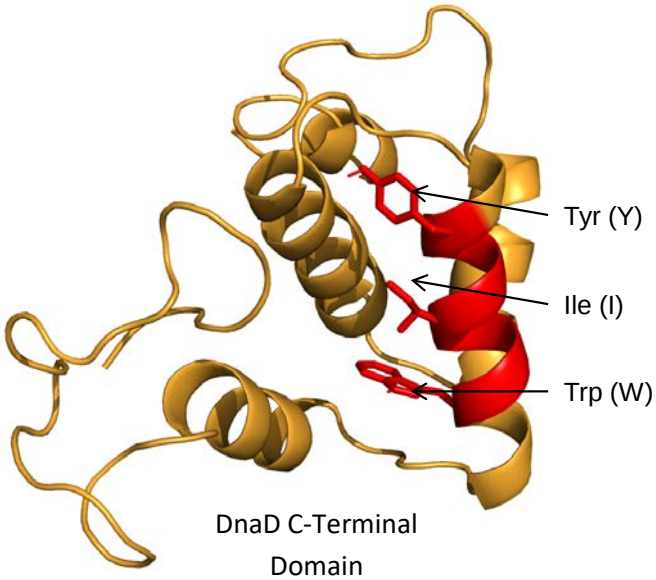
A


DnaD
NTD (WHD)
CTD
DnaB
NTD
MD
CTD
Function
Dimer/Tetramer Formation
Oligomerisation
DNA Binding

B


Winged-helix
(WH)
DnaD N-Terminal
Domain

C


Tyr (Y)
Ile (I)
Trp (W)
DnaD C-Terminal
Domain

D

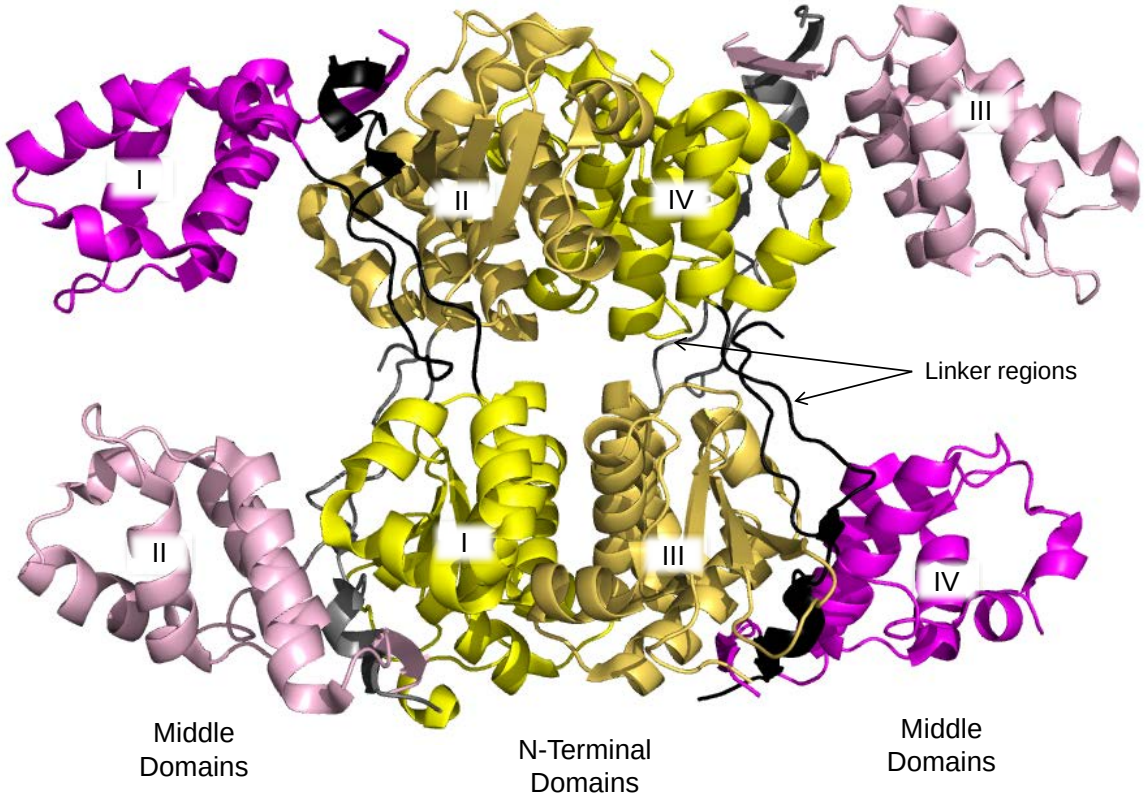


Figure 1.17. The *B. subtilis* accessory proteins DnaD and DnaB. (A) Schematic of the domain organisation of DnaD and DnaB along with the functions these domains perform. Structurally homologous domains between proteins are highlighted in the same colour. The pale orange CTD of DnaB is degenerate. **(B)** Structure of the *B. subtilis* DnaD Winged-helix domain (PDB ID 2V79). The winged-helix fold is highlighted. **(C)** Structure of the *B. subtilis* DnaD C-terminal domain (Marston *et al*., 2010) highlighting the conserved motif involved in dsDNA binding in red, with the highly conserved residues labelled. **(D)** Structure of a DnaB tetramer of the N-terminal and middle domains of *G. stearothermophilus* (PDB ID 5WTN).

1.5.2. DnaI: The replicative helicase loader

The replicative helicase of *B. subtilis*, DnaC, and the helicase loader, DnaI, are homologous to their *E. coli* counterparts. While this section focuses on the *B. subtilis* helicase loader, DnaI, several of the functions of the AAA+ domain are based upon those found for the homologous loader from *E.coli* due to similar activities being relatively unexplored in *Bacillus*.

DnaI is formed of two functional domains, mapped in Figure 1.18.A, an N-terminal (NTD) and a C-terminal (CTD) domain. The NTD (Figure 1.18.B) is required for the interaction of DnaI with the helicase and the N-terminal region of the *E. coli* helicase loader is also believed to be involved in helicase binding. The NTD of DnaI carries a zinc-binding fold (Figure 1.18.B) involved in the interaction with the helicase which is mediated by zinc ions (Loscha *et al.*, 2009).

The C-terminal domain (Figure 1.18.C) is a AAA+ domain (structurally homologous to DnaA domain III) capable of binding and hydrolysing ATP. In the *E. coli* helicase loader ATP binding induces a conformational change to the protein increasing its affinity for the helicase, while ADP-binding is a negative effector. The CTD is also capable of binding ssDNA, required for recruiting the helicase to the open origin complex. This ssDNA binding activity however, is only observed in the presence of the helicase (Ioannou *et al.*, 2006).

The observation that ssDNA binding is dependent upon helicase binding has led to the proposal that the loader NTD acts as a ‘molecular switch’ regulating the availability of the CTD DNA-binding site. This site is normally buried in the protein but the binding of the helicase to the loader NTD induces a conformational change, exposing the binding site so the loader can bind DNA and deliver the helicase to the unwound origin. ATP hydrolysis ejects the loader from the complex leaving an active helicase bound to the DNA. An overview for this model is shown in Figure 1.18.D (Ioannou *et al.*, 2006).

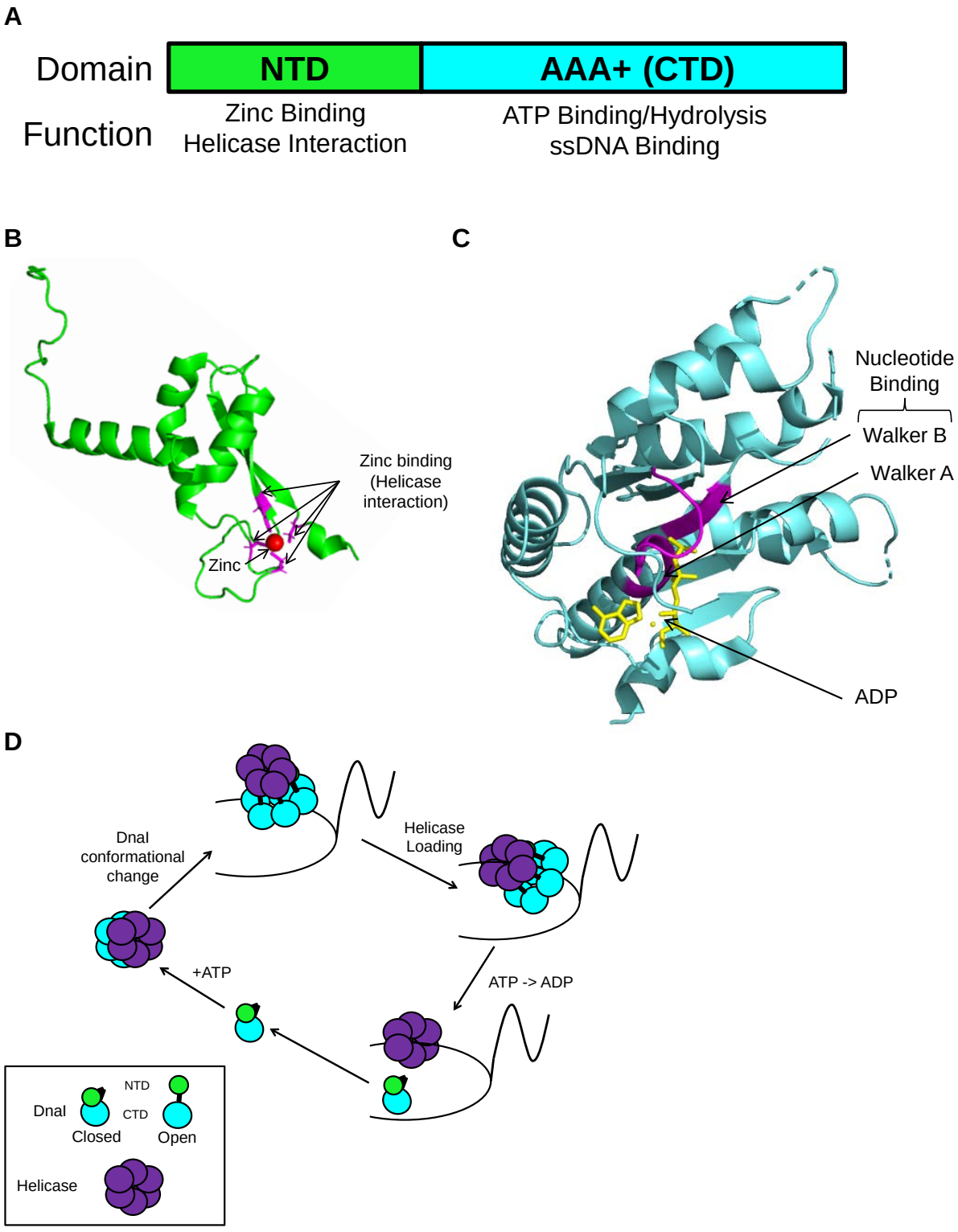


Figure 1.18. The *B. subtilis* helicase loader DnaI. (A) Schematic of the domain organisation of DnaI along with the functions these domains perform. **(B)** Structure of the *B. subtilis* DnaI N-terminal domain (PDB ID 2K7R) and **(C)** the *Geobacillus kaustophilus* DnaI AAA+ domain (PDB ID 2W58). Key residues and motifs are highlighted. **(D)** Model for DnaI-mediated helicase loading adapted from Ioannoui *et al.,* 2006.

1.5.3. The B. subtilis primosomal complex and initiation of DNA replication

In *B. subtilis* it has been determined that the initiator proteins are recruited to the origin in a hierarchical manner leading to the proposal of a helicase loading pathway which has been outlined in Figure 1.19.A (Smits *et al.*, 2010). The observations about DnaD, DnaB and DnaI discussed above have led to speculative models for the role these proteins play in the initiation of DNA replication.

As mentioned in section 1.5.1 DNA replication is membrane associated and DnaB and DnaD are proposed to be responsible for recruiting the initiation complex (DnaA bound *oriC*) to the membrane. The model for the recruitment of the initiation complex to the membrane and the association of the helicase/helicase loader complex with *oirC* is shown in Figure 1.19.B. As the model proposes DnaA binds the origin and interacts with DnaD. The interaction between DnaD and DnaB recruits the complex to the membrane. Finally the helicase/loader complex is recruited (Smits *et al.*, 2010).

As described in section 1.2.2 the *B. subtilis oriC* is bipartite formed of two distinct regions separated by the *dnaA* gene. Studies using electron microscopy have identified that DnaA bound to the regions upstream and downstream of *dnaA* interact causing the DNA to loop (Krause *et al.*, 1997). Another model for *B. subtilis* initiation has therefore been speculated where DnaD is recruited to both halves of the origin by DnaA, where it recruits additional DnaD molecules which bind the central DNA extending into a scaffold. This DnaD scaffold may bend the DNA bringing both halves of the bipartite *oriC* together and the two DnaA complexes into proximity for interaction, which could be stabilised by DnaD bridging. DnaB would then be recruited, the unwinding region opened and the helicase loader complex recruited (Figure 1.19.C). It is further speculated that the two initiator protein complexes forming either side of *dnaA* are involved in loading the helicase onto the opposite strands of the open complex (Briggs *et al.*, 2012).

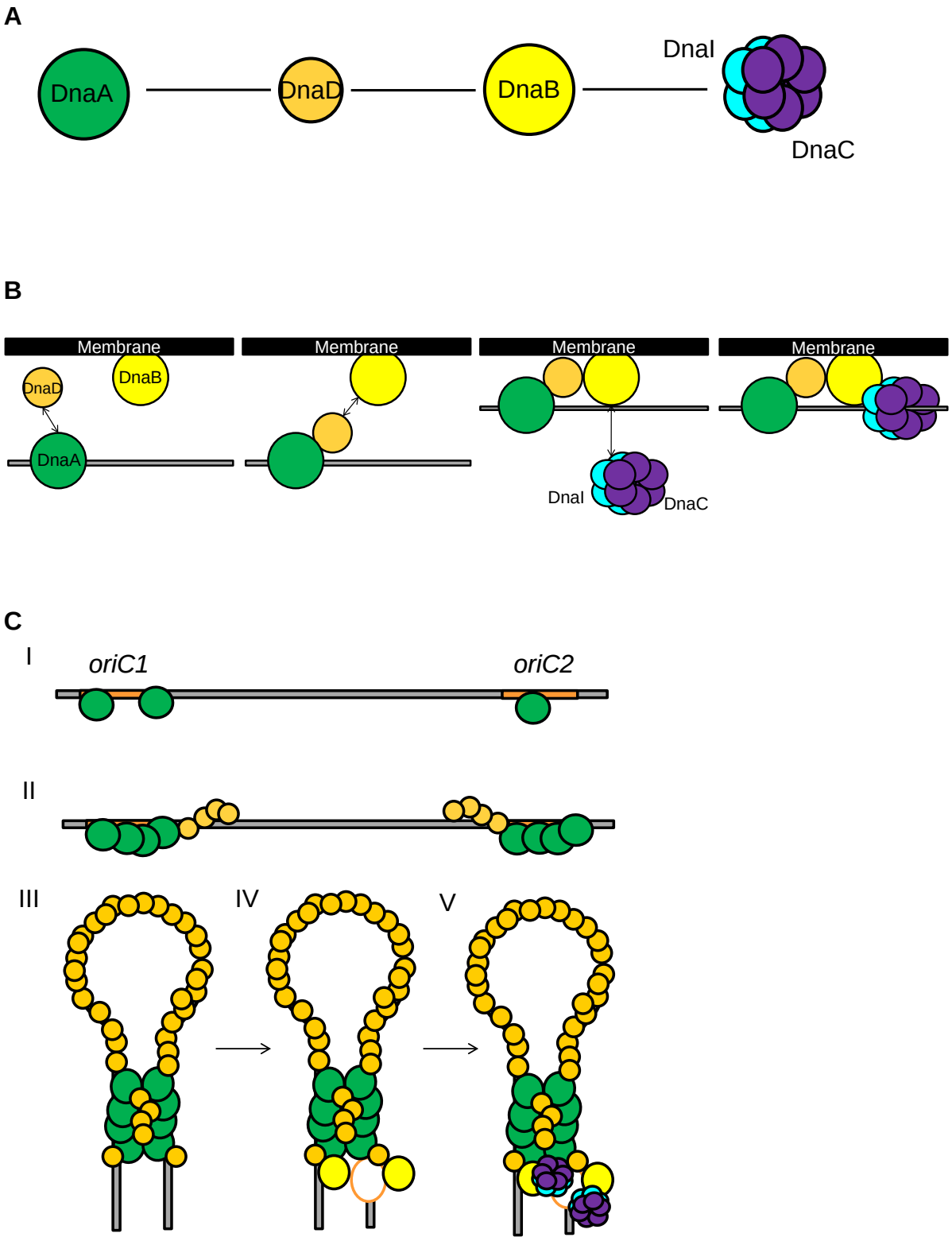


Figure 1.19. *B. subtilis* primosomal complex and the initiation of DNA replication. **(A)** Proposed helicase loading pathway of *B. subtilis* showing the hierarchical order of recruitment to *oriC.* **(B)** Model for the association of the helicase with *oriC* and recruitment of the initiation complex to the membrane. Adapted from Smits *et al*., 2010. **(C)** Speculative model for the role of DnaD in *B. subtilis* DNA replication initiation. Adapted from Briggs *et al*., 2012.

1.6. Regulation of DNA replication initiation in bacteria

During rapid growth bacteria initiate replication before the previous round of synthesis is complete, daughter cells therefore inherit chromosomes undergoing replication. This has given rise to mechanisms to regulate the frequency of initiation so it matches the frequency of cell division (Jameson *et al.*, 2014). Due to its role as the major protein for initiating replication, DnaA is a primary target for regulation (Katayama *et al.*, 2010).

Bacterial species have evolved many mechanisms to regulate initiation as outlined in Figure 1.20. The most common mechanism appears to be the inactivation of DnaA-ATP via ATP hydrolysis. DnaA is only able to initiate replication when bound by ATP, so the hydrolysis of ATP following initiation inactivates the protein preventing immediate re-initiation (Zakrzewska-Czerwinska *et al.*, 2007). Other strategies employed to regulate initiation appear to be more species specific.

Several of these strategies involve regulating *oriC* binding. Newly replicated DNA is unmethylated, thus generating an asymmetry where one strand is methylated and the other not. Hemimethylated DNA is the substrate for DNA adenine methylase (Dam) which re-methylates adenines at GATC sites (Russell and Zinder, 1987). The *E. coli oriC* contains multiple GATC sequences (highlighted as SeqA binding sites in Figure 1.6.A) which when hemimethylated provide a high affinity binding site for the regulatory protein SeqA. SeqA binds to the origin, outcompeting Dam for the GATC sites, sequestering *oriC* and inhibiting re-initiation. SeqA is not dissociated by Dam but rather spontaneously dissociates after a short period of time. This delays new rounds of replication allowing for the completion of elongation (Kang *et al.*, 1999). These GATC sites also overlap with low affinity DnaA-boxes, resulting in SeqA blocking DnaA filament formation providing another mechanism for inhibiting replication (Waldminghaus and Skarstad, 2009).

In bacteria that can undergo differentiation such as *B.* subtilis, *C. crescentus* and *Streptomyces coelicolor* it has been observed that developmentally expressed regulatory proteins (Spo0A, CtrA and AdpA respectively) bind to specific sites located within *oriC* and prevent DnaA binding to represses further rounds of DNA synthesis (Figure 1.20) (Castilla-Llorente *et al.*, 2006; Taylor *et al.*, 2011; Wolanski *et al.*, 2012).

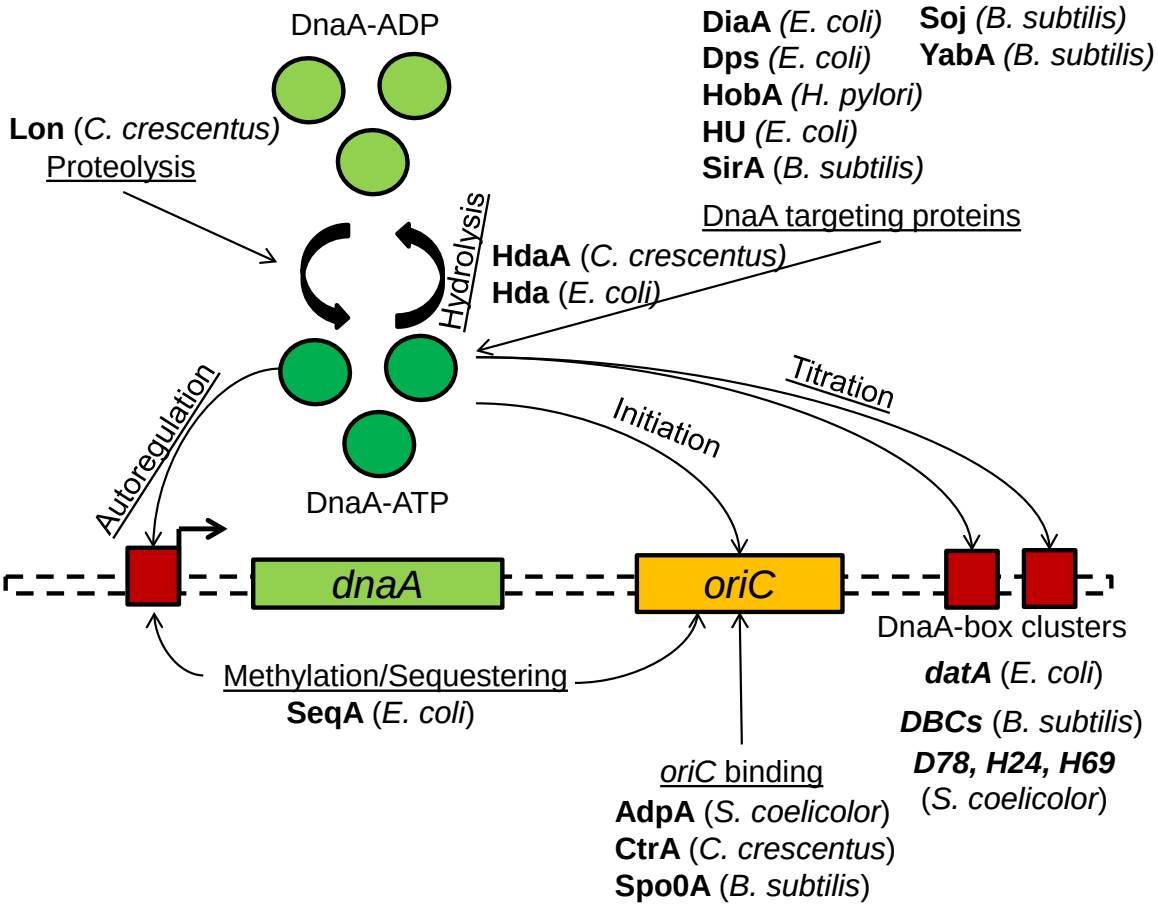


Figure 1.20. Stratergies for regulating the initation of DNA replication in Bacteria. Mechanisms bacteria use to regulate initation and where they act. Stratergies are underlined, with examples of proteins/genomic features utilised shown in bold. Figure adapted from Zakrzewska-Czerwinska *et al*., 2007.

Another strategy bacteria have evolved to regulate initiation is to control the levels of free cellular DnaA. For example SeqA can inhibit transcription of the *dnaA* gene to ensure replication only occurs once. The *E. coli dnaA* gene is located near *oriC* and contains GATC sequences within the promoter region which become hemi-methylated during replication and so bound by SeqA preventing transcription (Figure 1.20) (Katayama *et al.*, 2010). The *dnaA* gene is also believed to be autoregulated. DnaA binding sites have been identified in the *dnaA* promoter region, some of which require DnaA-ATP, suggesting that as protein levels increase, DnaA binds and autorepresses its promoter (Hansen and Atlung, 2018). A final strategy for limiting DnaA levels involves titrating DnaA away from *oriC*. In *E. coli* five DnaA-boxes are located at the *datA* locus (~1kb region) downstream of the origin (Kitagawa *et al.*, 1998), while in *S. coelicolor* three clusters of 5-6 DnaA-boxes (D78,H24,H69) are localised near *oriC* (Smulczyk-Krawczyszyn *et al.*, 2006), and finally in *B. subtilis* six DnaA-box clusters (DBCs) have been identified located away from *oriC* (Figure 1.20) (Okumura *et al.*, 2012). Following genome duplication titration of DnaA to these binding sites contributes to preventing re-initiation.

Table 1.3 lists some of the known proteins which regulate DnaA or initiation via DnaA. These are also highlighted in Figure 1.20.

Regulator	Role	Organism	Reference
DiaA	Binds DnaA and stimulates formation of a DnaA filament	*E. coli*	(Keyamura *et al.*, 2009)
Dps *(DNA-binding protein)*	Interacts with DnaA NTD, binds *oriC* non-specifically, affects origin opening	*E. coli*	(Chodavarapu *et al.*, 2008b)
Hda	Stimulates ATP hydrolysis Inactivating DnaA-ATP	*E. coli*	(Nakamura and Katayama, 2010)
HdaA	Stimulates ATP hydrolysis Inactivating DnaA-ATP	*C. crescentus*	(Frandi and Collier, 2019)
HobA	Interacts with DnaA, similar function to DiaA	*H. pylori*	(Zawilak-Pawlik *et al.*, 2007)
HU *(histone like U-factor)*	Interacts with DnaA NTD, binds *oriC* non-specifically, affects DnaA oligomer stability	*E. coli*	(Chodavarapu *et al.*, 2008a)
Lon	Degrades DnaA during proteotoxic stress	*C. crescentus*	(Jonas *et al.*, 2013)
SirA	Binds DnaA inhibiting initiation of DNA replication during sporulation	*B. subtilis*	(Rahn-Lee *et al.*, 2009)

Soj (Monomer)	Binds DnaA inhibiting oligomerisation	*B. subtilis*	(Scholefield *et al.*, 2012)
YabA	Binds DnaA inhibiting oligomerisation	*B. subtilis*	(Cho *et al.*, 2008)

Table 1.3. Proteins which regulate initiation through targeting DnaA.

Hda is a component of the regulatory inactivation of DnaA (RIDA) system in *E. coli* with homologues identified in other *Gammaproteobacteria*. Hda contains an AAA+ domain which interacts with DnaA. This interaction stimulates the hydrolysis of the ATP bound to DnaA resulting in the formation of inactive DnaA-ADP (Zakrzewska-Czerwinska *et al.*, 2007; Katayama *et al.*, 2010). HdaA from *C. crescentus* is proposed to inactivate DnaA through a similar mechanism (Frandi and Collier, 2019).

While ATP binding appears to be the mechanism used to inactivate DnaA in proteobacteria, oligomerisation has been suggested as the main regulatory target for *B. subtilis* DnaA. Monomeric Soj directly interacts with the AAA+ domain of DnaA and inhibits the formation of the helical filament required for opening the origin of replication. It is proposed that Soj could inhibit both the formation of the filament bound to dsDNA and that forming on ssDNA. It is further proposed Soj could prevent the bending of the domain III-IV junction (Figure 1.8) preventing formation of the compact state protein proposed as being required for ssDNA binding (Section 1.3.3) (Scholefield *et al.*, 2012). YabA is another negative regulator of DnaA that specifically inhibits DnaA oligomerisation, however it interacts with a different patch of the DnaA AAA+ domain to that of Soj. YabA is proposed to inhibit filament formation on either ds- or ssDNA (Cho *et al.*, 2008; Scholefield and Murray, 2013).

DiaA is a DnaA binding protein that is required for initiation to occur in a timely manner. DiaA has been shown to stimulate the formation of DnaA filaments at *oriC*, which in turn leads to origin opening and the initiation of DNA replication (Keyamura *et al.*, 2009). DiaA forms a tetramer which binds multiple DnaA molecules bringing them together and facilitating ATP-DnaA-DnaA interactions. It has been proposed that DiaA-DnaA binding overcomes the issue of DnaA diffusion during the limited time period when ATP-DnaA molecules must assemble at the origin to ensure timely initiation. DiaA has also been demonstrated to stimulate the interaction of ATP-DnaA with lower affinity binding sites, further supporting filament formation (Keyamura *et al.*, 2007). HobA from *H. pylori* is an essential structural homolog of DiaA that is proposed to

perform a similar a function of stimulating DnaA complex formation at the origin (Zawilak-Pawlik *et al.*, 2011).

During times of low nutrient availability certain Gram-positive bacteria can undergo a process of differentiation known as sporulation, which ultimately results in the formation of a highly resistant endospore, a state these cells can remain in until conditions improve (Veening *et al.*, 2009). Sporulation involves an asymmetric cell-division producing daughter cells of unequal size, a larger mother spore and smaller forespore, each of which must inherit a complete copy of the genome. The first stage of sporulation is defined by chromosome condensation and the anchoring of the origin of replication to the cell pole. DNA replication is regulated at the onset of sporulation to ensure just two intact copies of the chromosome are present and initiation is prevented to ensure just two origins are present (Jameson *et al.*, 2014; Tan and Ramamurthi, 2014).

DnaA contributes to this regulation though its role as a transcription factor. In *B. subtilis* the expression of sporulation-specific genes depends on the transcriptional regulator Spo0A which is phosphorylated following the decision to sporulate. DnaA activates expression of *sda* which encodes an inhibitor of the kinases which activate the phosphorelay that leads to the phosphorylation of Spo0A (Higgins and Dworkin, 2012). The regulation of *sda* by DnaA establishes a checkpoint preventing cells from attempting sporulation when DNA replication initiation is impaired (Burkholder *et al.*, 2001). The expression of intrinsically unstable Sda requires active DnaA and occurs in a pulsatile manner with a burst of expression at the onset of replication. This inhibits the initiation of sporulation while cells are actively replicating chromosomes, thereby avoiding inviable polyploidy spores (Veening *et al.*, 2009).

Re-initiation of replication is prevented by the inhibitor SirA, one the sporulation-specific proteins whose expression depends on Spo0A. SirA has been demonstrated to bind to DnaA inhibiting its functioning and preventing further initiation of DNA replication. The exact molecular mechanism for how this is achieved remains unknown (Rahn-Lee *et al.*, 2009).

Chapter 2

Materials and Methods

2.1. General techniques for DNA manipulation

2.1.1. Oligonucleotide synthesis

All oligonucleotides used during the course of this study were designed using Clone Manger version 9 and/or QC_Primer_Generator (Created by Theodor Sperlea) and purchased from Eurogentec at a concentration of 100 µM in Milli-Q water. The full list of oligonucleotides used is found in Table 2.4.

2.1.2. Polymerase Chain Reaction

PCR (Polymerase Chain Reaction) was used to amplify DNA fragments from either plasmid or genomic DNA templates. PCR was performed using either Q5 DNA polymerase (NEB) or Go Taq polymerase (Promega) in a reaction consisting of 1X the appropriate buffer, 200 µM dNTPs (Promega), 0.5 µM of each primer, ~1 ng template DNA or a bacterial colony and adjusted to a final volume of 25 or 50 µl with Milli-Q water. PCR was performed in a thermocycler (Techne or VWR) in a typical program of: 98°C for 60 seconds followed by 30 cycles of denaturation, annealing and extension (98°C for 10 seconds, annealing temperature for 30 seconds and 72°C for 1-5 minutes) and then a final extension of 72°C for 5 minutes. The annealing temperature was altered according to the specific primers and the extension time according to the size of product to be synthesised.

2.1.3. Agarose gel electrophoresis

DNA samples were added to 2 µl of bromophenol blue loading dye (Sigma-Aldrich) and 2 µl of SYBR gold (ThermoFisher Scientific) and visualised on a 1% agarose gel (Sigma-Aldrich) of 0.5X TBE (Tris-borate-EDTA) and ran in 0.5X TBE running buffer. Gels were visualised using UV or Blue light (~400 nm) transillumination with a 0.5 millisecond exposure time. A 1kb DNA ladder (Promega) was used to estimate molecular weight.

2.1.4. Gel extraction, PCR and plasmid purification

Purifications and extractions were performed using QIAquick Spin Miniprep kits (Qiagen) and centrifugations at 16200 x*g* for 1 minute (unless longer centrifugations were required).

DNA purification from a reaction was performed by adding 10X the reaction volume of buffer PB (5 M Gu-HCl, 30% isopropanol), loading into a spin column and centrifuging. 700 µl of PE (10 mM Tris-HCl pH 7.5, 80% ethanol) was added followed by two centrifugation steps. The column was transferred to a clean Eppendorf tube and 30 µl EB (10 mM Tris-HCl, pH 9.0) was added, stood for 5 minutes and centrifuged.

Gel extractions were performed by excising the required band from an agarose gel and dissolving in 500 µl QG (5.5 M guanidine thiocyanate (GuSCN), 20 mM Tris HCl pH 6.6) at 50°C. This was transferred to a spin column and centrifuged. 700 µl of PE (10 mM Tris-HCl pH 7.5, 80% ethanol) was added followed by two centrifugation steps. The column was transferred to a clean Eppendorf tube and 30 µl EB (10 mM Tris-HCl, pH 9.0) was added, stood for 5 minutes and centrifuged.

Plasmid extractions used 5 ml of overnight liquid culture centrifuged into a pellet. The pellet was re-suspended in 250 µl P1 (50 mM Tris-HCl pH 8.0, 10 mM EDTA, 100 µg/ml RNaseA), 250µl P2 (200 mM NaOH, 1% SDS) and 350 µl N3 (4.2 M Gu-HCl, 0.9 M potassium acetate, pH 4.8) sequentially and centrifuged for 10 minutes. The supernatant was transferred to a spin column and centrifuged. 500 µl buffer PB was applied to the spin column and centrifuged. 700 µl of PE was added followed by two centrifugation steps. The column was transferred to a clean Eppendorf tube and 30 µl EB was added, stood for 5 minutes and centrifuged.

2.1.5. Genomic DNA extraction

The extraction of genomic DNA for transformations started with a 2 ml overnight culture, used to inoculate 2.5 ml of LB in a 1:25 dilution and incubated for 3 hrs while shaking. The culture was pelleted after addition of 2.5 ml 1X SSC (saline sodium citrate buffer) at 7800 x*g* for 3 minutes, and then resuspended in 900 µl of 1X SSC. The suspension was incubated at 37°C with 20 µl of lysozyme (10 mg/ml stock) until clear. 1 ml 4 M NaCl was added and the suspension passed through a 0.45 µm Millipore filter.

Where genomic DNA was required as a PCR template the extraction and purification was performed using a DNeasy Blood and Tissue Kit (Qaigen) and 1 ml of an overnight liquid culture. Centrifugations were performed at 9600 x*g* for 1 minute unless otherwise stated. The culture was pelleted via centrifugation and resuspended in 180 µl lysis buffer (20 mM Tris-Cl pH 8, 2 mM EDTA and 1.2% triton X-100) with 2.5 µg/ml lysozyme and 0.25 µg/ml RNase, before incubation at 37°C for 20 minutes with shaking (800 rpm). 10 µl of Proteinase K and 200 µl Buffer AL was added before incubation at 56°C for 30 minutes with shaking (800 rpm). 200 µl of EtOH was added before transfer to a Qaigen column and centrifuged. 500 µl of Buffer AW1 then 500 µl of Buffer AW2 was added with centrifugation after each addition. The column was transferred to a clean Eppendorf tube and 200 µl Buffer AE (10 mM Tris-HCL pH 9.0, 0.5 mM EDTA) was added, stood for 1 minute and centrifuged at 12,000 rpm.

2.1.6. Restriction enzyme digestion

Restriction enzymes (NEB, Promega, Sigma-Aldrich) were used to digest ~1µg of DNA for either 3 hours at 37°C or overnight at 30°C, before purification (2.2.4). Reactions were performed in 1X of the appropriate buffer and adjusted to a final volume of 20 µl with Milli-Q water.

2.1.7. DNA fragment ligation

Restriction enzyme digested plasmids or purified PCR products were treated for 2 hrs with 1 µl thermosensitive alkaline phosphatase (Promega) prior to purification (2.1.4). The DNA fragments were then ligated in a 2:1 insert to vector ratio with 1 µl T4 DNA ligase (NEB) and 1x ligase buffer. The ligation was performed for 3 hrs at room temperature (20-23°C) or overnight at 15°C.

2.1.8. DNA sequencing

Plasmid and PCR products were sequenced using the MRC sequencing service in the School of Life Sciences at the University of Dundee, Scotland. Primers used for sequencing are listed in Table 2.4.

2.2. Plasmid construction

2.2.1. QuikChange mutagenesis

QuikChange (or site-directed) mutagenesis was used to generate *dnaA* mutant plasmids using Q5 DNA polymerase (NEB) and a template plasmid. The PCR reactants were assembled and the program run as described (2.1.2) using a 5 minute extension time. The template DNA was removed via digestion with DpnI (NEB) at 37°C for 1 hour. The primers used are listed in Table 2.4.

2.2.2. Plasmid construction via digestion and ligation

Some of the plasmids utilised during this study were constructed by the integration of DNA fragments generated via PCR into other plasmids to add or replace genes. The vector plasmid and insert fragment were digested with restriction enzymes (2.1.6.) and ligated (2.1.7) as described previously.

2.3. Maintenance and growth of strains

Nutrient agar (NA) (Oxoid) was utilised for the routine maintenance of bacterial strains on solid medium. For maintenance in liquid medium Luria-Bertani (LB) (1% peptone, 0.5% yeast extract, 1% NaCl) was used. For experiments in *B. subtilis* cells were grown in LB or in Minimal Media (Spizizen Minimal Medium (SMM) (0.2% ammonium sulphate, 1.4% dipotassium phosphate, 0.6% monopotassium phosphate, 0.1% sodium citrate, 0.02% $MgSO_4$, 0.5% glucose) supplemented with 1 µg/ml Fe-NH_4-citrate, 6 mM $MgSO_4$, 0.02 % casamino acids, 0.5% glucose and 0.02 mg/ml tryptophan).

Selection antibiotics and other supplements were used at the concentrations shown in Table 2.1. All chemicals and reagents were obtained from either Sigma-Aldrich or Oxoid.

Antibiotic/Chemical	Concentration
E. coli	
Ampicillin	100 µg/ml
Kanamycin	5 µg/ml
Spectinomycin	50 µg/ml
X-gal	80 µg/ml
B. subtilis	
Chloramphenicol	5 µg/ml
IPTG	0.1mM
Kanamycin	2 µg/ml
Spectinomycin	50 µg/ml
Xylose	1%
Zeocin	10 µg/ml

Table 2.1. Antibiotic and chemical supplement concentrations utilised during this study.

2.4. Competent cells and bacterial transformation

2.4.1. Producing chemically competent E. coli

Chemically competent *E. coli* were produced from a primary culture grown overnight at 37°C in LB, diluted 1:100 in 200 ml LB and incubated at 30°C until absorbance at 600 nm was 0.3. The culture was rapidly cooled in an ice water bath for 5 minutes, and then centrifuged at 3273 x*g* for 10 minutes at 4°C. After removal of the supernatant the pellet was resuspended in 32 ml of ice cold CCMB80 buffer (10 mM KOAc [pH 7], 80 mM $CaCl_2$, 20 mM $MnSO_4$, 10 mM $MgSO_4$ and 10% glycerol made up to 50 ml with Milli-Q water). Following 20 minutes of incubation on ice the cells were centrifuged again at 3273 x*g* for 10 minutes at 4°C and resuspended in 4 ml CCMB80. The cells were dispensed into appropriate aliquots, flash frozen in liquid nitrogen and stored at -80°C.

2.4.2. Transformation of E. coli

E. coli transformation was performed using chemically competent cells. Plasmid DNA was added to 200 µl of competent cell suspension, incubated at 37°C for 30 minutes, heat shocked at 42°C for 90 seconds, incubated on ice for 2 minutes and then finally incubated at 37°C for up to 2 hours in 1 ml LB. Transformed cells were pelleted, resuspended in 200 µl LB and plated on NA plates supplemented with antibiotics as required.

2.4.3. Transformation of B. subtilis

B. subtilis transformation was carried out using a primary culture grown overnight while shaking at 37°C in 2 ml of Minimal Media. The primary culture was diluted 1:16.7 in 3 ml fresh Minimal Media and grown for 3 hours with shaking at 37°C. 3 ml of pre-warmed Starvation Media (SMM supplemented with glucose (0.5%) and $MgSO_4$ (6 mM)) was added and incubated for a further 2 hours with shaking at 37°C. 300 µl of competent cell culture was incubated with 0.2 µl plasmid DNA and incubated with shaking at 37°C for 1 hour. 20 µl of transformed cells were then plated on NA plates supplemented with antibiotics as required and incubated at 37°C for 48 hours.

2.5. Strain construction

2.5.1. Generating B. subtilis amino acid substitution mutant strains

B. subtilis strains carrying amino acid substitution mutations were created using site directed mutagenesis (2.2.1) and then subcloning (via double enzyme digest (2.1.6) and ligation (2.1.7)) of the gene into a clean parental plasmid backbone. This new plasmid was transformed into competent cells (2.4.3) and the mutant gene and antibiotic selection marker integrated into the chromosome via homologous recombination. DNA sequencing (2.1.8) was performed after each step to confirm the correct genotype. A simplified overview of this process is shown in Figure 2.1. The plasmids constructed are listed in Table 2.3.

2.5.2. Constructing other B. subtilis strains

The *B. subtilis* strains constructed during this study that are not amino acid substitution mutants were created by transforming (2.4.3) constructed plasmids or genomic DNA into either wild type (168CA) or other desired strains. Constructed plasmids are listed in Table 2.3.

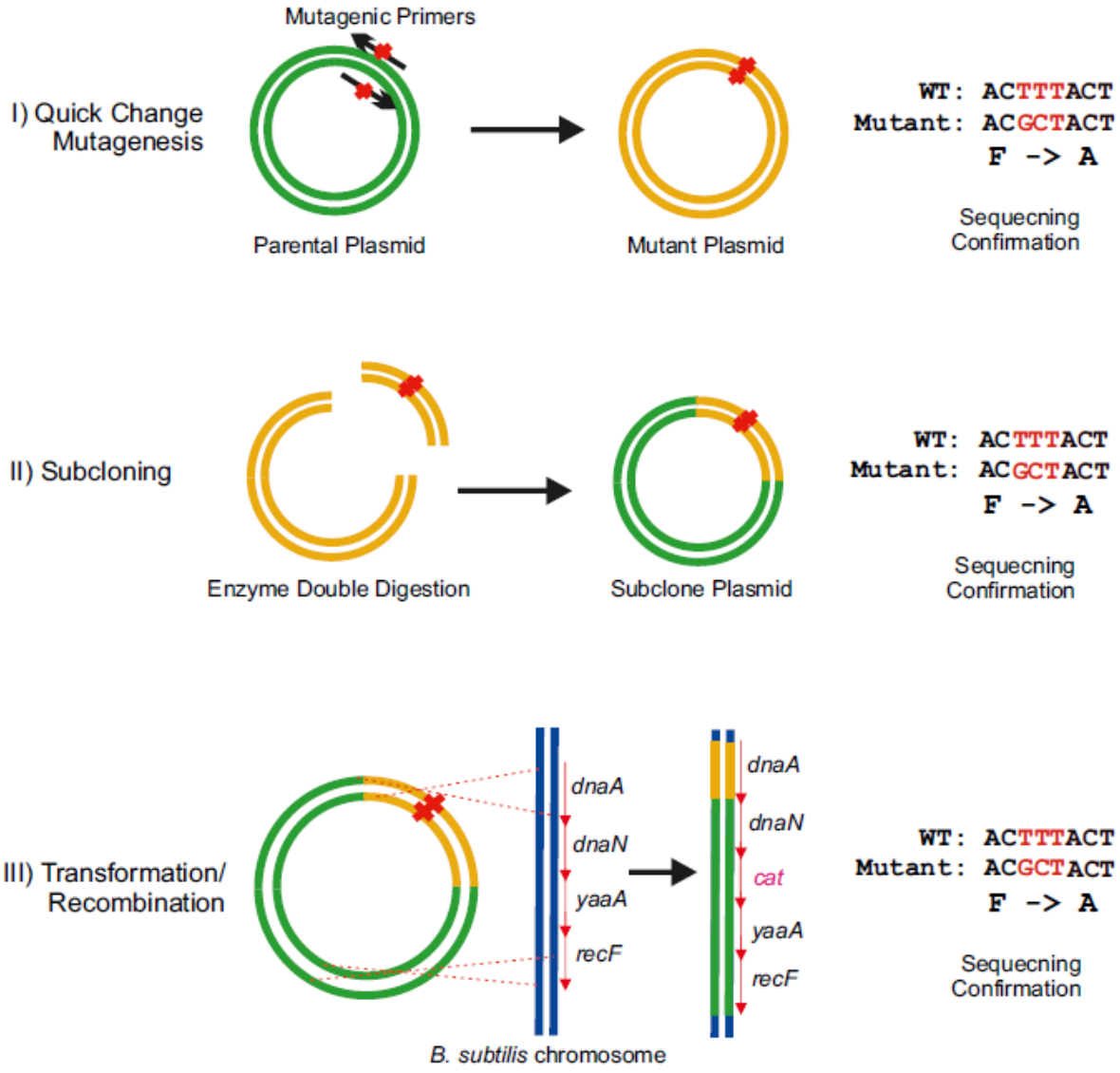


Figure 2.1. Generation of an amino acid substitution mutant *B. subtilis* strain. Overview for how amino acid substitution mutants were constructed in *B. subtilis in vivo*. In this example a native phenylalanine (F) residue within *dnaA* is substituted for alanine (A).

2.6. Phenotype analysis and growth assays

2.6.1. Spot titre assay

Strains were grown overnight at 37°C in Minimal Media containing appropriate supplements. 200 µl of the culture was serially diluted and 10 µl of each dilution point was spotted onto nutrient agar plates supplemented as required. The plate was then incubated at 37°C for 72 hours unless otherwise indicated.

2.6.2. Growth curve analysis

Strains were grown overnight at 37°C in Minimal Media containing appropriate supplements. Strains were then diluted 1:200 in 150 µl of LB in a 96 well plate (Falcon #353072) and incubated for ~24hrs while shaking (600rpm) at 37°C in a Tecan Sunrise plate reader. The absorbance at 600 nm was measured every 5 minutes.

2.7. Microscopy

To visualise cell membrane integrity starter cultures were grown overnight in Minimal Media supplemented appropriately then diluted 1:100 into LB similarly supplemented. Following 4 hours incubation at 37°C cells were stained with SYTOX Green in a 1:1 volume for 15 minutes. Cells were mounted on ~1.5% agarose pads and a 0.13- to 0.17-mm glass coverslip (VWR) was placed on top. Microscopy was performed on an inverted epifluorescence microscope (Nikon Ti) fitted with a phase contrast objective (Nikon Plan Apochomat DM 100x/1.40 Oil Ph3). Light was transmitted from a 300 Watt xenon arc lamp through a liquid light guide (Sutter Instruments), and images were collected using a Prime sCMOS camera (Photometrics). The GFP filter set was from Chroma: ET470/40x (EM), T495Ipxr (BS) and ET525/50m (EM). Digital images were acquired using METAMORPH software (version 7.7).

2.8. Western blot analysis

Primary cultures were grown at 37°C overnight in Minimal Media then diluted 1:100 into LB and grown to an A_{600} of 0.6. 1 ml of culture was centrifuged (9600 x*g*) to remove the supernatant and flash frozen in liquid nitrogen. The pellets were re-suspended in sonication buffer (Phosphate Buffer Saline (PBS) supplemented with 1 mM EDTA and one Roche mini-complete tablet) and sonicated twice (12 seconds at the lowest setting). 4X Laemmli Buffer and 10X Reducing agent (Invitrogen) were added and heated for 10 minutes at 80°C, before again centrifuging. 10 µl of sample was loaded

into a NuPAGE 4-12% Bis-Tris gradient gel (Invitrogen) in 1X MES buffer and separated by electrophoresis at 200 V for 30 minutes. The proteins were transferred to a Hybond-P PVDF membrane with transfer buffer (0.5X MES and 20% methanol) using a semi-dry apparatus. The membrane was then soaked in blocking buffer (5% semi-skimmed milk in PBSTween (PBST)) overnight at 4°C. The membrane was rinsed twice with PBST and washed three times (5 minutes soaking in PBST) and incubated with desired polyclonal antibodies (1:5000) using 5 ml blocking buffer in 20 ml PBST for 2 hours at room temperature. The membrane was rinsed and washed again and incubated for 1 hour with anti-rabbit horseradish peroxidase-linked secondary antibody (1:5000) using 5 ml blocking buffer in 20 ml PBST. The membrane was then rinsed and washed for a final time, before development in ECL solution (ThermoFisher scientific) and protein expression was detected via chemiluminescence of the membrane using an ImageQuant LAS 4000 mini digital imaging system (GE Healthcare).

2.9. Bacterial two-hybrid assay

10 µl of competent *E. coli* strain HM1784 was transformed with a combination of complementary plasmids (0.2 µl of each) and incubated on ice for 1 hour. The cells were then heat shocked at 42°C for 90 seconds, incubated on ice for 5 minutes and then used to inoculate 3 ml of LB supplemented with ampicillin and spectinomycin. These cultures were then incubated at 37°C to an A_{600} of 0.5. The cells were diluted 1:1000 in LB before 5 µl was spotted onto nutrient agar plates containing ampicillin, spectinomycin, and the indicator X-gal (80 µg/ml). Plates were incubated at 30°C for 48 hours and imaged using a digital camera.

2.10. DnaA protein purification

2.10.1. Protein expression

Wild-type and variant $DnaA^{CC}$ purifications began with a primary culture of BL21 *E. coli* cells transformed with the expression vector grown at 37°C overnight in 15 ml LB supplemented with 100 µg/ml ampicillin. This was added to 400 ml LB supplemented with 100 µg/ml ampicillin and grown to an A_{600} of 0.4, whereupon 1 mM of IPTG was added to induce expression and expressed at 30°C for 4 hours. Cells were harvested via centrifugation at 4754 x*g* for 10 minutes, re-suspended in 10 ml of Ni-binding buffer A (25 mM HEPES-KOH pH7.6, 250 mM potassium glutamate, 10 mM magnesium

acetate, 20% sucrose, 30 mM imidazole) and one Roche mini-complete tablet then stored at -80°C.

2.10.2. Protein purification

Frozen cell suspensions were thawed and disrupted by sonication (40W, 2 second pulse, 5 minutes) and pelleted at 69673 x*g* for 30 minutes before the supernatant was filtered through a sterile 0.45 µM Millipore filter. All subsequent steps were performed at 4°C unless stated otherwise. The lysate was applied at 1 ml/minute to a 1 ml HisTrap column (GE) equilibrated with Ni-binding buffer A (see above). The loaded column was washed with 10 ml of DnaA high salt wash buffer (25 mM HEPES-KOH pH7.6, 1 M potassium glutamate, 10 mM magnesium acetate, 20% sucrose, 30mM imidazole). This was followed by washing with 10 ml 1 step gradient of 10% DnaA Ni-elution buffer B (25 mM HEPES-KOH pH7.6, 250 mM potassium glutamate, 10 mM magnesium acetate, 20% sucrose, 500 mM imidazole) and then proteins were eluted using a 10 ml 1-step gradient of 100% of the same buffer. The eluted fractions were then applied at 1 ml/minute to a Heparin HP column (GE) equilibrated with DnaA heparin binding buffer A (25 mM HEPES-KOH pH7.6, 100 mM potassium glutamate, 10 mM magnesium acetate, 20% sucrose). Specifically-bound proteins were eluted using a 10 ml 1-step gradient of 100% DnaA heparin elution buffer B (25 mM HEPES-KOH pH7.6, 1 M potassium glutamate, 10 mM magnesium acetate, 20% sucrose).

The collected fractions were digested overnight on ice with His-Sumo Protease. Subsequently the reaction was loaded onto a 1 ml HisTrap column equilibrated with DnaA Ni-binding buffer A and unbound proteins eluted using DnaA Ni-binding buffer B. Eluted protein was concentrated using an Amicon centrifugal filter (Merck) at 3273 x*g* for 45 minutes. 20% PEG300 was added prior to aliquoting and storing at -80°C.

2.11. DnaA *in vitro* biochemical assays

2.11.1. Filament assembly in solution

Protein filament formation in solution was promoted by adding 3 µM of $DnaA^{CC}$ proteins (adjusted to 10 µl total volume with Milli-Q water) to 20 µl of oligomerisation buffer (25 mM HEPES-KOH pH7.6, 100 mM potassium glutamate, 50 mM NaCl, 10 mM $MgCl_2$) with 2 mM of nucleotide (ADP or ATP). Reactions were incubated at 37°C for 5 minutes before addition of 4 mM BMOE (ThermoFisher Scientific) followed by a further 6 minute incubation at 37°C. Reactions were then quenched with 60 mM cysteine and incubated

for 10 minutes at 37°C before fixing in NuPAGE loading dye at 98°C for 5 minutes. Complexes were resolved by running 10 µl of cross-linked DnaA from each reaction on a NuPAGE Novex 3-8% Tris-acetate gel (Invitrogen) then transferring to a Hybond 0.2 µm PVDF membrane (Amersham). Transfer was performed for 1 hour with TurboBlot followed by visualisation via western blotting with polyclonal anti-DnaA antibody (Section 2.8) (Scholefield *et al.*, 2012).

2.11.2. Filament assembly on DNA scaffolds

DNA scaffolds were prepared by combining oligonucleotides listed in Table 2.5 (final concentration 50 nM) in annealing buffer (10 mM HEPES-KOH pH7.6, 100 mM NaCl and 1 mM EDTA pH8.0) and heating to 98°C for five minutes, followed by gradual cooling to ~40°C. Filament formation was promoted by mixing $DnaA^{CC}$ proteins (200 nM final concentration) with the DNA scaffold (15 nM), total volume 10 µl in oligomerisation buffer (30 mM HEPES-KOH pH7.0, 100 mM potassium glutamate, 100 mM NaCl, 10 mM magnesium acetate, 25% glycerol, 0.01% Tween-20) and 2 mM of nucleotide (ADP or ATP). Reactions were incubated at 37°C for 20 minutes before addition of 4 mM BMOE (ThermoFisher Scientific) followed by a further 6 minute incubation at 37°C. Reactions were then quenched with 60 mM cysteine and incubated for 10 minutes at 37°C before fixing in NuPAGE loading dye at 98°C for 5 minutes. Complexes were resolved by running 10 µl of cross-linked DnaA from each reaction on a NuPAGE Novex 3-8% Tris-acetate gel (Invitrogen) then transferring to a Hybond 0.2 µm PVDF membrane (Amersham). Transfer was performed for 1 hour with TurboBlot followed by visualisation via western blotting with polyclonal anti-DnaA antibody (Section 2.8) (Richardson *et al.*, 2016).

2.11.3. DNA strand seperation

DNA scaffolds were prepared by combining oligonucleotides listed in Table 2.6 (final concentration 1 mM) in annealing buffer (30 mM HEPES-KOH pH 8, 100 mM potassium acetate and 5 mM magnesium acetate). Protein stocks (final concentration 780 nM) were prepared by combining appropriate volume of purified protein with 6 mM of nucleotide (ADP/ATP) in strand seperation buffer (10 mM HEPES-KOH pH 7.5, 100 mM potassium glutamate and 1 mM magnesium acetate). Final concentrations of 13 nM of DNA scaffold and 130 nM of protein were prepared in strand seperation buffer with 30% glycerol, 10% DMSO and Milli-Q water for a final volume of 30 µl. Protein/DNA mixes were loaded into a flat-bottom black polystyrene 96 well plate

(Costar #266) and shaken for 10 seconds at 200rpm. Fluorescence of 610-630 nm was measured every minute for 90 minutes at 25°C using a BMG Clariostar plate reader.

2.12. Strains, plasmids and oligonucleotides

Tables 2.2-2.6 list all strains, plasmids and oligonucleotides used in, or constructed during, the course of this study. Oligonucleotides used for quick-change mutagenesis show the mutant codon(s) underlined.

Table 2.2. List of Strains

Strain	Construction (Reference)	Genotype
E. coli		
DH5α		F– Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17 (rK–, mK+) *pho*A *sup*E44 λ– *thi*-1 *gyr*A96 *rel*A1
BL21		*fhuA2 [lon] ompT gal [dcm] ΔhsdS*
BTH101		F-, *cya-99*, *araD139, galE15, galK16, rpsL1 (Str r)*, *hsdR2, mcrA1, mcrB1.*
HM1784	H. Murray (Unpublished)	*BTH101* Δ*rnh*::*kan*
HM1792	H. Murray (Unpublished)	*DH5α* Δ*rnh*::*kan*
B. subtilis		
168CA	(Kunst *et al.*, 1997)	*trpC2*
CW162	C. Winterhalter (Unpublished)	*trpC2 amyE::spec*(P_{HSA+1T}-*dnaD-ssrA-lacI*$^{Q18M/W220F}$)
CW164	C. Winterhalter (Unpublished)	*trpC2 amyE::spec*(P_{HSA+1T}-*dnaD-ssrA-lacI*$^{Q18M/W220F}$) *cat::Δ(dnaD-nth)*
CW166	C. Winterhalter (Unpublished)	*trpC2 amyE::spec*(P_{HSA+1T}-*dnaD-ssrA-lacI*$^{Q18M/W220F}$) *dnaD*^{E95A}*::kan*
CW170	C. Winterhalter (Unpublished)	*trpC2 amyE::spec*(P_{HSA+1T}-*dnaD-ssrA-lacI*$^{Q18M/W220F}$) *dnaD*^{I83A}*::kan*
CW174	C. Winterhalter (Unpublished)	*trpC2 amyE::spec*(P_{HSA+1T}-*dnaD-ssrA-lacI*$^{Q18M/W220F}$) *dnaD*^{F51A}*::kan*
DS3	pDS10 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI* P_{spac}*-repN/oriN) dnaA* ^{N203A} *:: cat*
DS4	pDS14 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI* P_{spac}*-repN/oriN) dnaA* ^{R204A} *:: cat*
DS5	pDS2 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI* P_{spac}*-repN/oriN) dnaA* ^{S192A}*:: cat*
DS6	pDS22 transformed into HM1108 *(Richardson et al., 2019)*	*trpC2 aprE::kan(lacI* P_{spac}*-repN/oriN) dnaA* ^{E183A}*:: cat*
DS7	pDS23 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI* P_{spac}*-repN/oriN) dnaA* ^{F185A}*:: cat*
DS8	pDS24 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI* P_{spac}*-repN/oriN) dnaA* ^{T186A}*:: cat*
DS9	pDS25 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI* P_{spac}*-repN/oriN) dnaA* ^{F201A}*:: cat*
DS10	pDS26 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI* P_{spac}*-repN/oriN) dnaA* ^{D200A}*:: cat*

DS11	pDS30 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{A198G}:: cat*
DS12	pDS36 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA^{R194AD195A}:: cat*
DS13	pDS20 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{N207A}:: cat*
DS14	pDS25 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{V199A}:: cat*
DS15	pDS48 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{E188A}:: cat*
DS16	pDS44 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{K197A}:: cat*
DS17	pDS37 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA^{N196AK197A}:: cat*
DS18	pDS50 transformed into HM1108 *(Richardson et al., 2019)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{I190A}:: cat*
DS19	pDS49 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{F189A}:: cat*
DS20	pDS51 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{I193A}:: cat*
DS21	pDS45 transformed into HM1108 *(Richardson et al., 2019)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{R202A}:: cat*
DS22	pDS4 transformed into HM1108 *(Richardson et al., 2019)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{R206A}:: cat*
DS23	pDS47 transformed into HM1108 *(Richardson et al., 2019)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{N187A}:: cat*
DS24	pDS19 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{Y205A}:: cat*
DS25	pDS55 transformed into HM1108 *(Richardson et al., 2019)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{F128A}:: cat*
DS26	pDS56 transformed into HM1108 *(Richardson et al., 2019)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{F218A}:: cat*
DS27	pDS57 transformed into HM1108 *(Richardson et al., 2019)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{R321A}:: cat*
DS28	pDS41 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{R194A}:: cat*
DS29	pDS43 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{N196A}:: cat*
DS30	pDS42 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{D195A}:: cat*
DS31	pDS61 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{F49A}:: cat*
DS32	pDS62 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{D52A}:: cat*
DS33	pDS63 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{W53A}:: cat*
DS34	pDS64 transformed into HM1108 *(Richardson et al., 2019)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{L269A}:: cat*
DS35	pDS66 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{E48A}:: cat*
DS36	pDS46 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{K184A}:: cat*
DS37	pDS40 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{N191A}:: cat*

DS38	pDS38 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA^{R194AK197A}:: cat*
DS39	pDS39 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA^{D195AN196A}:: cat*
DS40	pDS21 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{S182A}:: cat*
DS41	pDS88 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{K252A}:: cat*
DS42	pDS98 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{S401A}:: cat*
DS43	pDS100 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{K404A}:: cat*
DS44	pDS90 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{D259A}:: cat*
DS45	pDS99 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{L402A}:: cat*
DS47	pDS96 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{E253A}:: cat*
DS48	pDS108 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{E253K}:: cat*
DS49	pDS109 transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{D259K}:: cat*
DS50	pDS121 transformed into HM1108 *(Richardson et al., 2019)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{H231A}:: cat*
DS51	pDS122 transformed into HM1108 *(Richardson et al., 2019)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA^{G317Q}:: cat*
DS52	pDS118 transformed into HM1108 *(Richardson et al., 2019)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{T225A}:: cat*
DS53	pDS116 transformed into HM1108 *(Richardson et al., 2019)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{K222A}:: cat*
DS54	pDS117 transformed into HM1108 *(Richardson et al., 2019)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{Q224A}:: cat*
DS56	pDS123 transformed into HM1108 *(Richardson et al., 2019)*	*trpC2 aprE::kan(lacI P_{spac}-repN/oriN) dnaA ^{R264A}:: cat*
DS57	TR244 Transformed into HM1423 *(Richardson et al., 2019)*	*trpC2 spoIIIJ*(359°)::(*oriN kan tet*) Δ*dnaA*::*zeo amyE*::*spc*(*xylR* P_{xyl}-*dnaAchi^{I190A}*)
DS58	TR262 Transformed into HM1423 *(Richardson et al., 2019)*	*trpC2 spoIIIJ*(359°)::(*oriN kan tet*) Δ*dnaA*::*zeo amyE*::*spc*(*xylR* P_{xyl}-*dnaAchi^{K222A}*)
DS59	TR265 Transformed into HM1423 *(Richardson et al., 2019)*	*trpC2 spoIIIJ*(359°)::(*oriN kan tet*) Δ*dnaA*::*zeo amyE*::*spc*(*xylR* P_{xyl}-*dnaAchi^{T225A}*)
DS60	TR313 Transformed into HM1423 *(Richardson et al., 2019)*	*trpC2 spoIIIJ*(359°)::(*oriN kan tet*) Δ*dnaA*::*zeo amyE*::*spc*(*xylR* P_{xyl}-*dnaAchi^{R264A}*)
DS61	TR480 Transformed into HM1423 *(Richardson et al., 2019)*	*trpC2 spoIIIJ*(359°)::(*oriN kan tet*) Δ*dnaA*::*zeo amyE*::*spc*(*xylR* P_{xyl}-*dnaAchi^{R202A}*)
DS62	TR481 Transformed into HM1423 *(Richardson et al., 2019)*	*trpC2 spoIIIJ*(359°)::(*oriN kan tet*) Δ*dnaA*::*zeo amyE*::*spc*(*xylR* P_{xyl}-*dnaAchi^{R206A}*)
DS64	TR483 Transformed into HM1423 *(Richardson et al., 2019)*	*trpC2 spoIIIJ*(359°)::(*oriN kan tet*) Δ*dnaA*::*zeo amyE*::*spc*(*xylR* P_{xyl}-*dnaAchi^{L269A}*)
DS65	TR486 Transformed into HM1423 *(Richardson et al., 2019)*	*trpC2 spoIIIJ*(359°)::(*oriN kan tet*) Δ*dnaA*::*zeo amyE*::*spc*(*xylR* P_{xyl}-*dnaAchi^{F218A}*)
DS66	TR488 Transformed into HM1423 *(Richardson et al., 2019)*	*trpC2 spoIIIJ*(359°)::(*oriN kan tet*) Δ*dnaA*::*zeo amyE*::*spc*(*xylR* P_{xyl}-*dnaAchi^{R321A}*)
DS67	pDS124 Transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI Pspac-repN/oriN) DnaA ^{E183AN187A}:: cat*

DS68	TR241 Transformed into HM1423 *(Richardson et al., 2019)*	*trpC2 spoIIIJ*(359°)::(*oriN kan tet*) Δ*dnaA*::*zeo amyE*::*spc*(*xylR* P_{xyl}-*dnaAchi*)
DS69	pDS110 Transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI* P_{spac}*-repN/oriN) dnaA* ^{K404E}*:: cat*
DS70	pDS111 Transformed into HM1108 *(This work)*	*trpC2 aprE::kan(lacI* P_{spac}*-repN/oriN) dnaA* ^{L402D}*:: cat*
HM1108	(Richardson *et al.*, 2016)	*trpC2 aprE*::*kan*(*lacI* P_{spac}-*repN/oriN*)
HM1424	(Murray and Koh, 2014)	*trpC2 spoIIIJ(*359°*)*::*(oriN kan tet) ΔdnaA::zeo*
HM1540	H. Murray (Unpublished)	*trpC2 aprE::kan(lacI* P_{spac}*-repN/oriN) dnaA* ^{T26A}*:: cat*
HM1541	H. Murray (Unpublished)	*trpC2 aprE::kan(lacI* P_{spac}*-repN/oriN) dnaA* ^{W27A}*:: cat*
HM1603	(Richardson *et al.*, 2016)	*trpC2 aprE::kan(lacI* P_{spac}*-repN/oriN) (dnaA-ΔincC-dnaN)::cat*
HM1683	(Richardson *et al.*, 2019)	*trpC2 aprE::kan(lacI* P_{spac}*-repN/oriN) amyE::spc(xylR Pxyl-dnaAchi)*
HM1694	(Richardson *et al.*, 2019)	*trpC2 aprE::kan(lacI* P_{spac}*-repN/oriN) incCart(DnaA-box#6/7)::cat amyE::spc(xylR Pxyl-dnaAchi)*
HM1834	(Richardson *et al.*, 2019)	*trpC2 aprE::kan(lacI* P_{spac}*-repN/oriN) incCart(DnaA-box#Tm45/6/7)::cat*
TR241	(Richardson *et al.*, 2019)	*trpC2 aprE::kan(lacI* P_{spac}*-repN/oriN) incC*^{art}*(DnaA-box#*Tm^{45}*/6/7)::cat amyE::spc(xylR* P_{xyl}*-dnaAchi)*
TR244	(Richardson *et al.*, 2019)	*trpC2 aprE::kan(lacI* P_{spac}*-repN/oriN) incC*^{art}*(DnaA-box#*Tm^{45}*/6/7)::cat amyE::spc(xylR* P_{xyl}*-dnaAchiI*^{190A}*)*
TR262	(Richardson *et al.*, 2019)	*trpC2 aprE::kan(lacI* P_{spac}*-repN/oriN) incC*^{art}*(DnaA-box#*Tm^{45}*/6/7)::cat amyE::spc(xylR* P_{xyl}*-dnaAchi*^{K222A}*)*
TR265	(Richardson *et al.*, 2019)	*trpC2 aprE::kan(lacI* P_{spac}*-repN/oriN) incC*^{art}*(DnaA-box#*Tm^{45}*/6/7)::cat amyE::spc(xylR* P_{xyl}*-dnaAchi*^{T225A}*)*
TR313	(Richardson *et al.*, 2019)	*trpC2 aprE::kan(lacI* P_{spac}*-repN/oriN) incC*^{art}*(DnaA-box#*Tm^{45}*/6/7)::cat amyE::spc(xylR* P_{xyl}*-dnaAchi*^{R264A}*)*
TR320	T. Richardson (Unpublished)	*trpC2 aprE::kan(lacI* P_{spac}*-repN/oriN), rpmH-erm dnaA*$^{\Delta I\text{-}II}$*)*
TR480	(Richardson *et al.*, 2019)	*trpC2 aprE::kan(lacI* P_{spac}*-repN/oriN) incC*^{art}*(DnaA-box#*Tm^{45}*/6/7)::cat amyE::spc(xylR* P_{xyl}*-dnaAchi*^{R202A}*)*
TR481	(Richardson *et al.*, 2019)	*trpC2 aprE::kan(lacI* P_{spac}*-repN/oriN) incC*^{art}*(DnaA-box#*Tm^{45}*/6/7)::cat amyE::spc(xylR* P_{xyl}*-dnaAchi*$^{R206A)}$
TR483	(Richardson *et al.*, 2019)	*trpC2 aprE::kan(lacI* P_{spac}*-repN/oriN) incC*^{art}*(DnaA-box#*$Tm^{45/}$*6/7)::cat amyE::spc(xylR* P_{xyl}*-dnaAchi*^{L269A}*)*
TR486	(Richardson *et al.*, 2019)	*trpC2 aprE::kan(lacI* P_{spac}*-repN/oriN) incC*^{art}*(DnaA-box#*Tm^{45}*/6/7)::cat amyE::spc(xylR* P_{xyl}*-dnaAchi*^{F218A}*)*
TR488	(Richardson *et al.*, 2019)	*trpC2 aprE::kan(lacI* P_{spac}*-repN/oriN) incC*^{art}*(DnaA-box#*Tm^{45}*/6/7)::cat amyE::spc(xylR* P_{xyl}*-dnaAchi*^{R321A}*)*

Table 2.3. List of Plasmids

Plasmid	Construction	Genotype	Reference
pCW213	*sirA* integrated into pHM642	*bla* P_{lacUV5}*-T18-MCS-dnaD sirA*	C. Winterhalter (Unpublished)
pDS1	Quick-change mutagenesis of pHM327 with oDS1/oDS2	*bla 'dnaA dnaN cat recF' dnaA*^{S192A}	This work
pDS2	pDS1 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{S192A}	This work
pDS3	Quick-change mutagenesis of pHM327 with oDS3/oDS4	*bla 'dnaA dnaN cat recF' dnaA*^{R206A}	(Richardson *et al.*, 2019)
pDS4	pDS3 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{R206A}	(Richardson *et al.*, 2019)

pDS5	Quick-change mutagenesis of pHM327 with oDS13/oDS14	*bla 'dnaA dnaN cat recF' dnaA*^{V199A}	This work
pDS6	Quick-change mutagenesis of pHM327 with oDS17/oDS18	*bla 'dnaA dnaN cat recF' dnaA*^{F201A}	This work
pDS8	Quick-change mutagenesis of pHM327 with oDS21/oDS22	*bla 'dnaA dnaN cat recF' dnaA*^{N203A}	This work
pDS9	Quick-change mutagenesis of pHM327 with oDS23/oDS24	*bla 'dnaA dnaN cat recF' dnaA*^{R204A}	This work
pDS10	pDS8 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{N203A}	This work
pDS11	Quick-change mutagenesis of pHM327 with oDS5/oDS6	*bla 'dnaA dnaN cat recF' dnaA*^{S182A}	This work
pDS12	Quick-change mutagenesis of pHM327 with oDS7/oDS8	*bla 'dnaA dnaN cat recF' dnaA*^{E183A}	(Richardson *et al.*, 2019)
pDS13	Quick-change mutagenesis of pHM327 with oDS15/oDS16	*bla 'dnaA dnaN cat recF' dnaA*^{D200A}	This work
pDS14	pDS9 subcloned into pHM327(PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{R204A}	This work
pDS15	Quick-change mutagenesis of pHM327 with oDS9/oDS10	*bla 'dnaA dnaN cat recF' dnaA*^{F185A}	This work
pDS16	Quick-change mutagenesis of pHM327 with oDS11/oDS12	*bla 'dnaA dnaN cat recF' dnaA*^{T186A}	This work
pDS17	Quick-change mutagenesis of pHM327 with oDS25/oDS26	*bla 'dnaA dnaN cat recF' dnaA*^{Y205A}	This work
pDS18	Quick-change mutagenesis of pHM327 with oDS27/oDS28	*bla 'dnaA dnaN cat recF' dnaA*^{N207A}	This work
pDS19	pDS17 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{Y205A}	This work
pDS20	pDS18 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{N207A}	This work
pDS21	pDS11 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{S182A}	This work
pDS22	pDS12 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{E183A}	(Richardson *et al.*, 2019)
pDS23	PDS15 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{F185A}	This work
pDS24	pDS16 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{T186A}	This work
pDS25	pDS5 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{V199A}	This work
pDS26	pDS13 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{D200A}	This work
pDS27	pDS6 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{F201A}	This work
pDS28	Quick-change mutagenesis of pHM327 with oDS29/oDS30	*bla 'dnaA dnaN cat recF' dnaA*^{A198G}	This work
pDS30	pDS28 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{A198G}	This work
pDS32	Quick-change mutagenesis of pHM327 with oDS35/oDS36	*bla 'dnaA dnaN cat recF' dnaA*^{R194AD195A}	This work
pDS33	Quick-change mutagenesis of pHM327 with oDS37/oDS38	*bla 'dnaA dnaN cat recF' dnaA*^{N196AK197A}	This work
pDS34	Quick-change mutagenesis of pHM327 with oDS39/oDS40	*bla 'dnaA dnaN cat recF' dnaA*^{R194AK197A}	This work

pDS35	Quick-change mutagenesis of pHM327 with oDS41/oDS42	*bla 'dnaA dnaN cat recF' dnaA*^{D195AN196A}	This work
pDS36	pDS32 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{R194AD195A}	This work
pDS37	pDS33 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{N196AK197A}	This work
pDS38	pDS34 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{R194AK197A}	This work
pDS39	pDS35 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{D195AN196A}	This work
pDS40	pTS019 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{N191A}	This work
pDS41	pTS016 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{R194A}	This work
pDS42	pTS017 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{D195A}	This work
pDS43	pTS018 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{N196A}	This work
pDS44	pTS014 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{K197A}	This work
pDS45	pTS015 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{R202A}	(Richardson *et al.*, 2019)
pDS46	pTS009 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{K184A}	This work
pDS47	pTS010 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{N187A}	(Richardson *et al.*, 2019)
pDS48	pTS011 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{E188A}	This work
pDS49	pTS012 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{F189A}	This work
pDS50	pTS002 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{I190A}	(Richardson *et al.*, 2019)
pDS51	pTS013 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{I193A}	This work
pDS52	Quick-change mutagenesis of pHM327 with oDS45/oDS46	*bla 'dnaA dnaN cat recF' dnaA*^{F128A}	(Richardson *et al.*, 2019)
pDS53	Quick-change mutagenesis of pHM327 with oDS47/oDS48	*bla 'dnaA dnaN cat recF' dnaA*^{F218A}	(Richardson *et al.*, 2019)
pDS54	Quick-change mutagenesis of pHM327 with oDS51/oDS52	*bla 'dnaA dnaN cat recF' dnaA*^{R321A}	(Richardson *et al.*, 2019)
pDS55	pDS52 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{F128A}	(Richardson *et al.*, 2019)
pDS56	pDS53 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{F218A}	(Richardson *et al.*, 2019)
pDS57	pDS54 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{R321A}	(Richardson *et al.*, 2019)
pDS58	Quick-change mutagenesis of pHM327 with oDS61/oDS62	*bla 'dnaA dnaN cat recF' dnaA*^{D52A}	This work
pDS59	Quick-change mutagenesis of pHM327 with oDS63/oDS64	*bla 'dnaA dnaN cat recF' dnaA*^{W53A}	This work
pDS60	Quick-change mutagenesis of pHM327 with oDS59/oDS60	*bla 'dnaA dnaN cat recF' dnaA*^{E48A}	This work
pDS61	pHM545 subcloned into pHM327 (AatII/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{F49A}	This work

pDS62	pDS58 subcloned into pHM327 (AatII/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{D52A}	This work
pDS63	pDS59 subcloned into pHM327 (AatII/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{W53A}	This work
pDS64	pDS60 subcloned into pHM327 (AatII/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{E48A}	This work
pDS65	Quick-change mutagenesis of pHM327 with oDS49/oDS50	*bla 'dnaA dnaN cat recF' dnaA*^{L269A}	(Richardson *et al.*, 2019)
pDS66	pDS65 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{L269A}	(Richardson *et al.*, 2019)
pDS84	pST25 w/dnaAF49A (PCR oHM665/oHM666) (BamHI-Asp718I)	*spc* P$_{lacUV5}$*-T25-MCS-dnaA*^{F49A}	This work
pDS88	Quick-change mutagenesis of pHM327 with oDS67/oDS68	*bla 'dnaA dnaN cat recF' dnaA*^{K252A}	This work
pDS89	Quick-change mutagenesis of pHM327 with oDS69/oDS70	*bla 'dnaA dnaN cat recF' dnaA*^{E253A}	This work
pDS90	Quick-change mutagenesis of pHM327 with oDS71/oDS72	*bla 'dnaA dnaN cat recF' dnaA*^{D259A}	This work
pDS91	Quick-change mutagenesis of pHM327 with oDS73/oDS74	*bla 'dnaA dnaN cat recF' dnaA*^{S401A}	This work
pDS92	Quick-change mutagenesis of pHM327 with oDS75/oDS76	*bla 'dnaA dnaN cat recF' dnaA*^{L402A}	This work
pDS93	Quick-change mutagenesis of pHM327 with oDS77/oDS78	*bla 'dnaA dnaN cat recF' dnaA*^{K404A}	This work
pDS95	pDS88 subcloned into pHM327 (SalI/FspI)	*bla 'dnaA dnaN cat recF' dnaA*^{K252A}	This work
pDS96	pDS89 subcloned into pHM327 (SalI/FspI)	*bla 'dnaA dnaN cat recF' dnaA*^{E253A}	This work
pDS97	pDS90 subcloned into pHM327 (SalI/FspI)	*bla 'dnaA dnaN cat recF' dnaA*^{D259A}	This work
pDS98	pDS91 subcloned into pHM327 (SalI/FspI)	*bla 'dnaA dnaN cat recF' dnaA*^{S401A}	This work
pDS99	pDS92 subcloned into pHM327 (SalI/FspI)	*bla 'dnaA dnaN cat recF' dnaA*^{L402A}	This work
pDS100	pDS93 subcloned into pHM327 (SalI/FspI)	*bla 'dnaA dnaN cat recF' dnaA*^{K404A}	This work
pDS102	Quick-change mutagenesis of pHM327 with oDS87/oDS88	*bla 'dnaA dnaN cat recF' dnaA*^{K252E}	This work
pDS103	Quick-change mutagenesis of pHM327 with oDS59/oDS96	*bla 'dnaA dnaN cat recF' dnaA*^{S401D}	This work
pDS104	Quick-change mutagenesis of pHM327 with oDS97/oDS98	*bla 'dnaA dnaN cat recF' dnaA*^{L402D}	This work
pDS105	Quick-change mutagenesis of pHM327 with oDS99/oDS100	*bla 'dnaA dnaN cat recF' dnaA*^{K404E}	This work
pDS106	Quick-change mutagenesis of pHM327 with oDS91/oDS92	*bla 'dnaA dnaN cat recF' dnaA*^{E253K}	This work
pDS107	Quick-change mutagenesis of pHM327 with oDS93/oDS94	*bla 'dnaA dnaN cat recF' dnaA*^{D259K}	This work
pDS108	pDS106 subcloned into pHM327 (SalI/FspI)	*bla 'dnaA dnaN cat recF' dnaA*^{E253K}	This work
pDS109	pDS107 subcloned into pHM327 (SalI/FspI)	*bla 'dnaA dnaN cat recF' dnaA*^{D259K}	This work
pDS110	pDS105 subcloned into pHM327 (SalI/FspI)	*bla 'dnaA dnaN cat recF' dnaA*^{K404E}	This work

pDS111	pDS104 subcloned into pHM327 (SalI/FspI)	*bla 'dnaA dnaN cat recF' dnaA*^{L402D}	This work
pDS112	pDS102 subcloned into pHM327 (SalI/FspI)	*bla 'dnaA dnaN cat recF' dnaA*^{K252E}	This work
pDS113	pDS103 subcloned into pHM327 (SalI/FspI)	*bla 'dnaA dnaN cat recF' dnaA*^{S401D}	This work
pDS116	pTS004 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{K222A}	(Richardson *et al.*, 2019)
pDS117	pTS005 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{Q224A}	(Richardson *et al.*, 2019)
pDS118	pTS006 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{T225A}	(Richardson *et al.*, 2019)
pDS119	pST25 w/dnaAT26A (PCR oHM665/oHM666) (BamHI-Asp718I)	*spc* P_{lacUV5}*-T25-MCS-dnaA*^{T26A}	This work
pDS120	pST25 w/dnaAW27A (PCR oHM665/oHM666) (BamHI-Asp718I)	*spc* P_{lacUV5}*-T25-MCS-dnaA*^{W27A}	This work
pDS121	pUT18C w/dnaDNTD (PCR oDS131/oDS132) (BamHI-Asp718I)	*bla* P_{lacUV5}*-T18-MCS-dnaD*^{NTD}	This work
pDS122	Quick-change mutagenesis of pHM327 with oDS138/oDS139	*bla 'dnaA dnaN cat recF' dnaA*$^{E183A/N187A}$	This work
pDS123	pTS001 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*^{R264A}	(Richardson *et al.*, 2019)
pDS124	pDS122 subcloned into pHM327 (PflmI/PacI)	*bla 'dnaA dnaN cat recF' dnaA*$^{E183A/N187A}$	This work
pDS125	pUT18C w/dnaDCTD (PCR oDS133/oDS134) (BamHI-Asp718I)	*bla* P_{lacUV5}*-T18-MCS-dnaD*^{CTD}	This work
pDS126	pST25 w/dnaDF51A (PCR oHM667/oHM668) (BamHI-Asp718I)	*spc* P_{lacUV5}*-T25-MCS-dnaD*^{F51A}	This work
pDS127	pST25 w/dnaDI83A (PCR oHM667/oHM668) (BamHI-Asp718I)	*spc* P_{lacUV5}*-T25-MCS-dnaD*^{I83A}	This work
pDS128	pTR74 subcloned into pSF014 (HindIII/PacI)	*bla14XHis-SUMO-dnaA*^{N191CA198C}	This work
pDS129	pTR168 subcloned into pDS128 (HindIII/PacI)	*bla14XHis-SUMO-dnaA*^{I190A}^{N191CA198C}	This work
pDS130	Quick-change mutagenesis of pDS128 with oDS162/oDS163	*bla14XHis-SUMO-dnaA*^{A}^{N191C} *I193*^{A198C}	This work
pDS131	Quick-change mutagenesis of pDS128 with oDS158/oDS159	*bla14XHis-SUMO-dnaA*^{I189A}^{N191CA198C}	This work
pDS132	pST25 w/dnaDE95A (PCR oHM667/oHM668) (BamHI-Asp718I)	*spc* P_{lacUV5}*-T25-MCS-dnaD*^{E95A}	This work
pDS137	Quick-change mutagenesis of pHM640 with oDS63/oDS64	*spc* P_{lacUV5}*-T25-MCS-dnaA*^{W53A}	This work
pHM327		*bla 'dnaA dnaN cat recF*	(Scholefield *et al.*, 2012)
pHM359	*sirA* integrated into pUT18C	*bla* P_{lacUV5}*-T18-MCS-sirA*	H. Murray (Unpublished)
pHM545	Quick-change mutagenesis of pHM327	*bla 'dnaA dnaN cat recF' dnaA*^{F49A}	H. Murray (Unpublished)
pHM638	*dnaA* integrated into pUT18C	*bla* P_{lacUV5}*-T18-MCS-dnaA*	H. Murray (Unpublished)
pHM640	*dnaA* integrated into pST25	*spc* P_{lacUV5}*-T25-MCS-dnaA*	H. Murray (Unpublished)
pHM642	*dnaD* integrated into pUT18C	*bla* P_{lacUV5}*-T18-MCS-dnaD*	H. Murray (Unpublished)

pHM644	*dnaD* integrated into pST25	*spc* P_{lacUV5}-*T25-MCS-dnaD*	H. Murray (Unpublished)
pHM646	*dnaA* domains II-IV integrated into pUT18C	*bla* P_{lacUV5}-*T18-MCS-dnaAII-IV*	H. Murray (Unpublished)
pHM648	*dnaA* domains II-IV integrated into pST25	*spc* P_{lacUV5}-*T25-MCS-dnaAII-IV*	H. Murray (Unpublished)
pHM650	*dnaB* integrated into pUT18C	*bla* P_{lacUV5}-*T18-MCS-dnaB*	H. Murray (Unpublished)
pHM652	*dnaB* integrated into pST25	*spc* P_{lacUV5}-*T25-MCS-dnaB*	H. Murray (Unpublished)
pHM654	*dnaI* integrated into pUT18C	*bla* P_{lacUV5}-*T18-MCS-dnaI*	H. Murray (Unpublished)
pHM656	*dnaI* integrated into pST25	*spc* P_{lacUV5}-*T25-MCS-dnaI*	H. Murray (Unpublished)
pSF014		*bla14XHis-SUMO-dnaA*	S. Fenyk (Unpublished)
pST25		*spc* P_{lacUV5}-*T25-MCS*	(Ouellette *et al.*, 2014)
pTS001	Quick-change mutagenesis of pHM327 with oAK369/oAK370	*bla 'dnaA dnaN cat recF' dnaA*^{R264A}	T. Sperlea (Richardson *et al.*, 2019)
pTS002	Quick-change mutagenesis of pHM327 with oAK361/oAK362	*bla 'dnaA dnaN cat recF' dnaA*^{I190A}	T. Sperlea (Richardson *et al.*, 2019)
pTS004	Quick-change mutagenesis of pHM327 with oAK365/oAK366	*bla 'dnaA dnaN cat recF' dnaA*^{K222A}	T. Sperlea (Richardson *et al.*, 2019)
pTS005	Quick-change mutagenesis of pHM327 with oAK367/oAK368	*bla 'dnaA dnaN cat recF' dnaA*^{Q224A}	T. Sperlea (Richardson *et al.*, 2019)
pTS006	Quick-change mutagenesis of pHM327 with oAK373/oAK374	*bla 'dnaA dnaN cat recF' dnaA*^{T225A}	T. Sperlea (Richardson *et al.*, 2019)
pTS009	Quick-change mutagenesis of pHM327 with oTS001/oTS003	*bla 'dnaA dnaN cat recF' dnaA*^{K184A}	T. Sperlea (unpuplished)
pTS010	Quick-change mutagenesis of pHM327 with oTS004/oTS005	*bla 'dnaA dnaN cat recF' dnaA*^{N187A}	T. Sperlea (Richardson *et al.*, 2019)
pTS011	Quick-change mutagenesis of pHM327 with oTS006/oTS007	*bla 'dnaA dnaN cat recF' dnaA*^{E188A}	T. Sperlea (unpuplished)
pTS012	Quick-change mutagenesis of pHM327 with oTS008/oTS009	*bla 'dnaA dnaN cat recF' dnaA*^{F189A}	T. Sperlea (unpuplished)
pTS013	Quick-change mutagenesis of pHM327 with oTS014/oTS015	*bla 'dnaA dnaN cat recF' dnaA*^{I193A}	T. Sperlea (unpuplished)
pTS014	Quick-change mutagenesis of pHM327 with oTS022/oTS023	*bla 'dnaA dnaN cat recF' dnaA*^{K197A}	T. Sperlea (unpuplished)
pTS015	Quick-change mutagenesis of pHM327 with oTS024/oTS025	*bla 'dnaA dnaN cat recF' dnaA*^{R202A}	T. Sperlea (Richardson *et al.*, 2019)
pTS016	Quick-change mutagenesis of pHM327 with oTS016/oTS017	*bla 'dnaA dnaN cat recF' dnaA*^{R194A}	T. Sperlea (unpuplished)
pTS017	Quick-change mutagenesis of pHM327 with oTS018/oTS019	*bla 'dnaA dnaN cat recF' dnaA*^{D195A}	T. Sperlea (unpuplished)
pTS018	Quick-change mutagenesis of pHM327 with oTS020/oTS021	*bla 'dnaA dnaN cat recF' dnaA*^{N196A}	T. Sperlea (unpuplished)

pTS019	Quick-change mutagenesis of pHM327 with oTS011/oTS012	*bla 'dnaA dnaN cat recF' dnaA*^{N191A}	T. Sperlea (unpuplished)
pTR74		*bla* $_{pT7}$(his6-link-Xa-*dnaA*^{N191CA198C})	(Richardson *et al.*, 2016)
pTR168		*bla* $_{pT7}$(his6-link-Xa-*dnaA*^{I190AN191CA198C})	(Richardson *et al.*, 2016)
pUT18C		*bla* P$_{lacUV5}$-*T18-MCS*	(Karimova *et al.*, 1998)

Table 2.4. Oligonucleotides used for plasmid construction or sequencing

Oligonucleotide	Sequence (5′→3′)
oAK361	GAATTCGCTAACTCTATCCGAGATAATAAAGCCGTC
oAK362	GAGTTAGCGAATTCGTTTGTAAATTTCTCAGAAGAC
oAK365	CGGGGGCTGAACAAACCCAGGAAGAATTTTTCCATAC
oAK366	GTTCAGCCCCCGCTAAAAATTGAATATCATCTATC
oAK367	AAGAAGCTACCCAGGAAGAATTTTTCCATAC
oAK368	GGGTAGCTTCTTTCCCCGCTAAAAATTGAATATC
oAK369	GCTCAGCTTTTGAATGGGGACTTATTACAGATATC
oAK370	TCAAAAGCTGAGCGCAATCTGTCTTCAAGTGTCGG
oAK373	GAACAAGCTCAGGAAGAATTTTTCCATACATTTAAC
oAK374	CCTGAGCTTGTTCTTTCCCCGCTAAAAATTGAATATC
oDS1	CAACGCTATCCGAGATAATAAAGCCGTC
oDS2	GGATAGCGTTGATGAATTCGTTTGTAAATTTC
oDS3	CTATGCTAATGTTGATGTGCTTTTGATAG
oDS4	CATTAGCATAGCGATTGCGGAAGTCGAC
oDS5	GTCTGCTGAGAAATTTACAAACGAATTC
oDS6	GTAAATTTCTCAGCAGACAGATAAACCACTTTGG
oDS7	CTTCTGCTAAATTTACAAACGAATTCATC
oDS8	GTAAATTTAGCAGAAGACAGATAAACCACTTTG
oDS9	GAAAGCTACAAACGAATTCATCAACTCTATC
oDS10	GTTTGTAGCTTTCTCAGAAGACAGATAAAC
oDS11	GAAATTTGCTAACGAATTCATCAACTCTATC
oDS12	CGTTAGCAAATTTCTCAGAAGACAGATAAAC
oDS13	GCCGCTGACTTCCGCAATCGCTATCGAAATG
oDS14	GAAGTCAGCGGCTTTATTATCTCGGATAG
oDS15	GTCGCTTTCCGCAATCGCTATCGAAATG
oDS16	GGAAAGCGACGGCTTTATTATCTCGGATAG
oDS17	GACGCTCGCAATCGCTATCGAAATGTTG
oDS18	GATTGCGAGCGTCGACGGCTTTATTATCTC
oDS21	CGCGCTCGCTATCGAAATGTTGATGTG
oDS22	GATAGCGAGCGCGGAAGTCGACGGCTTTATTATC
oDS23	CAATGCTTATCGAAATGTTGATGTGCTTTTG
oDS24	GATAAGCATTGCGGAAGTCGACGGCTTTATTATC
oDS25	CGCGCTCGAAATGTTGATGTGCTTTTG
oDS26	CATTTCGAGCGCGATTGCGGAAGTCGACGG
oDS27	CGAGCTGTTGATGTGCTTTTGATAGATG

oDS28	CAACAGCTCGATAGCGATTGCGGAAGTC
oDS29	GATAATAAAGGTGTCGACTTCCGCAATCGCTATC
oDS30	CGACACCTTTATTATCTCGGATAGAGTTG
oDS35	CTATCGCAGCTAATAAAGCCGTCGACTTC
oDS36	CTTTATTAGCTGCGATAGAGTTGATGAATTCGTTTG
oDS37	GATGCAGCTGCCGTCGACTTCCGCAATC
oDS38	GGCAGCTGCATCTCGGATAGAGTTGATGAATTC
oDS39	CTATCGCAGATAATGCAGCCGTCGACTTC
oDS40	GGCTGCATTATCTGCGATAGAGTTGATGAATTCGTTTG
oDS41	CGAGCAGCTAAAGCCGTCGACTTCCGC
oDS42	CTTTAGCTGCTCGGATAGAGTTGATGAATTC
oDS45	CGAGCTGCACATGCTGCTTCCCTCGCAG
oDS46	GTGCAGCTCGGTTTCCAGATCCGATGAC
oDS47	CAAGCTTTAGCGGGGAAAGAACAAACC
oDS48	CTAAAGCTTGAATATCATCTATCAAAAG
oDS49	GGAGCTATTACAGATATCACACCGCCTG
oDS50	GTAATAGCTCCCCATTCAAAACGTGAGC
oDS51	CATTAATCGCTGTTGTCGCTTATTCATCTTTAATTAATAAAG
oDS52	CAACAGCGATTAATGCTCCTTCGAGTTC
oDS59	CAATGCTTTTGCCAGAGACTGGCTGGAG
oDS60	CAAAAGCATTGGGAGCCGTGATTGTTAATG
oDS61	CAGAGCTTGGCTGGAGTCCAGATACTTG
oDS62	GCCAAGCTCTGGCAAATTCATTGGGAG
oDS63	GACGCTCTGGAGTCCAGATACTTGCATC
oDS64	CCAGAGCGTCTCTGGCAAATTCATTGG
oDS65	GAAGCTGCCAGAGACTGGCTGGAGTCC
oDS66	CTGGCAGCTTCATTGGGAGCCGTGATTG
oDS67	CCAGCTGAAATTCCGACACTTGAAGAC
oDS68	GAATTTCAGCTGGCGGCCGGTCACTTGAAATG
oDS69	CAAAGGCTATTCCGACACTTGAAGACAGATTG
oDS70	GAATAGCCTTTGGCGGCCGGTCACTTG
oDS71	GAAGCTAGATTGCGCTCACGTTTTGAATG
oDS72	CAATCTAGCTTCAAGTGTCGGAATTTCCTTTG
oDS73	GATTCCGCTCTTCCTAAAATCGGTGAAGAG
oDS74	GAAGAGCGGAATCAGTCATTTCCCTTG
oDS75	CTCTGCTCCTAAAATCGGTGAAGAGTTTG
oDS76	GATTTTAGGAGCAGAGGAATCAGTCATTTCCC
oDS77	CCTGCTATCGGTGAAGAGTTTGGAGGAC
oDS78	CGATAGCAGGAAGAGAGGAATCAGTCATTTC
oDS79	CATACGGCTGTTATTCATGCGCATGAAAAAATTTC
oDS80	GAATAACAGCCGTATGATCACGTCCTCCAAAC
oDS84	ATGTGCTTTTGATAGATGATATTC
oDS85	GCCGGTCACTTGAAATGACG
oDS86	TTATGCCGAAACCGCAAGTC
oDS87	CCAGAAGAAATTCCGACACTTGAAGAC
oDS88	GAATTTCTTCTGGCGGCCGGTCACTTGAAATG

oDS91	CAAAGAAAATTCCGACACTTGAAGACAGATTG
oDS92	GAATTTTCTTTGGCGGCCGGTCACTTG
oDS93	GAAAAAAGATTGCGCTCACGTTTTGAATG
oDS94	CAATCTTTTTTCAAGTGTCGGAATTTCCTTTG
oDS95	GATTCCGACCTTCCTAAAATCGGTGAAGAG
oDS96	GAAGGTCGGAATCAGTCATTTCCCTTG
oDS97	CTCTGACCCTAAAATCGGTGAAGAGTTTG
oDS98	GATTTTAGGGTCAGAGGAATCAGTCATTTCCC
oDS99	CCTGAAATCGGTGAAGAGTTTGGAGGAC
oDS100	CGATTTCAGGAAGAGAGGAATCAGTCATTTC
oDS106	GAAAGCTCAAACCCAGGAAGAATTTTTC
oDS107	GTTTGAGCTTTCCCCGCTAAAAATTGAATATC
oDS124	CAGAACTTGGGTTTCTGTGATGACG
oDS131	ACGTGGATCCCAAAAAACAGCAATTTATTGATATGC
oDS132	ACGTGGTACCCGGCTTTTTTGTTCCCCTTCTGCTTTTC
oDS133	ACGTGGATCCCAGCCTTTATACCATTTTTGAGGAAG
oDS134	ACGTGGTACCCGTTGTTCAAGCCAATTGTAAAAAGG
oDS136	CCATCCCCAATCTTCTGCTCAC
oDS137	CGCCTTTGGAGTGTGAAACG
oDS138	CTTCTGCTAAATTTACAGCAGAATTCATCAACTCTATCC
oDS139	AAATTTAGCAGAAGACAGATAAACCACTTTG
oDS142	TGTTCCCTGCTCCTGCATATC
oDS143	GGCGTGTACAAGCAAAGCAG
oDS154	CAGAGAAAGACGCCATTTTCTAAGAAAAGG
oDS155	AGTTATTCACACTTTCCCCGATTGATCCCCGGTC
oDS158	GAAGCAATCTGCTCTATCCGAGATAATAAATG
oDS159	CAGATTGCTTCGTTTGTAAATTTCTCAGAAG
oDS162	CTCTGCACGAGATAATAAATGCGTCGAC
oDS163	CTCGTGCAGAGCAGATGAATTCGTTTG
oHM272	GGGGGGACGTCTAAGAAAATATATTAGACCTGTGGAAC
oHM665	ACGTGGATCCCGAAAATATATTAGACCTGTGGAAC
oHM666	ACGTGGTACCCGTTTAAGCTGTTCTTTAATTTC
oHM667	ACGTGGATCCCAAAAAACAGCAATTTATTGATATGCAGG
oHM668	ACGTGGTACCCGTTGTTCAAGCCAATTGTAAAAAGGAAC
oTS001	GAGGCATTTACAAACGAATTCATCAAC
oTS003	GTAAATGCCTCAGAAGACAGATAAACCAC
oTS004	GAAATTTACAGCAGAATTCATCAACTCTATCCGAG
oTS005	GAATTCTGCTGTAAATTTCTCAGAAGACAG
oTS006	CAAACGCATTCATCAACTCTATCCGAGATAATAAAG
oTS007	GATGAATGCGTTTGTAAATTTCTCAGAAG
oTS008	GAAGCAATCAACTCTATCCGAGATAATAAAG
oTS009	GTTGATTGCTTCGTTTGTAAATTTCTCAG
oTS011	CATCGCATCTATCCGAGATAATAAAGCC
oTS012	GATAGATGCGATGAATTCGTTTGTAAATTTC
oTS014	CTCTGCACGAGATAATAAAGCCGTCGAC
oTS015	CTCGTGCAGAGTTGATGAATTCGTTTG

oTS016	CTATCGCAGATAATAAAGCCGTCGACTTC
oTS017	CTTTATTATCTGCGATAGAGTTGATGAATTCGTTTG
oTS018	CGAGCAAATAAAGCCGTCGACTTCCGC
oTS019	CTTTATTTGCTCGGATAGAGTTGATGAATTC
oTS020	GATGCAAAAGCCGTCGACTTCCGCAATC
oTS021	CTTTTGCATCTCGGATAGAGTTGATGAATTC
oTS022	GATAATGCAGCCGTCGACTTCCGCAATCGC
oTS023	CGGCTGCATTATCTCGGATAGAGTTGATG
oTS024	CTTCGCAAATCGCTATCGAAATGTTGATG
oTS025	GATTTGCGAAGTCGACGGCTTTATTATC

Table 2.5. Oligonucleotides used to assemble DNA Scaffolds

Oligonucleotide	**Sequence (5′→3′)**
5’ Tail DnaA-trio Sequence	
oTR602	ACTTATCCACAAATCCACAGGCCC
oTR609	TGAATAGGTGTTTAGGTGTCCGGGATGATAATGAAGATGATAAAAA ATATTTATATATA
3’ Tail	
oTR603	TGAATAGGTGTTTAGGTGTCCGGG
oTR608	ACTTATCCACAAATCCACAGGCCCTACTATTACTTCTACTATTTTT TATAAATATATAT
Δ DnaA-boxes	
oTR612	ACCTTCCTTGCTTCCTTGCGGCCC
oTR611	TGGAAGGAACGAAGGAACGCCGGGATGATAATGAAGATGATAAAAA ATATTTATATATA
5’ Tail Complementary Sequence	
oTR602	ACTTATCCACAAATCCACAGGCCC
oTR619	TGAATAGGTGTTTAGGTGTCCGGGTACTATTACTTCTACTATAAAA ATATTTATATATA

Table 2.6. Oligonucleotides used for DNA strand seperation assays

Oligonucleotide	**Sequence (5′→3′)**
Wild-type	
oHM778	ACTTATCCACAAATCCACAGGCCC
oHM588	CCTTAGCAGGAGACATTCCTTC
Cy5-oHM590	TACTATTACTTCTACTA
Δ DnaA-trios	
oHM778	ACTTATCCACAAATCCACAGGCCC
oHM595	AATCATTTACATCTTCTGGGCCTGTGGATTTGTGGATAAGT
Cy5-oHM596	AGAAGATGTAAATGATT

Chapter 3

Activities specifically required by DnaA proteins delivered to the site of DNA unwinding from an upstream origin subregion

Chapter 3 – Introduction

As described in Chapter One, the essential process of triggering new rounds of DNA replication in bacteria occurs at the origin of replication (*oriC*) and is coordinated by DnaA, the master initiator protein. DnaA performs several key activities to begin initiation: I) recognising and binding to *oriC*, II) assembling into a filament upon the DnaA-trios, and III) promoting unwinding of the DNA duplex (Richardson *et al.*, 2016). As discussed in Sections 1.2 and 1.3 DnaA is guided in performing these activities by several essential DNA sequence motifs that constitute *oriC*.

Investigation into the key DNA sequences of the unwinding region of the *Bacillus subtilis* origin (*incC*) identified the minimal architecture required to guide DnaA and form a functional origin (Richardson *et al.*, 2019). These essential motifs, highlighted in Figure 3.1.A, are the DnaA-trios and a small number of DnaA-boxes. Systematic genetic analysis has indicated that the two DnaA-boxes proximal to the DnaA-trios (#6 and #7) and a single upstream box are the only specific double stranded *incC* DNA sequences required to support origin unwinding by DnaA.

Further analysis determined that for functional origin activity the upstream DnaA-box needed to be located a minimum of 44 base pairs upstream of DnaA-box#6. This position corresponds to the location of the native DnaA-box#3. This DnaA-box could still support origin activity if moved further upstream but displayed phasing as it could only support function if located on the same face of the DNA helix. The result was independent of which direction the DnaA-box was facing. Critically, replication initiation was still supported when the upstream box was moved 132, 297 and 462 base pairs from DnaA-box#6. These observations led to the hypothesis that the distal DnaA-box is acting through a DNA loop (Richardson *et al.*, 2019).

The variability in positioning of the upstream DnaA-box suggested the DnaA-loop might function to increase local concentration of DnaA at the downstream region. To test this hypothesis it was determined if the requirement for the distal DnaA-box could be bypassed by DnaA overexpression.

Figure 3.1.B presents a schematic of a strain carrying a xylose inducible second copy of *dnaA* (and *dnaN*). As Figure 3.1.C shows (data taken from Richardson *et al.,* 2019), increasing DnaA concentration by inducing expression of the second copy of *dnaA* rescues growth of a strain carrying an artificial *incC* composed of just the proximal DnaA-boxes (#6 and #7) and all DNA upstream of DnaA-box#6 replaced with a synthetic sequence.

Therefore, the current model for the unwinding of the origin supported by this recent data is that the DnaA-boxes proximal to the site of unwinding promote the loading of DnaA onto the DnaA-trios. This DnaA will then assemble into a filament to promote DNA strand separation, most likely by preventing complimentary base-pairing through stretching the DNA strand containing the DnaA-trios (Duderstadt *et al.*, 2011).

In the model described above, the upstream DnaA-box functions to increase the local concentration of DnaA at the site of unwinding, with the protein being delivered through a DNA loop forming within *incC*. If so, and the mechanism for strand separation is stretching of the DnaA-trios, then the DnaA protein delivered from the upstream DnaA-box should require both filament formation and single-stranded DNA binding activities.

This chapter investigates the activities required by the DnaA protein specifically bound to the upstream DnaA-box using an *in vivo* genetic approach.

In order to investigate the activities required *in vivo*, there were two main obstacles that needed to be overcome. Firstly in order to determine which DnaA activities are required, the protein needed to be rendered non-functional for these activities but still capable of being studied *in vivo*. Secondly, to identify which activities are specifically required at the upstream binding site there needed to be a way to differentiate between the functions required by the protein binding upstream to that binding downstream, when the same protein could bind to either site.

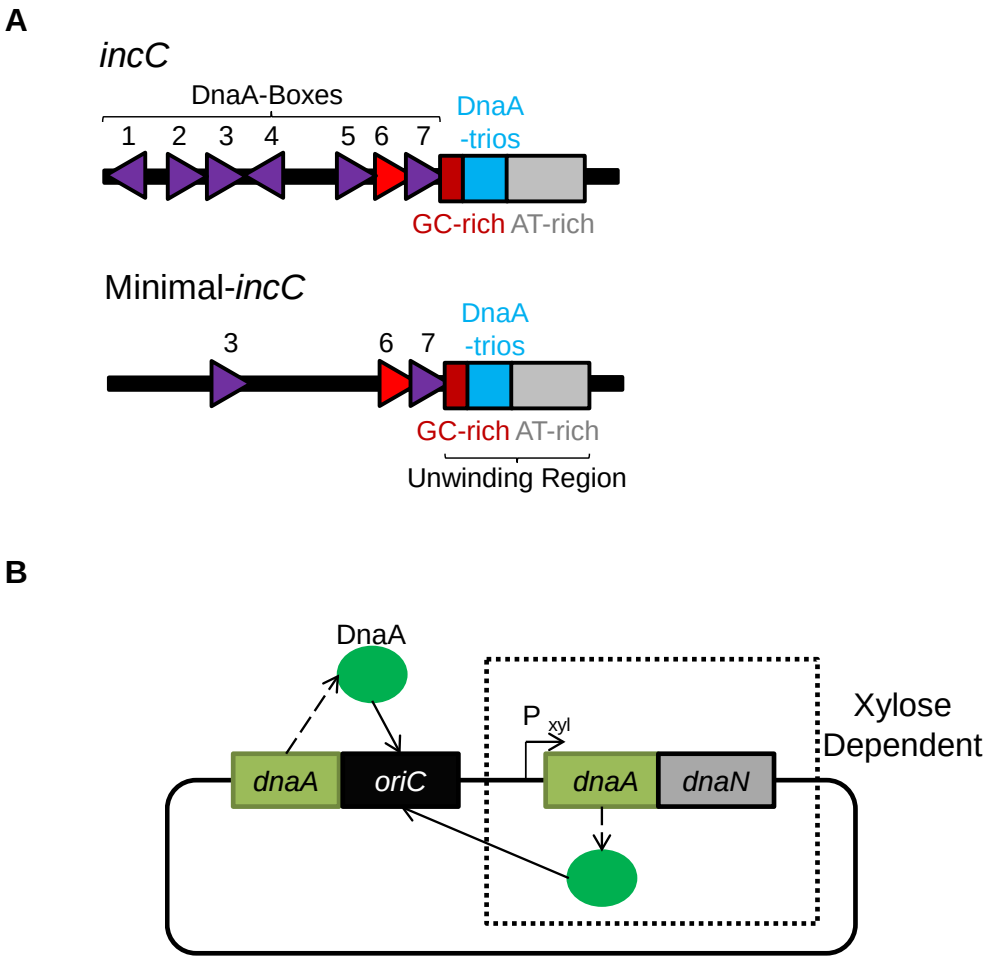


Figure 3.1. Minimal *incC* architecture required to promote DNA unwinding. (A) The unwinding region of *oriC*, *incC*, and the minimum DNA sequences required for a functional origin. DnaA-boxes carrying the consensus sequence are highlighted in red. **(B)** Schematic of a strain capable of overexpressing DnaA (and DnaN) in a xylose dependent manor. **(C)** DnaA overexpression rescues the origin activity of an artificial *incC* strain formed only of DnaA-boxes 6 and 7 and synthetic DNA upstream (orange) (Richardson *et al.,* 2019).

Chapter 3 – Results

3.1. Essential DnaA activities can be investigated *in vivo* by bypassing *oriC*

The first challenge for the *in vivo* investigation into the activities required by the DnaA protein binding to the distal DnaA-box, was finding a way to render the protein non-functional for these essential activities, while still allowing the cells to grow.

To enable the investigation of lethal *dnaA* mutations, a strain capable of initiating replication through either the native *oriC*-DnaA system or independently of it, in an inducible manner, was utilised. A schematic of this strain is shown in Figure 3.2.A. This strain can replicate from a secondary plasmid origin, *oriN*, integrated into the chromosome (Richardson *et al.*, 2016). Replication from *oriN* requires the initiator protein RepN, the expression of which is under the control of an IPTG (Isopropyl β-D-1-thiogalactopyranoside) inducible promotor, allowing activity at *oriN* to be shut off once *dnaA* mutations are introduced to observe their effects. As shown in Figures 3.2.B and C where *dnaA* and *oriC* have been rendered non-functional respectively, RepN and *oriN* can provide for the essential activities of a replication origin.

Cells that are dependent upon oriN/RepN for initiating replication show a lighter appearance compared to wild type cells (Figure 3.2). The reason is unclear at this stage but is likely due to these cells no longer being able to undergo sporulation. *B. subtilis* spores are opaque and so when these form they give the colonies the darker appearance identifiable at the later time points in Figure 3.2.B and C. While it is unknown why these cells are unable to sporulate, a possible explanation is the loss of the ability to regulate re-initiation of replication from oriN.

As has been explored in Sections 1.3.3 and 1.3.4, multiple amino acid residues have been implicated in either filament formation (Duderstadt *et al.*, 2010; Ozaki *et al.*, 2012) or single strand DNA binding (Ozaki *et al.*, 2008; Duderstadt *et al.*, 2011) based on information obtained via protein structures or *in vitro* assays using *E. coli*, *T. maritima* and *A. aeolicus*. To determine which of these amino acids are physiologically relevant and essential for *B. subtilis* DnaA activity *in vivo,* the native *dnaA* gene was subjected to site-directed mutagenesis (see 2.5.1) within the inducible *oriN* strain.

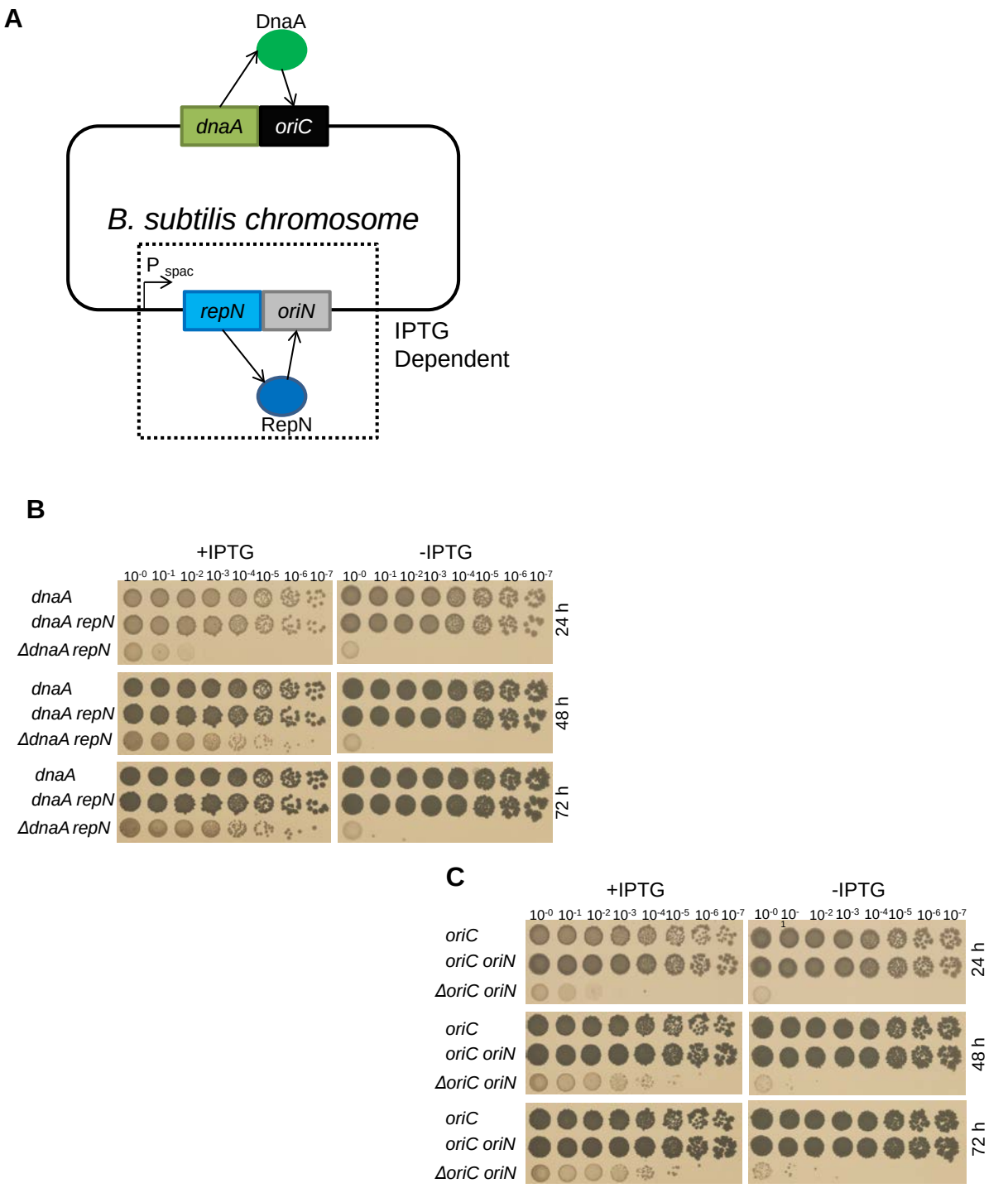


Figure 3.2. A tool for investigating DnaA and *oriC in vivo*. (A) Schematic of the chromosome of the *B. subtilis* strain capable of replicating both through *oriC* and independent of it. **(B)** Growth of strains carrying a non-functional DnaA protein (*ΔdnaA*) are dependent upon the activity of *repN* which can be shut on or off in an IPTG dependent manner. **(C)** Growth of strains carrying a non-functional origin (*ΔoriC*) are dependent upon the activity of *oriN* which can be shut on or off in an IPTG dependent manner. *ΔdnaA* is a truncated protein lacking domains I and II, *ΔoriC* has the *incC* region of the origin deleted. *dnaA/oriC* (168CA), *dnaA repN/oriC oriN* (HM1108), *ΔdnaA repN* (TR320), *ΔoriC oriN* (HM1603).

3.2. Residues proposed as being required for DnaA filament formation are essential in *Bacillus subtilis in vivo*

The first DnaA activity investigated was filament formation. As discussed in Section 1.3.3 many residues within the AAA+ domain (domain III) of DnaA have been implicated in forming the interfaces of the AAA+/AAA+ interaction required for guiding DnaA oligomerisation into a filament. The majority of the evidence implicating these residues is from studies utilising several Gram-negative bacteria (*E. coli*, *T. maritima* and *A. aeolicus*). Most residues being proposed are based on insights from crystal structures and the *in vitro* activity of purified proteins, with only a small selection of residues investigated *in vivo* (Duderstadt *et al.*, 2010; Ozaki *et al.*, 2012). As such, the physiological relevance of many of these residues remains unknown, as does their relevancy in *B. subtilis* (with the exception of the arginine finger residue (R264) which has been previously investigated *in vitro* (Scholefield *et al.*, 2012)). These residues are highlighted in Figure 3.3.A with residues previously investigated *in vivo* underlined.

The relevancy of these residues in *B. subtilis* was investigated by determining which amino acids were essential for a functional DnaA protein *in vivo*. This was achieved though site-directed mutagenesis of the endogenous *dnaA* gene, followed by characterisation of protein variants using the inducible *oriN* system discussed above.

The proposed residues were individually substituted for alanine (A), except for the glycine at position 317 which was substituted for glutamine (Q) (due to unknown difficulty substituting it with alanine). The results revealed that eight of the residues (R202, R206, F218, H231, R264 (the arginine finger), L269, G317, R231) significantly inhibit growth when substituted (Figure 3.3.B). The remaining residues (F128, E183, N187) reduce growth rate when substituted (see Figure 3.4.B), but do not fully inhibit it as the strains do not become IPTG dependent. An immunoblot showed the mutant proteins are all being stably expressed at levels similar to wild-type (Figure 3.3.C).

Taken together the results suggest that all the residues proposed as being involved in filament formation are required for optimal growth, with eight being absolutely essential for DnaA function *in vivo*.

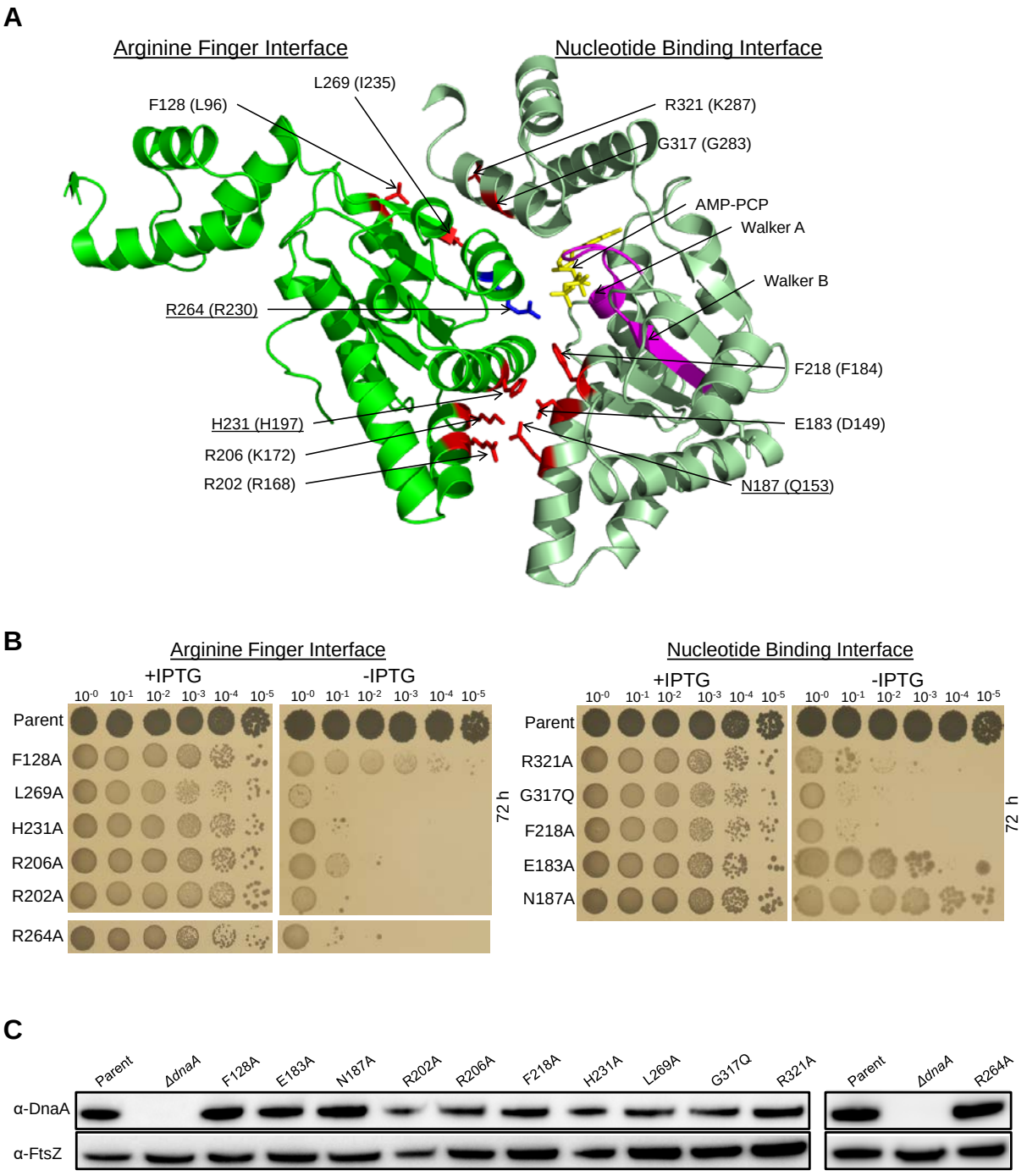


Figure 3.3. *In vivo* requirement for residues implicated in DnaA filament formation in *B. subtilis*. (A) DnaA residues implicated in forming the AAA+/AAA+ interfaces critical for filament formation mapped onto the crystal structure from *A. aeolicus* (PDB ID 2HCB). Residues are labelled according to the *B. subtilis* numbering, with *A. aeolicus* shown in parentheses. Residues investigated *in vivo* previously in homologs are underlined. **(B)** Analysis of *B. subtilis* DnaA substitution mutants using the *oriN* strain. Residues have been divided according to the two separate AAA+ interfaces and growth is shown after 72 hours incubation at 37°C. **(C)** Immunoblot analysis of the DnaA substitution mutants with the tubulin homolog FtsZ as a loading control and a strain lacking *dnaA* as an antibody control. Parent (HM1108), F128A (DS25), E183A (DS6), N187A (DS23), R202A (DS21), R206A (DS22), F218A (DS26), H231A (DS50), R264A (DS56), L269A (DS34), G317Q, (DS51), R321A (DS27), *ΔdnaA* (HM1424).

One interesting result to arise from the investigation into the proposed filament formation residues was the contrast on the effect of viability between the residues E183 and N187 with R202, R206 and H231. In the structure of the AAA+/AAA+ interaction (Figure 3.3.A), these residues are in close proximity to one another and appear to potentially interact. A space fill of this section of the crystal structure makes this more clear (Figure 3.4.A). It was surprising therefore that the residues on one of the interfaces of the filament were absolutely required for viability while those on the other were not.

To further investigate these residues, E183 and N187 were substituted together within the *oriN* strain. The result shows that while the single alanine substitutions give a slow growing phenotype independent of IPTG, the double mutant is inviable (Figure 3.4.B). The mutant protein was expressed similar to wild-type as shown by immunoblot (Figure 3.4.C). This result suggests that for viability at least one of either the E183 or N187 residues is required, presumably to interact with the essential residues on the adjacent DnaA protein.

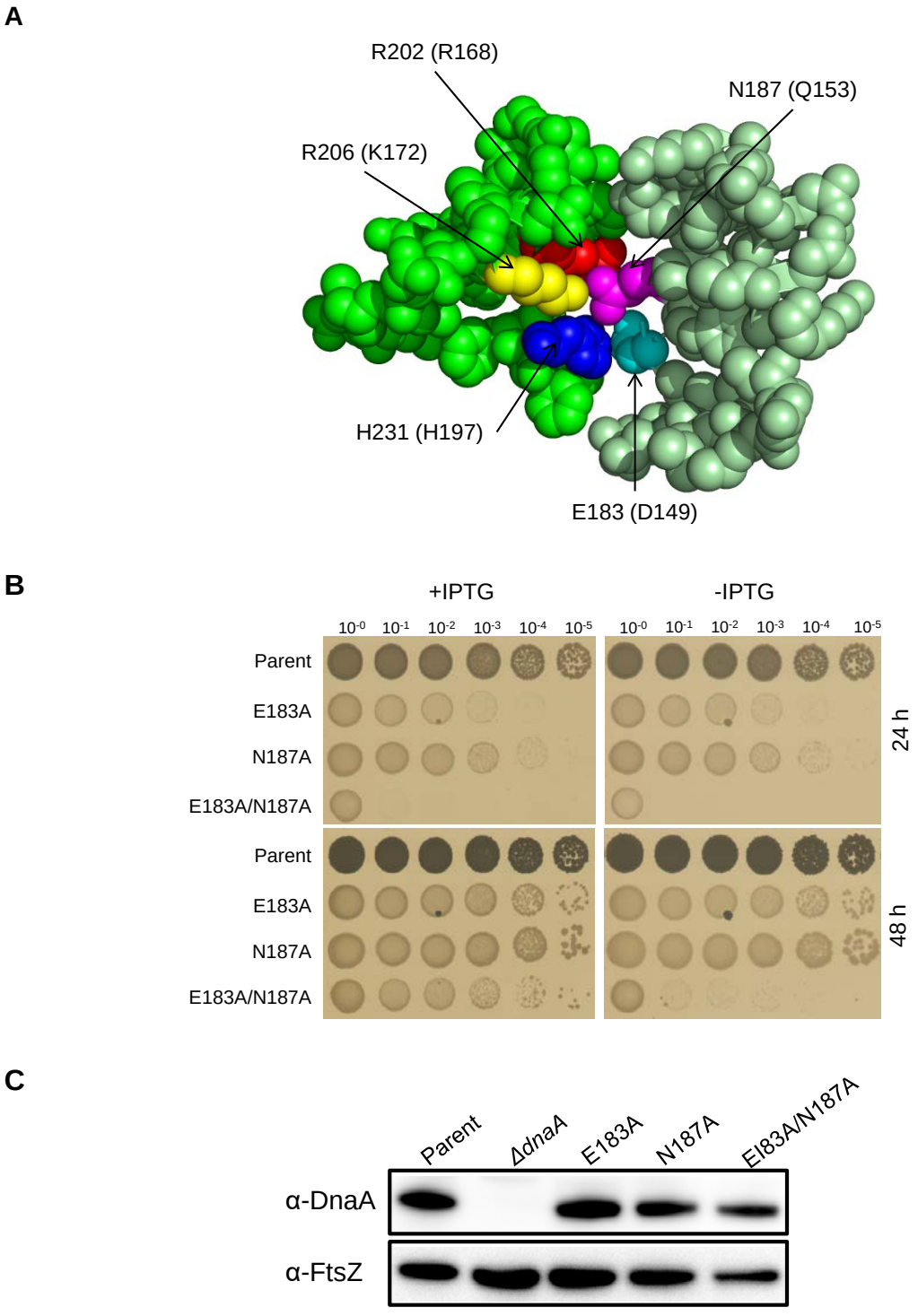


Figure 3.4. Investigating the slow growing filament formation substitutions. (A) Space filled crystal structure (PDB ID 2HCB) of a subsection of the AAA+/AAA+ interfaces from *A. aeolicus* highlighting residues that appear to directly interact with one another during DnaA filament formation. Residues are labelled according to the *B. subtilis* numbering, with *A. aeolicus* shown in parentheses. **(B)** Analysis of the substitution mutants for E183, N187 and a double substitution of both in the *oriN* strain. The result is following 24 and 48 hours incubation at 37°C. **(C)** Immunoblot analysis of the DnaA substitution mutants with the tubulin homolog FtsZ as a loading control and a strain lacking *dnaA* as an antibody control. Parent (HM1108), E183A (DS6), N187A (DS23), E183A/N187A (DS67), *ΔdnaA* (HM1424).

3.3. The DBD/AAA+ DnaA filament interface is not physiologically relevant in *Bacillus subtilis*

3.3.1. The proposed DnaA DBD/AAA+ interaction interface residues are not required in vivo when substituted for alanine

As discussed in Section 1.3.3 it has been proposed that there is a second interaction interface within the DnaA filament between the AAA+ domain (domain III) and the DNA binding domain or DBD (domain IV) of the adjacent protomer (Duderstadt *et al.*, 2010). It has been proposed that the DnaA-filament undergoes a conformational change to enable the binding of single-stranded DNA, with the DBD of one protomer packing into the AAA+ domain of another (Duderstadt *et al.*, 2010). Again the evidence for this second filament interface is based on crystal structures and *in vitro* protein assays, with very few residues investigated *in vivo* (and none in *B. subtilis*).

The amino acids implicated in forming the proposed DBD/AAA+ interface are highlighted in Figure 3.5.A, with residues previously investigated *in vivo* underlined. These positions were again substituted for alanine in the *oriN*-RepN strain, and the growth phenotype of these substitutions was determined on both solid and liquid media (Figure 3.5.B-C). The results suggest that none of the residues are individually required for protein function as all of the substitutions could support growth in the absence of IPTG. However, some cells did display a lighter colour on solid media (most noticeably those carrying substitutions to the residues K252 and L402) and this correlated with a possibly reduced growth rate in liquid media (Figure 3.5.C). An immunoblot confirmed the DnaA variants were all being stably expressed at levels similar to wild-type (Figure 3.5.D). Growth rate in liquid media was determined by measuring the amount of light passing through the culture overtime, with the assumption that less light passing through correlates with a denser cell population and so a higher number of cells. A number of factors other than bacterial cell numbers can influence this result, including changes in cell size or shape, as well as dead cells still contributing to the density. As such the results shown in Figure 3.5.C may not be a genuine change in growth rate and other assays, such as time-lapse microscopy, may provide a better indication of differences in growth rate.

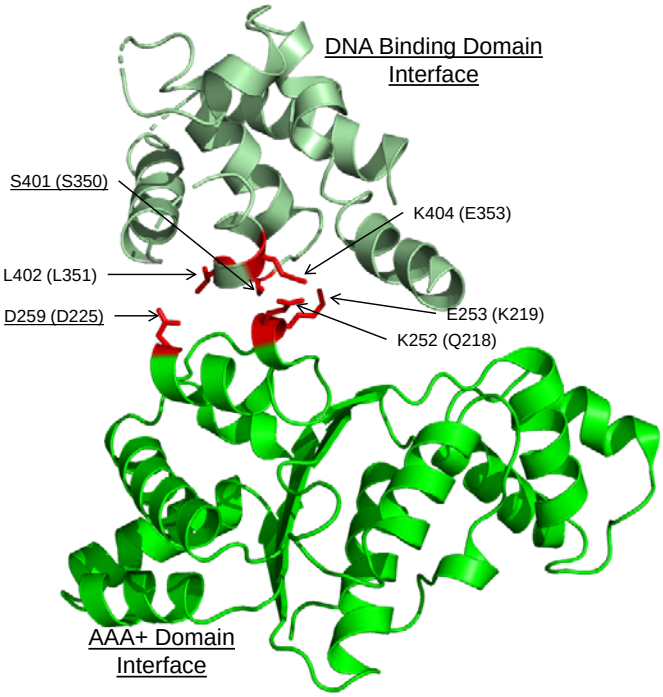
As mentioned when the residues at position 252 and 402 were substituted the resultant strains displayed a lighter colour and slight reduction in growth rate in liquid media. During times of low nutrient availability *B. subtilis* can undergo sporulation, a

developmental pathway which results in it forming a hardened endospore in which it can survive until conditions improve. The endospore is opaque giving the bacteria the darker appearance. One explanation for the lighter phenotype is that the cells are no longer able to form spores, however this proved not to be the case (data not shown). Further to this it is possible that cells are unable to sporulate when replication is dependent upon *oriN* (Section 3.1) and the lighter colour seen with these mutant cells does not appear comparable to these non-sporulating cells. The ability of these cells to sporulate also suggests the lighter colouring and potential slower growth rate may not be related to initiation.

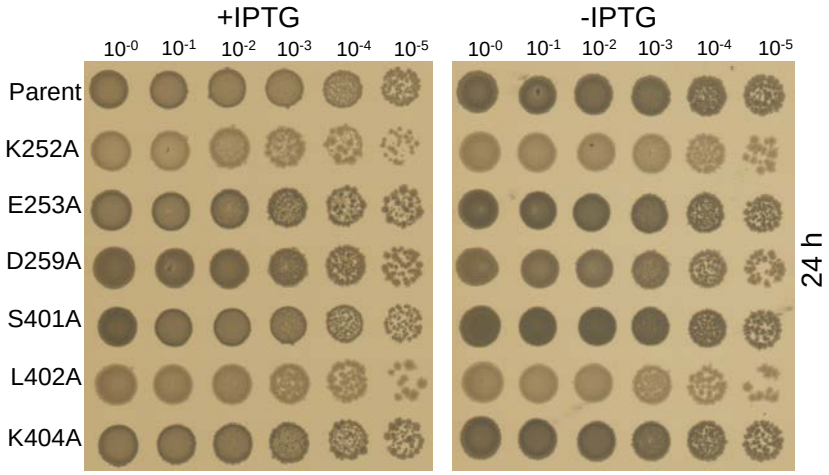
Another explanation is that the cells are undergoing lysis, which could also explain the slight growth rate reduction, if there is one. If lysis is occurring then cell membranes will be compromised which can be detected using a fluorescent nucleic acid dye (Figure 3.5.E). The dye used, SYTOX green, cannot cross intact membranes but will penetrate a damaged or otherwise compromised membrane. Very few cells gave a positive signal when staining the K252A and L402A cells with SYTOX green, suggesting the membranes of these cells were intact and the phenotype was not due to lysis (Figures 3.5.E and F) furthermore the phase contrast images show the cells appear to be intact providing further evidence the cells are not lysing. Heat shocked cells were readily stained with SYTOX green, confirming that the assay was operational under these imaging conditions (Figure 3.5.E and F). As such the reason for the lighter phenotype remains unknown.

Taken together, the results indicate that all of the amino acid side chains at the proposed DBD/AAA+ interface can be replaced with alanine without disrupting DnaA function. Thus, the DBD/AAA+ interface may not form in the DnaA filament of *B. subtilis*, or it may form and not be physiologically relevant *in vivo*.

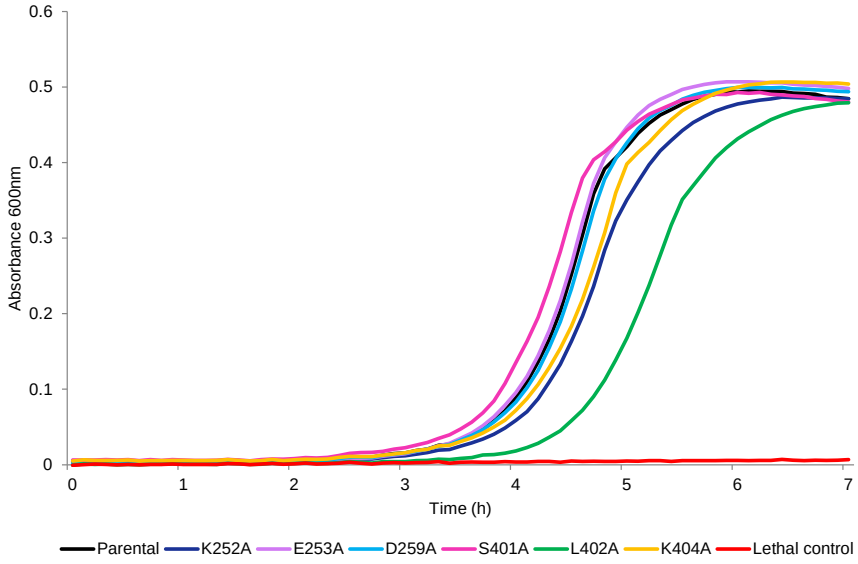
A


DNA Binding Domain Interface
S401 (S350)
K404 (E353)
L402 (L351)
D259 (D225)
E253 (K219)
K252 (Q218)
AAA+ Domain Interface

B


+IPTG
-IPTG
10-0 10-1 10-2 10-3 10-4 10-5
Parent
K252A
E253A
D259A
S401A
L402A
K404A
24 h

C


Absorbance 600nm
0.6
0.5
0.4
0.3
0.2
0.1
0
0 1 2 3 4 5 6 7
Time (h)
Parental K252A E253A D259A S401A L402A K404A Lethal control

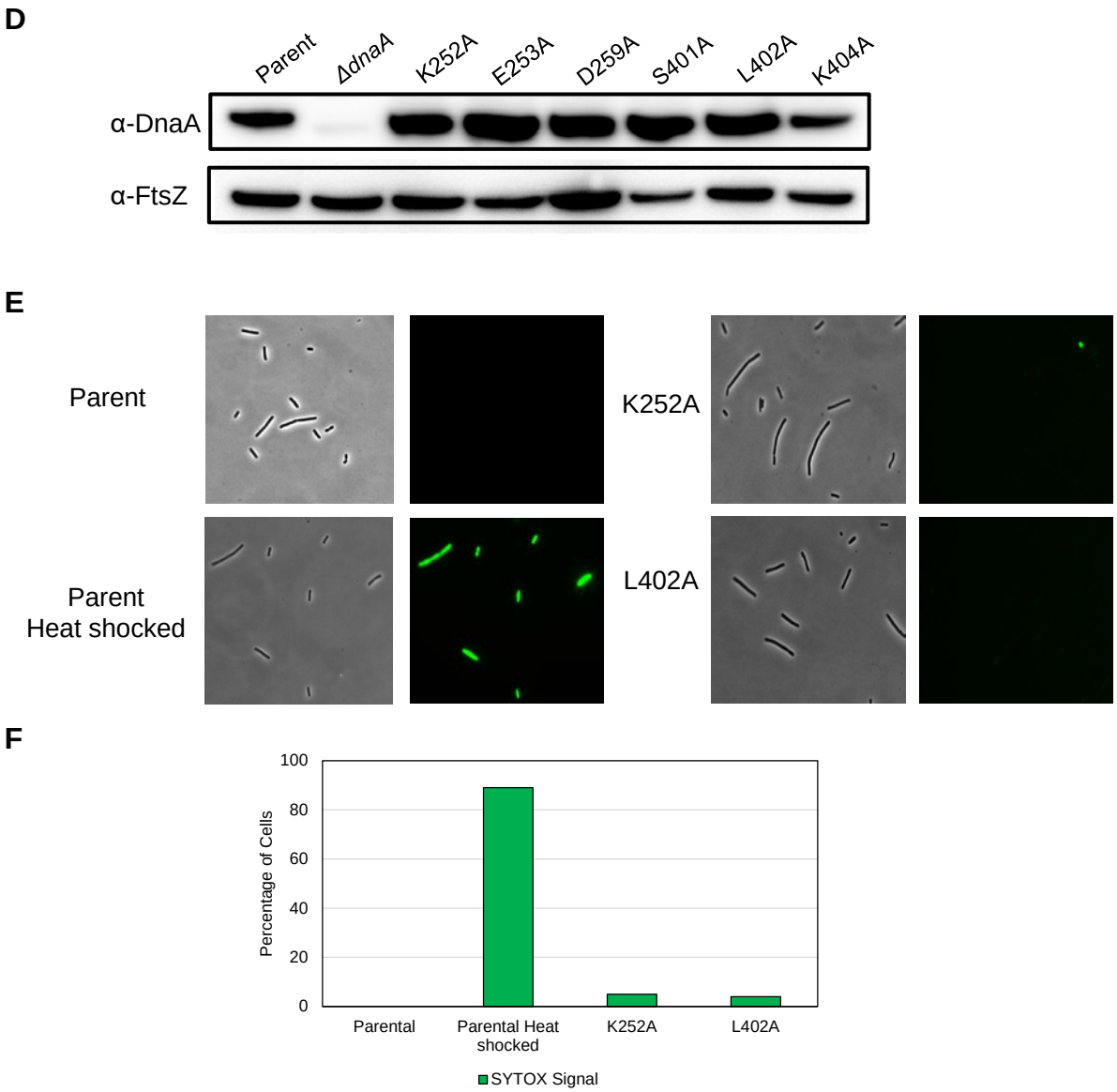


Figure 3.5. Essentiality of residues implicated in the DBD/AAA+ DnaA filament interface in *B. subtilis* when substituted for alanine. (A) DnaA residues implicated in forming the DBD/AAA+ interfaces of the DnaA filament mapped onto the crystal structure from *A. aeolicus* (PDB ID 2HCB). Residues are labelled according to the *B. subtilis* numbering, with *A. aeolicus* shown in parentheses. Residues investigated *in vivo* previously in homologs are underlined. **(B)** Analysis of *B. subtilis* DnaA alanine substitution mutants using the *oriN*/*repN* strain. The result is shown after 24 hours incubation at 37°C. **(C)** Analysis of the growth of the same DnaA substitution mutants from B in liquid culture at 37°C compared to a strain carrying a non-functional DnaA (ΔDnaA domains I-II). **(D)** Immunoblot analysis of the DnaA substitution mutants using the tubulin homolog FtsZ as a loading control and a strain lacking *dnaA* as an antibody control. **(E)** Fluorescence microscopy showing SYTOX green staining of the parental strain before and after heat-shocking and the K252A and L402A DNA variant strains. **(F)** Quantification of SYTOX staining. Percentage of 100 cells from the same strains/conditions as E giving a positive (Green) signal following SYTOX staining.

Parent (HM1108), K252A (DS41), E253A (DS47), D259A (DS44), S401A (DS42), L402A (DS45), K404A (DS43), ΔDnaA domains I-II (TR320), *ΔdnaA* (HM1424).

3.3.2. Residues of the proposed DnaA DBD/AAA+ interaction interface are required in vivo when substituted more dramatically

Section 3.3.1 has shown that residues within the proposed DBD/AAA+ interface can be substituted for alanine without impacting viability, suggesting this interface may not be physiologically relevant. To further investigate this interface, a series of more dramatic substitutions were introduced *in vivo* using the previously described *oriN*-RepN strain. These substitutions are highlighted in Table 3.1.

Position	Native Residue	Charge	Substitution	Charge
252	Lysine (K)	Positive	Glutamic acid (E)	Negative
253	Glutamic acid (E)	Negative	Lysine (K)	Positive
259	Aspartic acid (D)	Negative	Lysine (K)	Positive
401	Serine (S)	-	Aspartic acid (D)	Negative
402	Leucine (L)	-	Aspartic acid (D)	Negative
404	Lysine (K)	Positive	Glutamic acid (E)	Negative

Table 3.1. Dramatic substitutions made to residues within the proposed DBD/AAA+ interface.

Figures 3.6.A-B shows the effect of these substitutions on growth and viability on solid and in liquid media respectively. Substitutions to residues 253 and 404 had no effect on growth or viability, as seen with the alanine substitutions. Substituting the residue at position 259 resulted in a slower growth rate in contrast to the alanine substitution. Finally substitution of residues at positions 252, 401 and 402 was lethal, with the cells dependent upon IPTG, suggesting that these protein variants are no-longer functional. An immunoblot showed that mutant proteins are stably expressed at wild-type levels (Figure 3.6.C).

Although these results revealed some residues at the proposed DBD/AAA+ interface that were essential, they only became essential after the introduction of dramatic amino acid substitutions, several of which involved the introduction of a charge. This is in contrast to the results of the AAA+/AAA+ interface where several alanine substitutions were sufficient for loss of DnaA function. Taking all the results together, it appears the proposed DBD/AAA+ interface is not critical for *B. subtilis* DnaA, at least under the conditions tested.

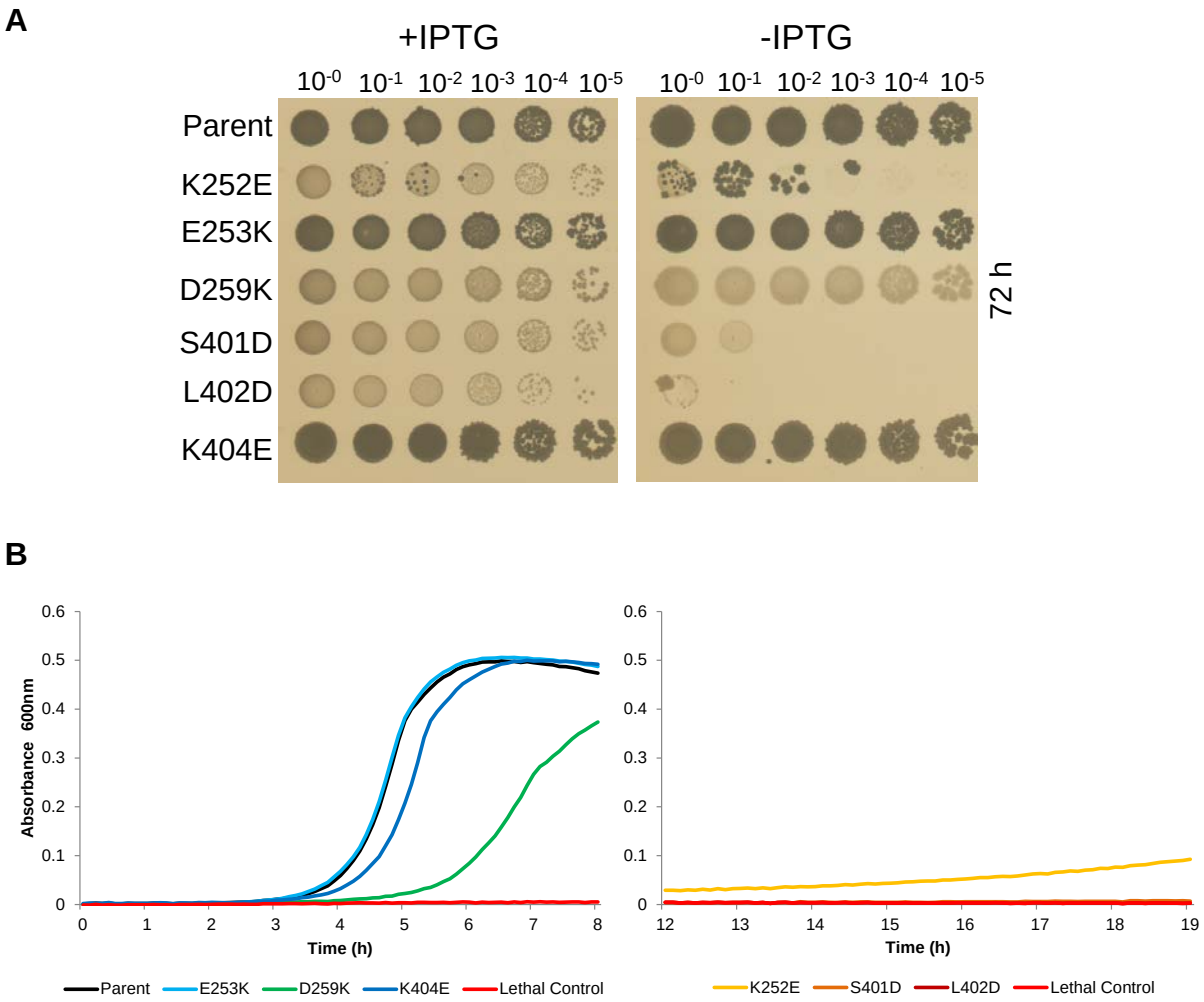


Figure 3.6. Essentiality of residues implicated in the DBD/AAA+ DnaA filament interface in *B. subtilis* when substituted more dramatically. (A) Analysis of *B. subtilis* DnaA substitution mutants using the *oriN*/*repN* strain. The result is shown after 72 hours incubation at 37°C. **(B)** Analysis of the growth of the same DnaA substitution mutants from B in liquid culture without IPTG at 37°C compared to a strain carrying a non-functional DnaA (ΔDnaA domains I-II). **(C)** Immunoblot analysis of the DnaA substitution mutants using the tubulin homolog FtsZ as a loading control and a strain lacking *dnaA* as an antibody control.

Parent (HM1108), K252E (DS71), E253K (DS48), D259K (DS49), S401D (DS72), L402D (DS70), K404E (DS69), ΔDnaA domains I-II (TR320), *ΔdnaA* (HM1424).

3.4. Residues proposed as being required for DnaA single-stranded DNA binding are essential in *Bacillus subtilis in vivo*

Several residues within the AAA+ domain of DnaA have been proposed as being involved in single strand DNA binding (Section 1.3.4). As in previous sections these results are based mainly on crystal structures and *in vitro* studies using *E.coli*, *T. maritima* and *A. aeolicus* (Ozaki *et al.*, 2008; Duderstadt *et al.*, 2011). Most of the residues have been previously identified as essential in *E. coli* via plasmid complementation tests. For *B. subtilis* DnaA, only one residue, Ile190 has previously been investigated *in vitro* with no investigation *in vivo* (Scholefield *et al*., 2012; Richardson *et al*., 2016). As such, the physiological relevance of these residues in *B. subtilis* is again unknown.

The implicated residues are highlighted in Figure 3.7.A and were substituted for alanine *in vivo* using the inducible *oriN*/RepN strain. The results show that I190 and K222 are required for growth, T225 is not essential but the alanine substitution does cause a growth defect, and Q224 is not required as substituting this residue did not produce a growth phenotype. An immunoblot showed that the mutant proteins are all stably expressed at levels compatible with wild-type (Figure 3.7.C).

These results indicate that this region of the protein is critical for DnaA function *in vivo,* but suggests that in *B. subtilis* the precise single-stranded DNA binding mechanism of DnaA may occur slightly differently to that of *A. aeolicus* with a synthetic DNA substrate (poly-A_{12}).

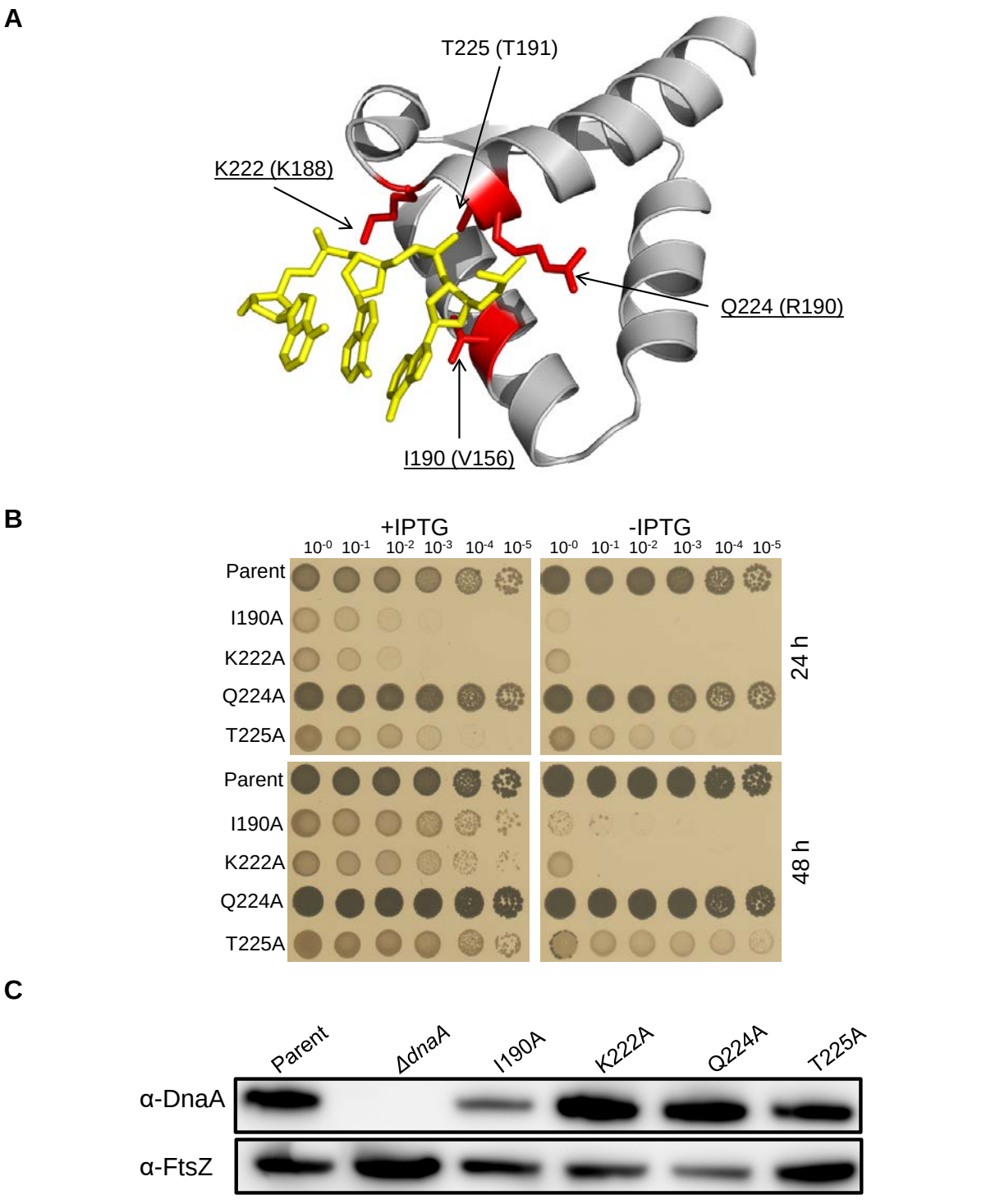


Figure 3.7. *In vivo* requirement for residues implicated in DnaA single-stranded DNA binding in *B. subtilis*. (A) DnaA residues implicated in binding single-stranded DNA mapped onto the crystal structure from *A. aeolicus* (PDB ID 1L8Q) labelled according to *B. subtilis* numbering with *A. aeolicus* in parentheses. Residues investigated *in vivo* previously in homologs are underlined. **(B)** Analysis of *B. subtilis* DnaA substitution mutants in the *oriN*/*repN* strain. The result is shown following 24 and 48 hours incubation at 37°C. **(C)** Immunoblot analysis of the DnaA substitution mutants with the tubulin homolog FtsZ as a loading control and a strain lacking *dnaA* as an antibody control. Parent (HM1108), I190A (DS18), K222A (DS53), Q224A (DS54), T225A (DS52), *ΔdnaA* (HM1424).

3.5. A DnaA chimera can enable investigation into the protein functions required by DnaA binding specifically to the upstream *incC* subregion

Essential DnaA residues implicated in filament formation and ssDNA binding have now been identified. It is now possible to render the protein non-functional for these activities by subsitituing these residues. To determine if filament formation or ssDNA binding are functions specifically required by the protein binding to the upstream *incC* binding site an assay is required which can differentiate between the protein activities required by DnaA binding specifically to the upstream site, from that binding downstream.

To investigate DnaA activities specifically required by DnaA binding to the upstream DnaA-box an *in vivo* genetic system was developed. As shown in Section 1.2.2 *B. subtilis* DnaA, like DnaA from a majority of bacteria, recognises the consensus DnaA-box sequence (5′-TTATCCACA-3′). However the DnaA protein of *T. maritima* has several amino acid substitutions within the DNA binding domain altering the DnaA-box sequence it binds to (5′-AAACCTACCACC-3′). This allows the creation of a DnaA chimera composed of *B. subtilis* domain I-III and domain IV of *T. maritima* as shown in Figure 3.8.A (Noguchi *et al.*, 2015). The chimeric DnaA ($DnaA^{chi}$) will bind specifically to the *T. maritima* DnaA-box while retaining the native domains required for filament formation and ssDNA binding. As such the $DnaA^{chi}$ should still be able to perform the required activities in relation to the B. subtilis DnaA-trio sequence, despite the shorter trio sequence in *T. maritima* (see Figure 1.7), as the AAA+ domain (domain III) is proposed to be required for this interaction.

To establish a system for investigating activities required at the upstream DnaA-box in *B. subtilis* a *T. maritima* DnaA-box sequence was introduced into the artificial minimal *incC* (see introduction to this chapter) in the position of the native DnaA-box 3 (Figure 3.8.B). This hybrid *incC* (*$incC^{chi}$*) was created within the *oriN* strain with a wild-type chimeric *dnaA* gene introduced under the control of a xylose-dependent promoter. The schematic for this strain is shown in Figure 3.8.C.

As shown in Figure 3.8.D the hybrid *$incC^{chi}$* origin only supports growth in the presence of xylose (or IPTG) revealing the origin is non-functional without the wild-type $DnaA^{chi}$. This result confirms the native DnaA protein cannot recognise the *T. maritima* box. Equally importantly a functional origin requires the *T. maritima* box, demonstrating the wild-type chimera is providing the activities required by DnaA binding to the upstream

box. This result validates that the chimeric system can be used to determine which activities of DnaA are required by the protein bound to the upstream sub-region.

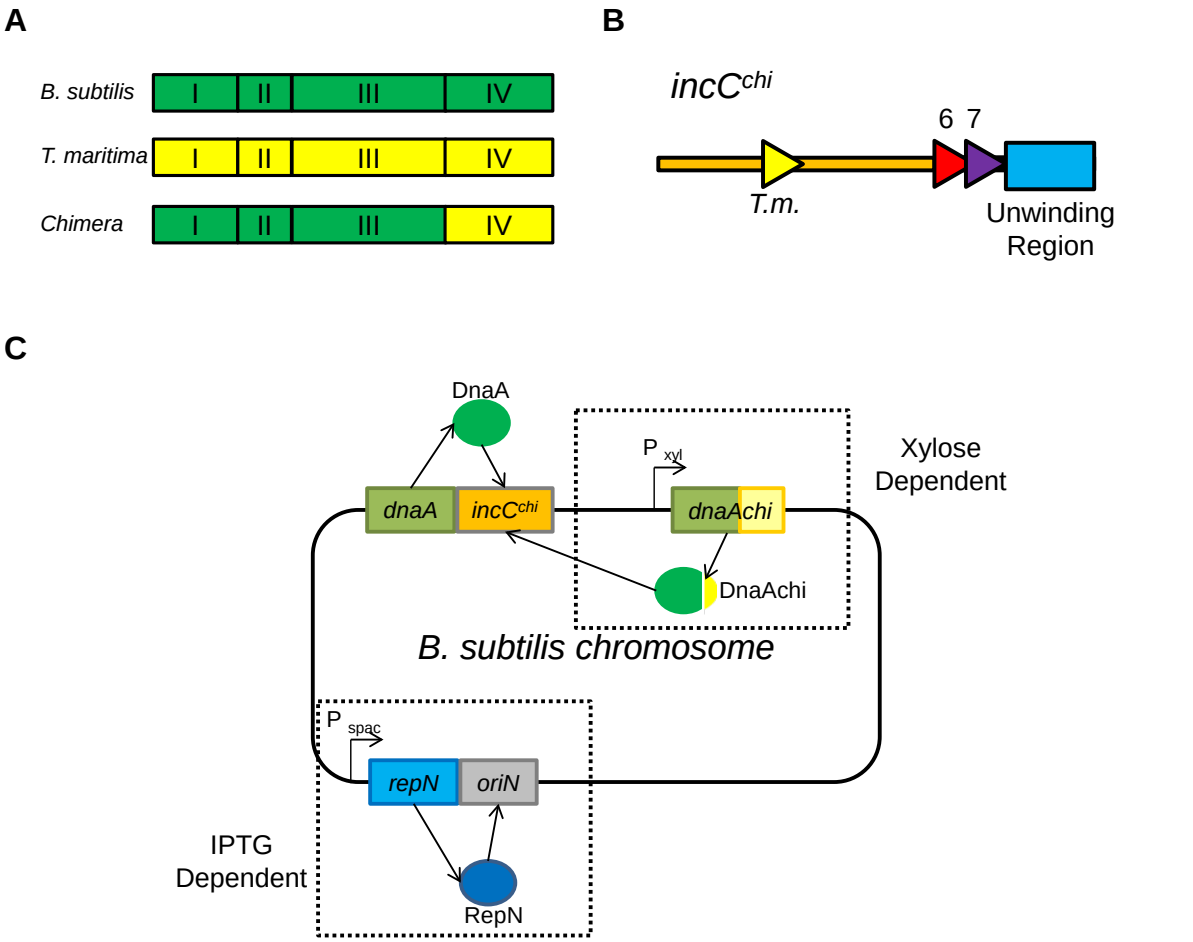


Figure 3.8. A system for investigating DnaA activities at the *incC* upstream subregion. (A) Schematic of the DnaA chimera. **(B)** Schematic of the *incC* chimera hybrid origin. Synthetic DNA is shown in orange. **(C)** Diagrammatic representation of the strain used for investigating the activities required at the upstream subregion of *incC*. **(D)** Analysis of the ability of the chimeric strain to support the growth of strains carrying either a wild type, artificial minimal or *incC*^{chi} hybrid origin in the presence and absence of xylose and IPTG. The result is shown following 48 hours incubation at 37°C. I (HM1108), II (HM1683), III (HM1694), IV (TR241).

3.6. The DnaA being delivered from the *incC* upstream subregion requires residues implicated in filament formation and single-stranded DNA binding activities

The chimeric system established in Section 3.5 was utilised to determine which activities are required by DnaA bound to the upstream *incC* DnaA-box. Sections 3.2 and 3.4 determined which residues implicated in filament formation or single-stranded DNA binding were essential *in vivo*. A selection of these lethal substitutions was introduced into the inducible *dnaA* chimera within the strain carrying the hybrid *incC*^{chi} origin. The results show that, in contrast to the wild-type chimera, all of the protein variants displayed severe growth defects (Figure 3.9.A).

To confirm the DnaA^{chi} variants were being properly expressed, the mutant genes were transformed into a *ΔdnaA* strain carrying a constitutive *oriN*. This allowed for the specific detection of the chimeric DnaA proteins via immunoblot in the presence and absence of xylose. The result showed that all DnaA^{chi} variants were stably expressed upon induction with xylose (Figure 3.9.B).

The results using the DnaA chimera system indicates that the DnaA protein being delivered to the site of unwinding from the upstream subregion requires essential amino acid residues previously demonstrated to be involved in both filament assembly and single-stranded DNA binding. This is consistent with the notion that DnaA is indeed being delivered to the site of unwinding from the upstream *incC* box where it actively participates in the strand opening reaction.

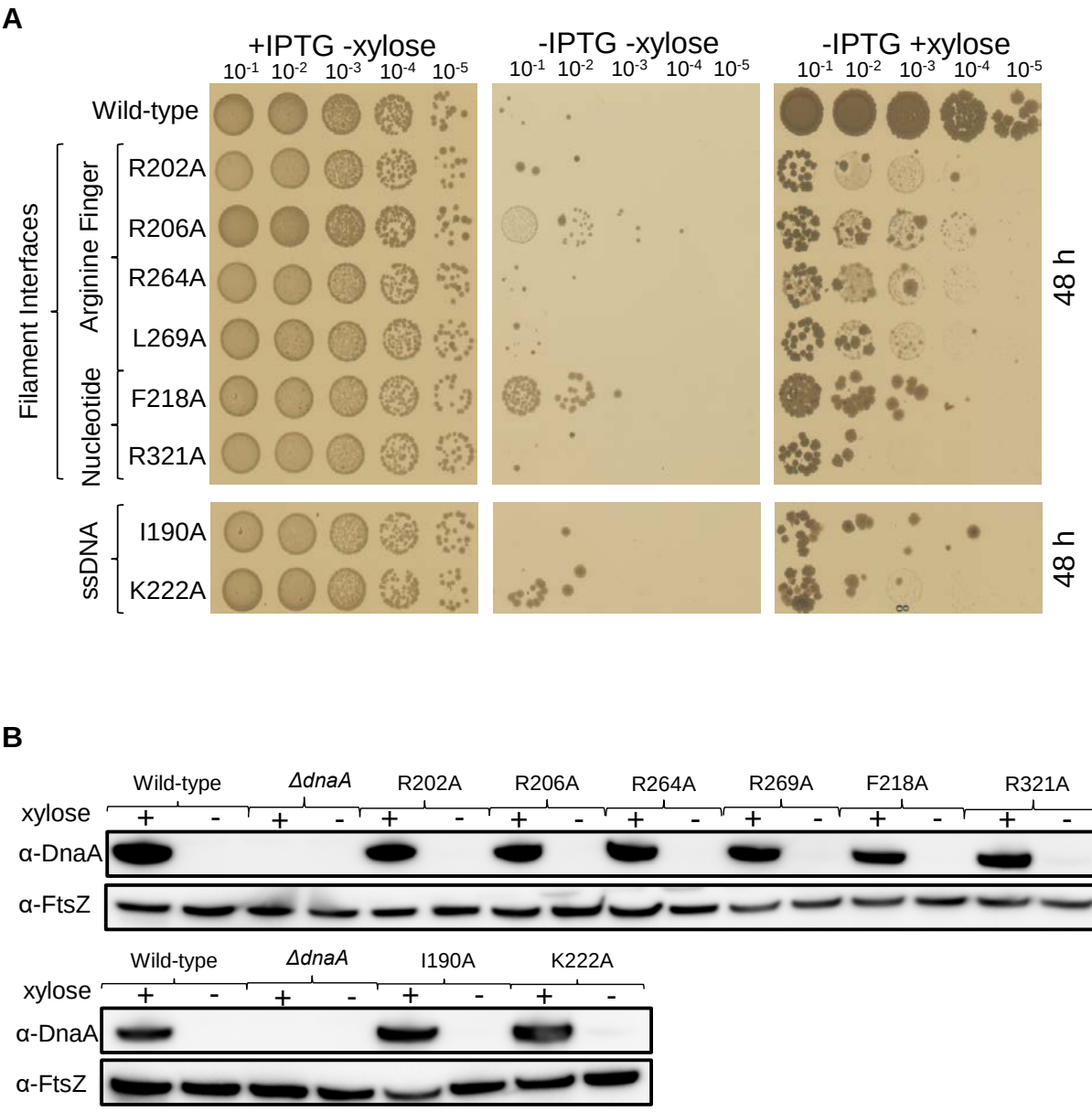


Figure 3.9. DnaA binding to the upstream DnaA-box in a minimal *incC* requires residues implicated in filament formation and ssDNA binding. (A) *In vivo* analysis of chimeric DnaA protein variants binding to the upstream DnaA-box. The result is shown following 48 hours incubation at 37°C. Wild-type (TR241), R202A (TR480), R206A (TR481), R264A (TR313), L269A (TR483), F218A (TR486), R321A (TR488), I190A (TR244), K222A (TR262). **(B)** Immunoblot analysis of the chimeric DnaA mutant variants in a *ΔdnaA* strain using the tubulin homolog FtsZ as a loading control and a strain lacking *dnaA* as an antibody control.

Wild-type(DS68), *ΔdnaA* (HM1424), R202A (DS61), R206A (DS62), R264A (DS60), L269A (DS64), F218A (DS65), R321A (DS66), I190A (DS57), K222A (DS58).

Chapter 3 – Discussion

The role of the upstream *incC* subregion in *Bacillus subtilis* origin function

It has been hypothesised that the DnaA-box distal to the DnaA-trios functions to increase the local concentration of protein, via a DNA loop, at the site of unwinding to assist in separating the DNA strands. If this is the case then the protein being delivered from the upstream subregion should require oligomerisation and ssDNA binding activities. This hypothesis was investigated by use of a genetic system involving a DnaA chimera that enabled the activities specifically required by the DnaA protein binding to the distal box to be determined.

Introducing lethal substitutions to amino acid residues implicated in filament formation or single-stranded DNA binding into the DnaA protein functioning from the upstream subregion resulted in severe growth defects. These results suggest that the DnaA specifically binding to the distal DnaA-box requires filament formation and single-stranded DNA binding activities. This supports the hypothesis that the distal DnaA-box is delivering DnaA to the site of unwinding, where it can assist in unwinding the DNA strand and opening the origin. As single-stranded DNA binding is a required activity, and not just filament formation, the results further support the notion that the mechanism for strand separation is more likely to be the DnaA filament stretching the DnaA-trios to induce unwinding rather than another mechanism such as super helical strain.

How exactly, and in what conformation, the DnaA is delivered by the distal DnaA-box to the site of unwinding remains unknown. It has been shown that the DnaA-box must be a minimum number of base pairs upstream of the boxes proximal to the DnaA-trios suggesting a loop forms within *incC* delivering protein to the site of unwinding (Richardson *et al., 2019*). There are several potential possibilities for what is being delivered by the *incC* loop to the DnaA-trios and these are highlighted in Figure 3.10. It is possible a single DnaA monomer bound to the DnaA-box is being delivered by a loop to directly engage either a DnaA-trio, or the early stages of a DnaA-filament forming from the proximal DnaA-boxes. Alternatively the DnaA bound to the distal DnaA-box could be promoting oligomerisation around the box prior to looping and so a longer DnaA oligomer could be being delivered to engage multiple DnaA-trios (Richardson *et al.*, 2019). Also remaining to be determined is whether unwinding of the duplex occurs after the DnaA protein is delivered via the loop or whether the extra

DnaA proteins assist an already started process to extend or maintain the open complex.

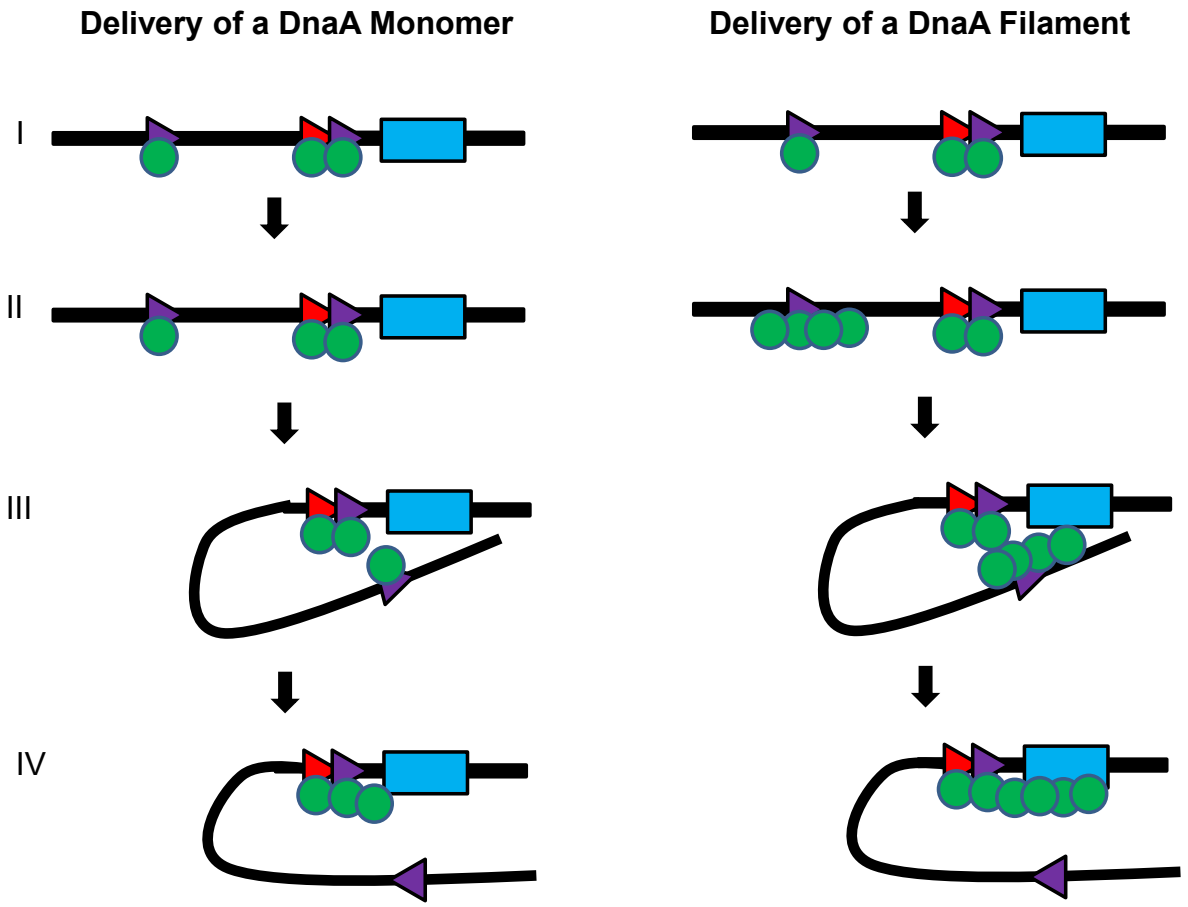


Figure 3.10. Models for the delivery of DnaA, to the DnaA-trios, from the distal *incC* subregion. Two models for how either a DnaA monomer or a DnaA filament is delivered to the DnaA-trios, and the site of unwinding, from the upstream *incC* subregion during initiation of DNA replication in *Bacillus subtilis*. Only the minimal *incC* architecture is shown for simplicity. DnaA-boxes are purple except the consensus DnaA-box (box #6) which is in red. DnaA is green and the DnaA-trios are blue.

Filament formation and single-stranded DNA binding in *Bacillus subtilis*

To render the DnaA protein non-functional for certain activities, this first required determining which of the amino acid residues previously implicated in these activities were essential *in vivo*.

Firstly, the investigation into the residues implicated in filament formation indicated that the AAA+/AAA+ filament interface is critical for a fully functional protein and therefore physiologically relevant. However, viability was not completely lost with a small number of residues suggesting some residues, while required for optimal growth and protein function are not absolutely essential. Interestingly, these non-essential residues were more common on the nucleotide binding interface of the inter-protomer interactions than the arginine finger interface. This result suggests the residues of the arginine finger interface are potentially more important to the interaction than those forming the opposite interface. DnaA oligomerisation occurs in a head-to-tail manner with the arginine finger of one AAA+ domain recognising the ATP molecule bound to another (Kawakami *et al.*, 2005). The arginine finger therefore guides oligomerisation, with the residues forming this interface presumably assisting in further guiding and stabilising the initial interactions potentially explaining the relative importance of the arginine finger interface over the nucleotide binding interface. The helicase loader (DnaI in *B. subtilis*) is proposed to interact with the AAA+ domain of DnaA using it as a docking site to bind to the end of a DnaA filament (Section 1.3.5) (Mott *et al.*, 2008). It is possible that the arginine finger interface is the part of the filament utilised by the loader and the residues located here are involved in this interaction providing another potential explanation for their relative importance.

Single-stranded DNA binding is another essential activity of DnaA. Several of the residues implicated in this activity proved to be required for DnaA function. Unexpectedly, the glutamine at position 224 could be substituted for alanine with no apparent loss of viability or fitness suggesting this residue isn't required for protein function *in vivo*. The homologous position in *T. maritima* has been shown to be both essential *in vivo* and defective for single-stranded DNA binding (Ozaki *et al.*, 2008). These results suggest that single-stranded DNA binding by *B. subtilis* DnaA is potentially slightly different than the current proposal suggests, with it also being possible there is a level of redundancy in the system with only certain residues being absolutely required for binding. There could also be other amino acids involved that

have yet to be implicated, and it is possible that the specific DNA sequence used for crystallisation versus the DnaA-trios could account for the observed differences. The results suggest that for *B. subtilis* at least, the mechanism of DnaA binding DnaA-trios is not yet fully understood.

The DBD/AAA+ interaction in *Bacillus subtilis*

During the investigation into the residues implicated in filament formation the physiological relevance of the proposed DBD/AAA+ interaction was also determined. Cells could tolerate alanine substitutions to any of the residues within the proposed interface, with the majority of substitutions growing like wild-type. Substituting the proposed residues of the DBD/AAA+ interfaces with more dramatic amino acids was much less tolerable for the cells, with several substitutions causing growth defects.

Two of the lethal substitutions involved introducing negatively charged residues into the DBD. While these could be involved in a DBD/AAA+ interaction, it is necessary to note that these two positions (S401 and L402) form the start of the helix-turn-helix double-stranded DNA binding motif as re-outlined in Figure 3.11.A (Erzberger *et al.*, 2002). It seems likely that these substitutions would disrupt DnaA-box binding by repelling the negatively charged DNA backbone. Indeed, residues in positions homologous to these in *E. coli* have been shown, via a crystal structure of DnaA bound to a DnaA-box, to be making contacts with phosphate groups (Fujikawa *et al.*, 2003) (Figure 3.11.B).

Swapping the charge of two residues located in the AAA+ domain caused severe growth defects. Again these could be involved in a DBD/AAA+ interaction, however if they are not, then as highlighted in Figure 3.11.C, these residues would be exposed on the surface of the DnaA filament away from the AAA+ ssDNA binding site. As discussed in Section 1.3.5 it has been proposed from studies involving *A. aeolicus* that DnaA interacts directly with the helicase loader protein (DnaI in *B. subtilis*) (Mott *et al.*, 2008). The proposed mechanism of interaction is that the AAA+ of the loader docks against the AAA+ domain of DnaA. This interaction is ATP dependent so is proposed to occur between filaments of both proteins. The AAA+ residues identified here could be involved in an interaction between DnaA and DnaI with the charge swaps resulting in some form of repulsion, preventing the interaction, and explaining the lethal phenotype.

Since this investigation was purely *in vivo,* it did not differentiate between whether the growth defects are due to a disruption of a DBD/AAA+ interaction, an effect on duplex DNA binding or an interaction with the helicase loader. Fully differentiating between these could be achieved through investigation *in vitro*. Once purified the DnaA variant proteins ability to bind duplex DNA/DnaA-boxes could be investigated through BMOE crosslinking assays to capture oligomeric filaments forming on the DNA (Chapter 4). Pull down assays involving the AAA+ mutant variants could determine if the residues identified are required for a DnaA-DnaI interaction.

A final consideration is that the proposed DBD/AAA+ interaction could be conditional and not required under the normal growth conditions utilised in this investigation. In such a scenario the residues required for the interaction would only become essential under the adverse conditions. Investigating the requirement of the proposed amino acid residues under different temperatures or stresses (such as salts or antibiotics) could help identify if the DBD/AAA+ interface is conditionally physiologically relevant.

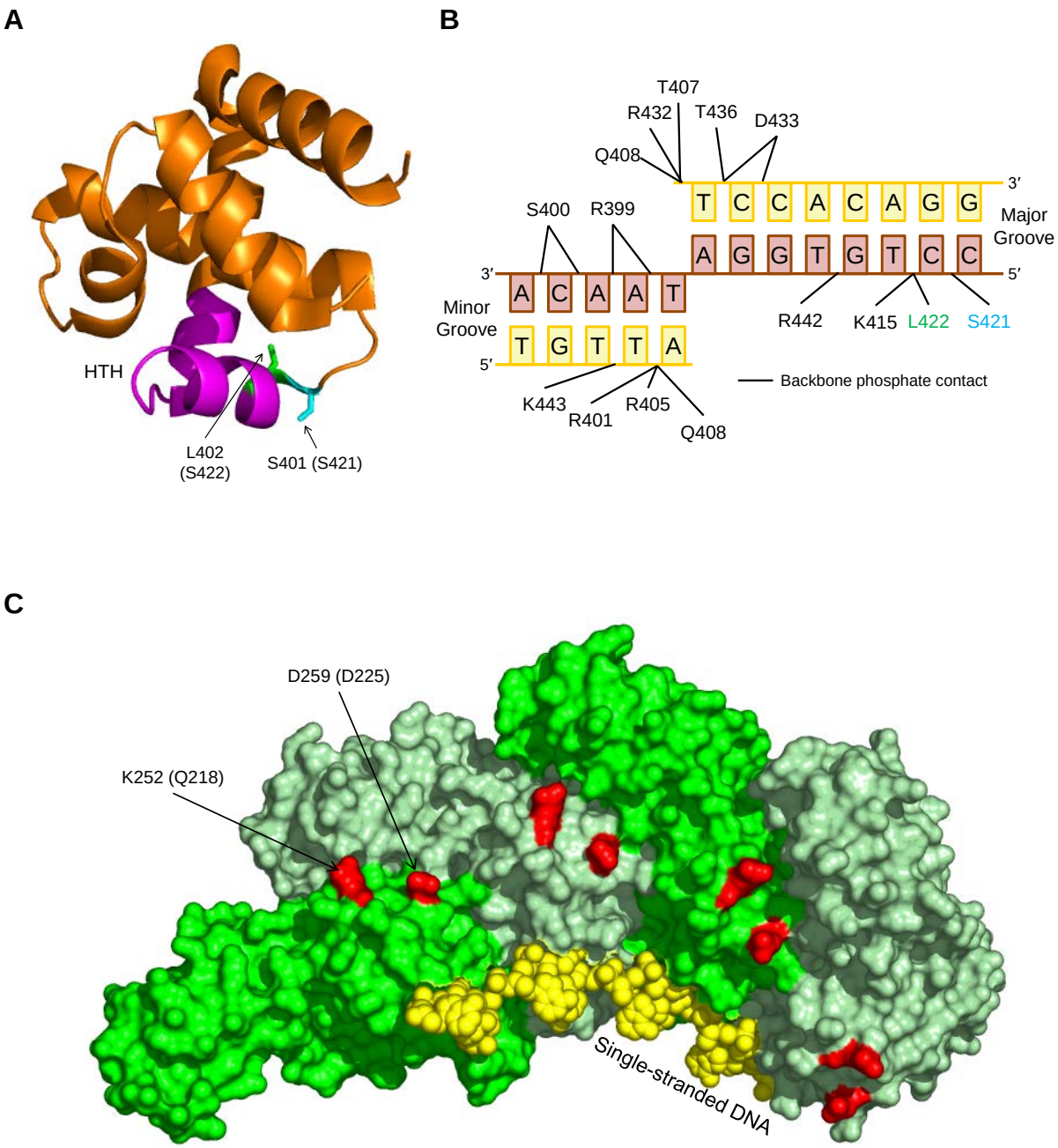


Figure 3.11. Functions for the DBD/AAA+ interaction interface residues. (A) Structure of the DBD of *E. coli* (PDB ID 1J1V) highlighting the location of the helix-turn-helix motif (HTH) and the position of the conserved serine and leucine residues labelled according to *B. subtilis* numbering with *E. coli* in parentheses. **(B)** Schematic highlighting the phosphate backbone contacts by residues within DnaA domain IV with a DnaA-box adapted from Fujikawa *et al.*, 2003. The conserved serine and leucine residues highlighted in A are highlighted. **(C)** Structure of an *A. aeolicus* DnaA filament bound to ssDNA (yellow) showing the AAA+ domains of four protomers in alternating shades of green. Residues of interest are highlighted red and labelled according to *B. subtilis* numbering with *A. aeolicus* in parentheses.

Chapter 4

The DnaA initiator specific motif (ISM) and specificity for the DnaA-trios in *Bacillus subtilis*

Chapter 4 – Introduction

The current model(s) for how the origin of replication is unwound by DnaA has been discussed in Chapters 1 and 3. The proposed sequence of events in *B. subtilis* is that DnaA monomers recognise *oriC* by binding to the double-stranded DnaA-boxes; subsequently DnaA-ATP filaments assemble and are loaded onto one strand of the dsDNA where they promote DNA strand separation by engaging and stretching the DNA. The essential repeating trinucleotide motif termed the DnaA-trio (Section 1.2.3) has been identified to promote and stabilise a DnaA filament onto a specific single DNA strand, with each motif interacting with a single DnaA protein (Duderstadt and Berger, 2013; Richardson *et al.*, 2016).

The molecular mechanism underpinning specificity for the DnaA-box sequence has been appreciated for several decades (Section 1.3.2) (Fuller *et al.*, 1984). However, the mechanisms behind the specific recognition of the DnaA-trio sequence by DnaA remains unknown. As outlined in Section 1.3.4 it has previously been established that DnaA binds to ssDNA, non-specifically, using two pairs of alpha helices, α3-α4 (the initiator specific motif) and α5-α6 within the AAA+ domain (Duderstadt *et al.*, 2011). Binding to the DnaA-trio motif was shown to require the ssDNA binding residue Ile190 located within the AAA+ initator speicifc motif (ISM) of DnaA *in vitro* (Richardson *et al.*, 2016).

The ISM of AAA+ domains from various initiator proteins are used for binding DNA across the domains of life. These include DnaA (Duderstadt *et al.*, 2011), the bacterial helicase loader (Arias-Palomo *et al.*, 2019), the ORC1 initiator protein from archaea (Dueber *et al.*, 2007) and the ORC subunits of the eukaryotes *Saccharomyces cerevisiae* (Li *et al.*, 2018) and *Drosophila melanogaster* (Bleichert *et al.*, 2018). Interestingly, while most of the DNA binding mechanisms involve amino acid interactions with the phosphate backbone, some of the mechanisms employed involve base specific contacts. For example the ISM of ORC1 from the archaea *A. pernix* inserts into the minor DNA groove and residues in the loop and helix of the motif make

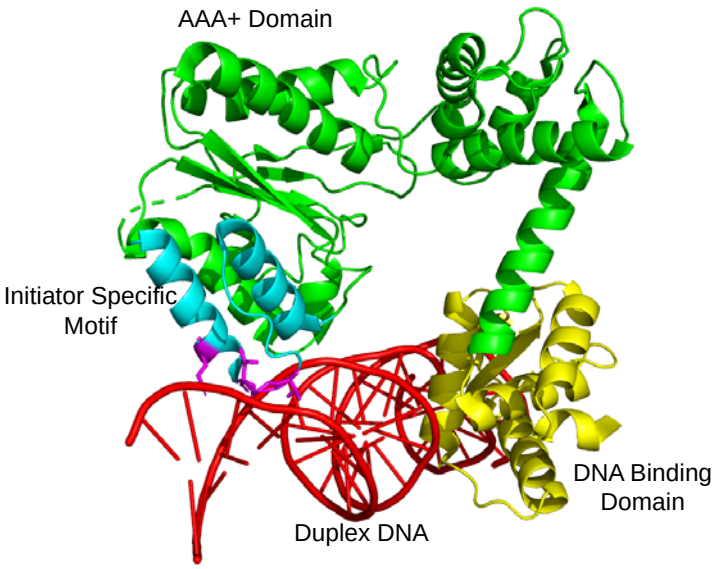
direct and indirect base and backbone interactions (Figure 4.1.A) (Dueber *et al.*, 2007; Gaudier *et al.*, 2007).

Figure 4.1.B shows the loop of the ISM of *B. subtilis* contains an arginine (R) amino acid residue, an asparagine (N) and a lysine (K). These are residues commonly used by DNA binding proteins to form van der Waals contacts and hydrogen bonds with the phosphate backbone and bases of DNA (Luscombe *et al.*, 2001; Sathyapriya *et al.*, 2008). I hypothesised that these residues at the tip of the ISM might contribute the specific recognition of the DnaA-trio motif by DnaA.

As mentioned the ISM has already been proposed to be involved in non-specific ssDNA binding. This proposal is based on the findings from a structure of *A. aeolicus* DnaA bound to a synthetic DNA substrate (poly-A_{12}). The residues implicated from this structure were investigated in Chapter 3 where it was concluded that, for *B. subtilis* at least, the precise single-stranded DNA binding mechanism of DnaA likely occurs slightly differently to that proposed. As such the investigation into the residues of the ISM may also help create a clearer picture of the general mechanism of single-stranded DNA binding by *B. subtilis* DnaA.

This chapter shows the investigation into the structure and function of the *B. subtilis* initiator specific motif and the potential role some of the amino acid residues may play in ssDNA binding, origin unwinding and DnaA-trio specificity.

A



B

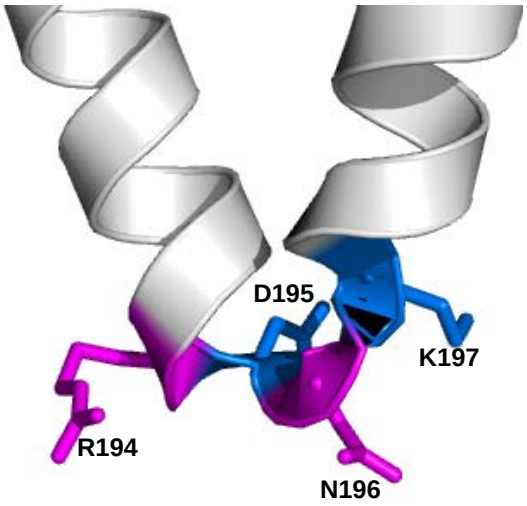


Figure 4.1. The loop of the initiator specific motif contains residues which could be contacting DNA. (A) Crystal structure of the ORC1 initiator protein from the archaea *A. pernix* binding double-stranded DNA (PDB ID 2V1U). The DNA is shown in red, the ORC1 DNA binding domain in yellow, the AAA+ domain in green with the ISM shown in cyan and the ISM residues which contact DNA in magenta. **(B)** Model of the loop of the ISM of *Bacillus subtilis* with residues highlighted in alternating colours. Model created by mutating the residues of the *A. aeolicus* crystal structure (PDB ID 2HCB) in PyMOL.

Chapter 4 – Results

4.1. Alanine substitution of several ISM residues results in growth defects

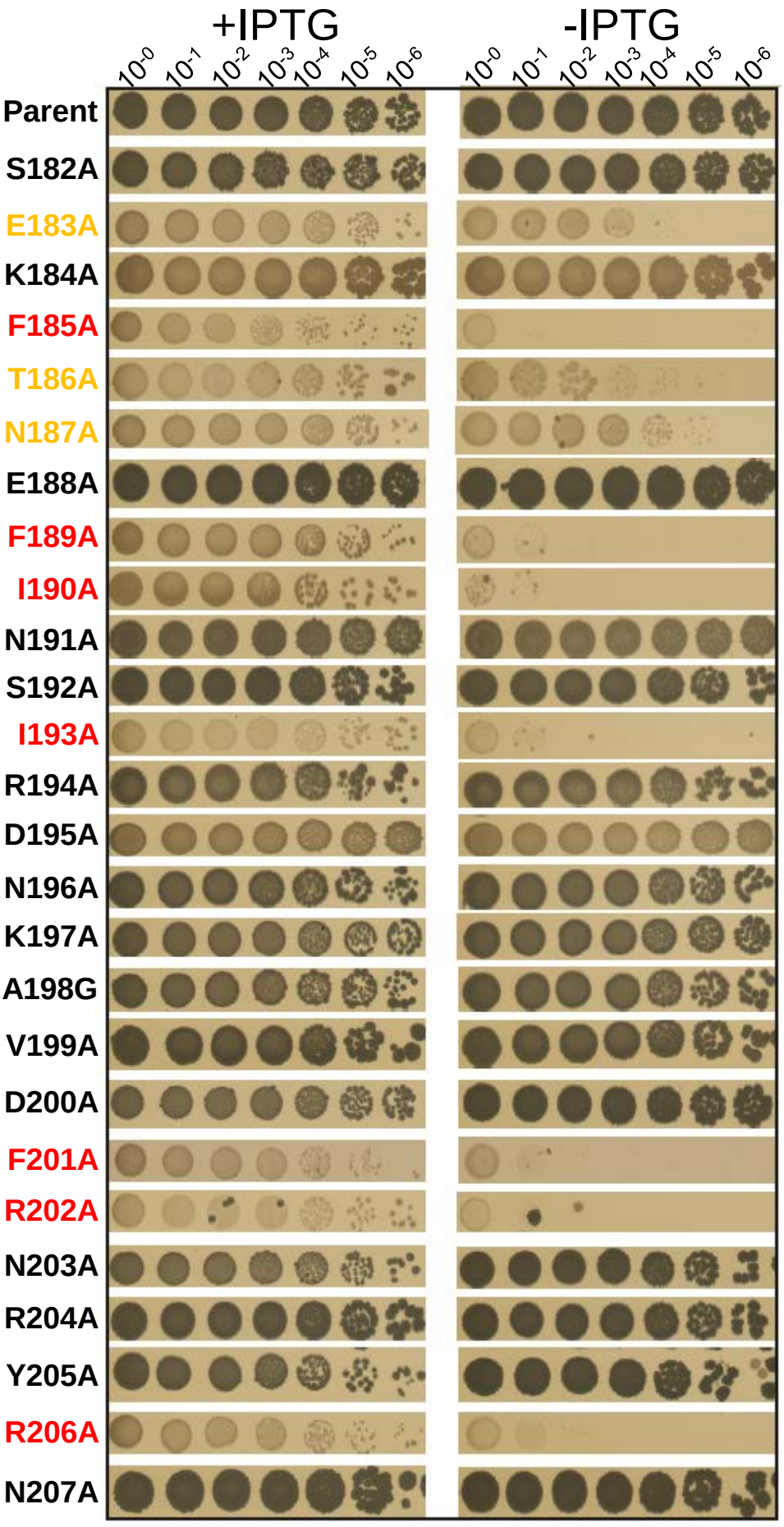
4.1.1. Residues of the Bacillus subtilis DnaA initiator specific motif are essential

To determine which of the residues of the initiator specific motif of *B. subtilis* DnaA are required for a functional protein *in vivo,* all 26 residues were substituted individually for alanine via site directed mutagenesis (Section 2.5.1) within the inducible second origin strain (*oriN*) outlined in Chapter 3. Note that the native residue at position 198 is already an alanine so for completeness this was substituted for a glycine (G).

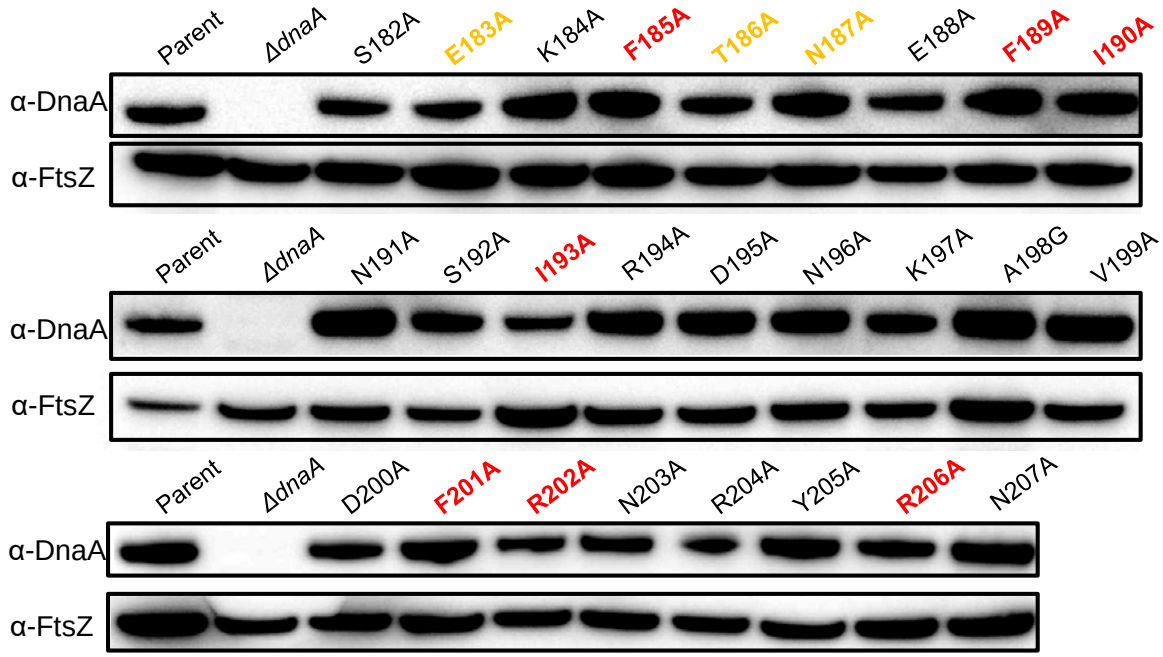
The results of the alanine scan revealed that 16 of the substitutions had no effect on colony growth or viability and so these residues were classified as non-essential (Figure 4.2.A, labelled black). In contrast 7 of the substitutions resulted in a complete loss of viability suggesting these residues are required for a functional protein *in vivo* and as such these residues were classified as essential (F185, F189, I190, I193, F201, R202, R206; Figure 4.2 labelled red). The remaining residues displayed a slow growing phenotype, where the cells are still viable in the absence of IPTG but grow at a noticeably slower rate to wild-type (E183, T186, N187; Figure 4.2.A labelled orange).

Immunoblots were performed (Figure 4.2.B) and indicated all the mutant proteins were being expressed at levels similar to wild type. The residues determined by the alanine scan to be essential or intermediate were mapped onto the crystal structure of the ISM (from the homologous *A. aeolicus*) as shown in Figure 4.2.C. Showing the surface representation of the structure and rotating it 180° revealed that all the residues that had any effect on growth are located on one face of the ISM with the majority appearing to be surface exposed (Figure 4.2.D). The essential residues also appear to form two major clusters, one in the top right (F201, R202 and R206) and the other the bottom left (F189, I190 and I193).

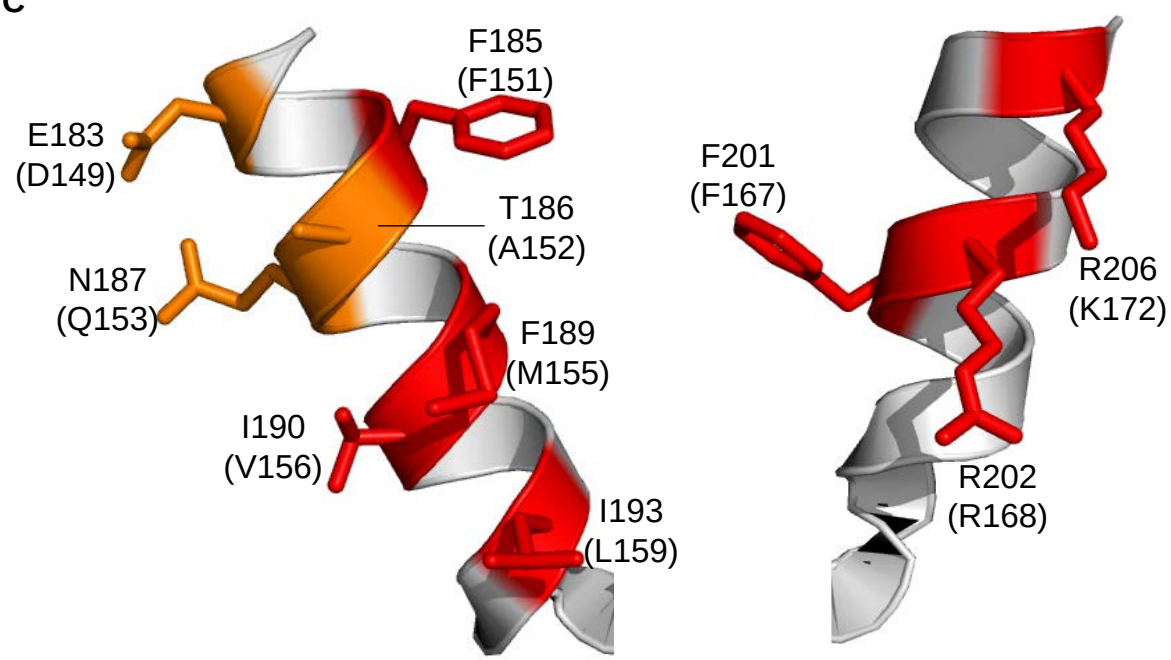
A


+IPTG
-IPTG
10-0 10-1 10-2 10-3 10-4 10-5 10-6
Parent
S182A
E183A
K184A
F185A
T186A
N187A
E188A
F189A
I190A
N191A
S192A
I193A
R194A
D195A
N196A
K197A
A198G
V199A
D200A
F201A
R202A
N203A
R204A
Y205A
R206A
N207A
72 h

B


Parent
ΔdnaA
S182A
E183A
K184A
F185A
T186A
N187A
E188A
F189A
I190A
α-DnaA
α-FtsZ
Parent
ΔdnaA
N191A
S192A
I193A
R194A
D195A
N196A
K197A
A198G
V199A
α-DnaA
α-FtsZ
Parent
ΔdnaA
D200A
F201A
R202A
N203A
R204A
Y205A
R206A
N207A
α-DnaA
α-FtsZ

C


F185
(F151)
E183
(D149)
T186
(A152)
N187
(Q153)
F189
(M155)
I190
(V156)
I193
(L159)
F201
(F167)
R206
(K172)
R202
(R168)

D

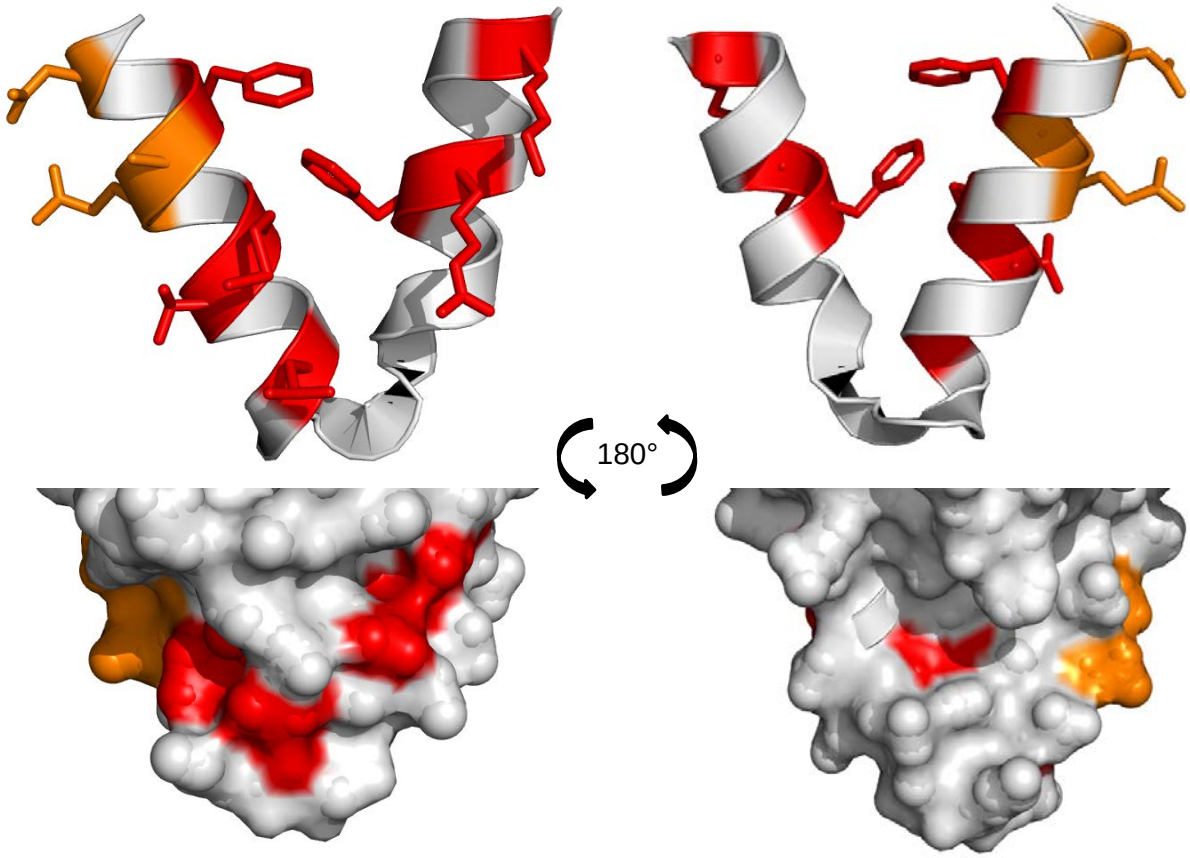


Figure 4.2. Essentiality of the residues of the *B. subtilis* DnaA initiator specific motif. (A) Analysis of *B. subtilis* DnaA substitution mutants in the *oriN*/RepN strain. Growth is shown after 72 hours incubation at 37°C. Residues highlighted red are essential while those highlighted orange are slow growing. **(B)** Immunoblot analysis of the DnaA substitution mutants with the tubulin homolog FtsZ as a loading control and a strain lacking *dnaA* as an antibody control. Highlighting is the same as A. **(C)** Results from A mapped onto the crystal structure of the ISM from *A. aeolicus* (PDB ID 2HCB). Residues are labelled according to the *B. subtilis* numbering, with *A. aeolicus* shown in parentheses. Only residues showing growth defects are highlighted the same as A. **(D)** Structure from C shown as cartoon and surface representation in the same orientation and after a 180° rotation.

Parent (HM1108), S182A (DS40), E183A (DS6), K184A (DS36), F185A (DS7), T186A (DS8), N187A (DS23), E188A (DS15), F189A (DS19), I190A (DS18), N191A (DS37), S192A (DS5), I193A (DS20), R194A (DS28), D195A (DS30), N196A (DS29), K197A (DS16), A198G (DS11), V199A (DS14), D200A (DS10), F201A (DS9), R202A (DS21), N203A (DS3), R204A (DS4), Y205 (DS24), R206A (DS22), N207A (DS13), *ΔdnaA* (HM1424).

4.1.2. Residues forming the loop of the ISM are not required for a functional protein in vivo

As discussed in the introduction to this chapter, part of the original rational for investigating the initiator specific motif was based upon the observation that the loop of the motif in the homologous ORC1 protein directly contacts the DNA during double-stranded DNA binding (Figure 4.1.A). This region of the *B. subtilis* DnaA ISM contains charged and polar residues commonly involved in DNA binding. The results of the alanine scan surprisingly revealed that none of the residues located here (K194, D195, N196, R197), re-highlighted in Figure 4.3.A, had any effect on viability or growth when substituted.

To fully rule this region out as being important for a functional DnaA *in vivo,* a series of double mutations were introduced, substituting two residues for alanine simultaneously using the same approach as for creating the single substitutions. The result of these substitutions, and an immunoblot, revealed that the substitutions had no effect on colony growth or viability, and as expected the mutant proteins were expressed at normal levels (Figure 4.3.B and 4.3.C respectively).

While these results do not completely rule out this region as potentially being involved in DNA binding, or stabilising any protein-DNA interaction, it does rule out that any single amino acid residue located here as being required for protein function *in vivo*. As single-strand DNA binding, and presumably the specific recognition of DnaA-trios, is essential, this region was not investigated further for a role in DnaA-trio binding.

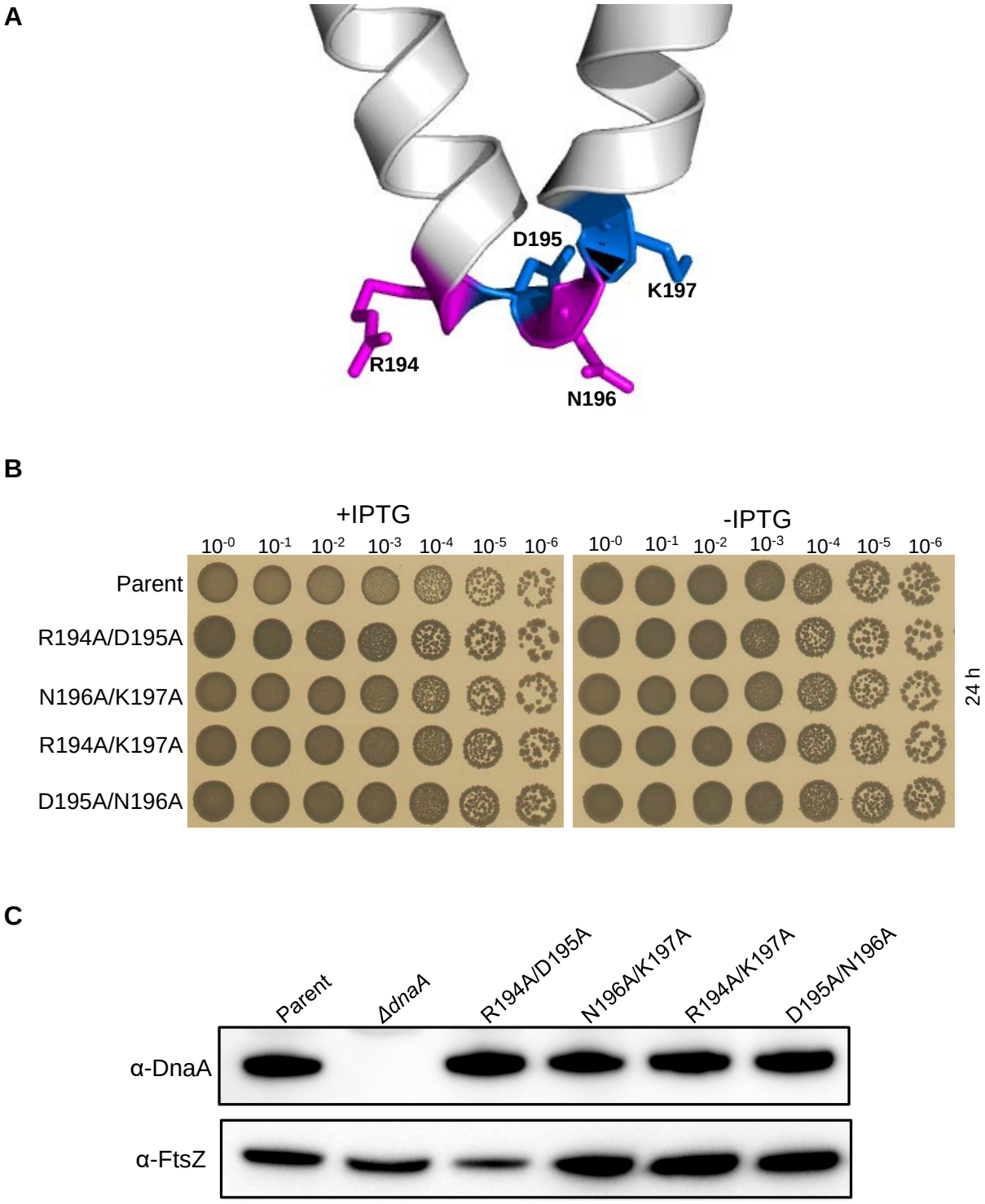


Figure 4.3. Essentiality of the residues which form the loop of the ISM. (A) Model of the residues that form the loop of the *B. subtilis* ISM. Model the same as figure 4.1.B. **(B)** Analysis of *B. subtilis* DnaA substitution mutants using the *oriN* strain, growth is shown after 24 hours incubation at 37°C. **(C)** Immunoblot analysis of the DnaA substitution mutants with the tubulin homolog FtsZ as a loading control and a strain lacking *dnaA* as an antibody control.

Parent (HM1108), R194A/D195A (DS36), N196A/K197A (DS37), R194A/K197A (DS38), D195A/N196A (DS39), *ΔdnaA* (HM1424).

4.1.3. Residues of the initiator specific motif are required for optimal cell growth

As discussed in section 4.1.1, a subset of the residues of the ISM was classified as having intermediate defects on growth. To investigate these residues further they were grown under altered variables to determine if the residues were required for growth in certain conditions. Prior to investigation the mutant strains were re-sequenced to ensure the correct mutation was still present and no further mutations had occurred.

As shown in Figure 4.4.A the substitutions still present an intermediate phenotype, however the E183A and N187A mutant strains appear to be growing better than the T186A mutation. This slow growing phenotype is also seen in liquid culture where the strains all grow at a comparable rate regardless of whether the second origin is switched on or off (+ or – RepN) (Figure 4.4.B). Here, while the N187A variant grows best again, the growth rate in liquid of the E183A and T186A mutations appear to be more comparable.

All three strains carrying the substituted versions of DnaA appear to be cold sensitive, becoming dependent upon RepN for growth at 20°C (Figure 4.4.C). This provides further indication that while these residues may not be absolutely essential for viability they are required for optimal growth. This was the same conclusion drawn about the residues E183 and N187 in section 3.2, where it was concluded both residues are probably involved in filament formation and, while survival was possible with just one of either residue, they are required for optimal growth. The results obtained here provide further support for this conclusion.

The ISM structure from Figure 4.2.C shows the T186 residue is located between the essential F185 residue and the probable filament formation involved N187 residue. Based on this the simplest explanation for the result of the T186A substitution is that this alteration has shifted the spatial positions of the flanking amino acids enough to affect their essential functions. As the E183 and N187 residues are most likely involved in filament formation and the T186 substitution is definitely viable, these residues were ruled out for further investigation into a role in DnaA-trio binding.

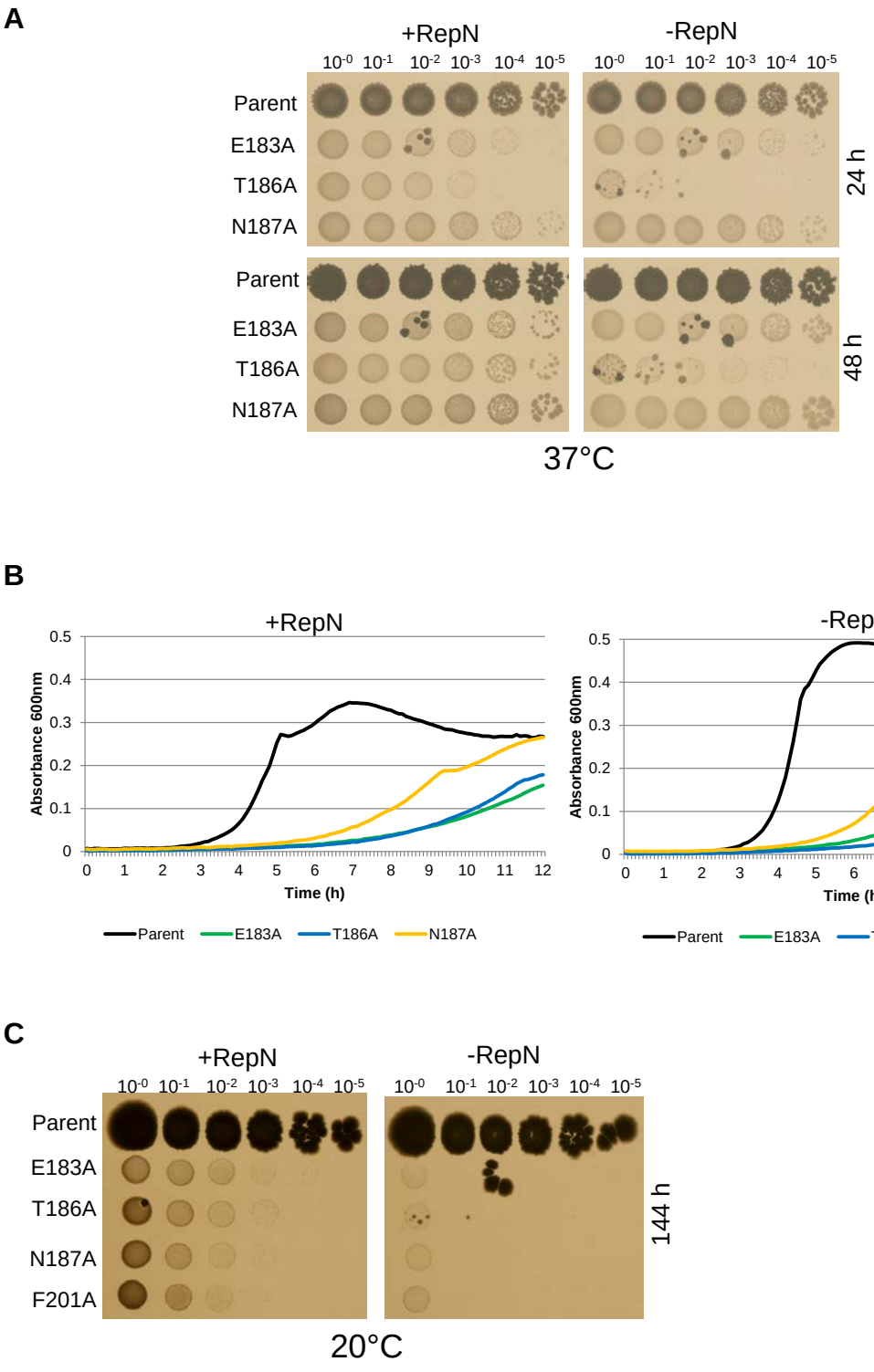


Figure 4.4. Reinvestigating the intermediate residues of the initiator specific motif. (A) Analysis of *B. subtilis* DnaA substitution mutants using the *oriN* strain, growth is shown after 24 and 48 hours incubation at 37°C. **(B)** Analysis of the growth of the same DnaA substitution mutants from A in liquid culture at 37°C. **(C)** Analysis of the growth of the same DnaA substitution mutants from A following 144 hours incubation at 20°C. F201A mutation was used as a control. The colour contrast for the image has been increased to show the results more clearly.

Parent (HM1108), E183A (DS6), T186A (DS8), N187A (DS23), F201A (DS9).

4.2. The *Bacillus subtilis* DnaA initiator specific motif contains a patch of essential residues of unknown function

Section 4.1 outlined the results of the *in vivo* investigation into the amino acid residues that constitute the ISM of DnaA in *B. subtilis*. The results revealed that the substitution of 10 residues (highlighted in Figure 4.2.C) had some effect on growth or viability. Elimination of the slow growing intermediate residues (section 4.1.3) has narrowed the number of residues to investigate for a role in protein function to the 7 that are essential.

The 7 essential amino acids have been highlighted on the ISM crystal structure from *A. aeolicus* in Figure 4.5.A. As the space fill of the structure shows, the two phenylalanine residues at positions 185 and 201 are buried within the protein, suggesting that these are likely involved in protein folding or maintaining the overall three-dimensional shape.

The two arginine residues at positions 202 and 206 have been discussed in previous chapters as having been implicated in filament formation. The work of previous studies provided strong evidence that the equivalent residues of R202 (Duderstadt *et al.*, 2010) and R206 (Ozaki *et al.*, 2012) in homologs are required for oligomerisation. In addition to this it has also been shown that the equivalent mutation to R202A in *A. aeolicus* is still able to bind ssDNA with high affinity (Duderstadt *et al.*, 2010). As such, it is highly likely these two residues are required for the same function in *B. subtilis* and unlikely to be involved in ssDNA binding.

Based on the information presented above the number of candidate residues most likely involved in DnaA-trio recognition is reduced by elimination to the 3 highlighted in Figure 4.5.B, F189, I190 and I193. The I190 residue has been discussed previously (Section 1.3.4) and has been shown to be involved in non-specific single-stranded DNA binding. Interestingly, the work performed here has identified two neighbouring surface exposed residues as both being required for viability *in vivo* (Section 4.1.1, Figure 4.5.B). Further, as shown in the space fill in Figure 4.5.B, the three residues appear to form a surface that could be a potential site for interacting with DNA. These three DnaA protein variants were selected for further analysis *in vitro* to investigate their biochemical properties.

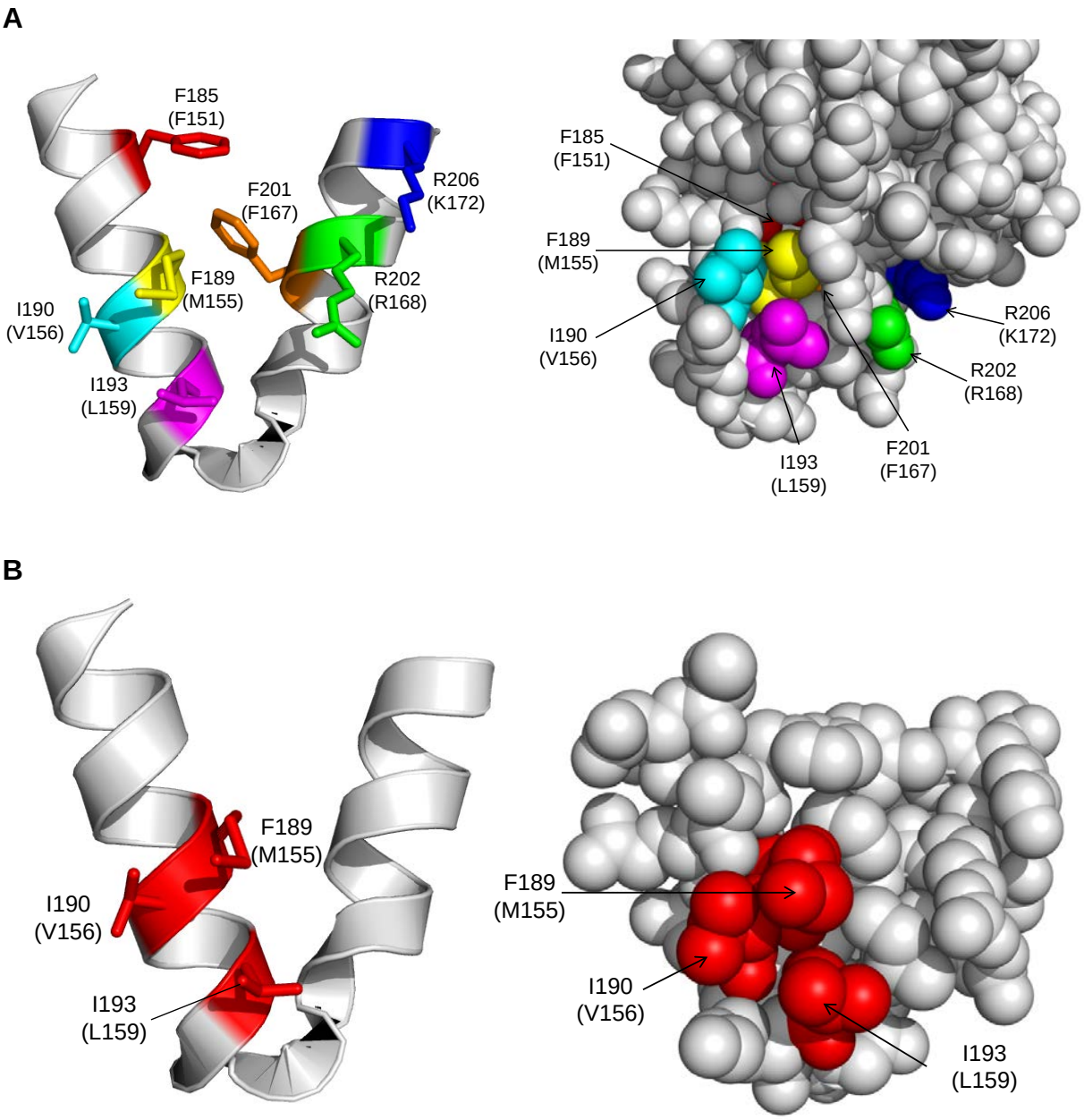


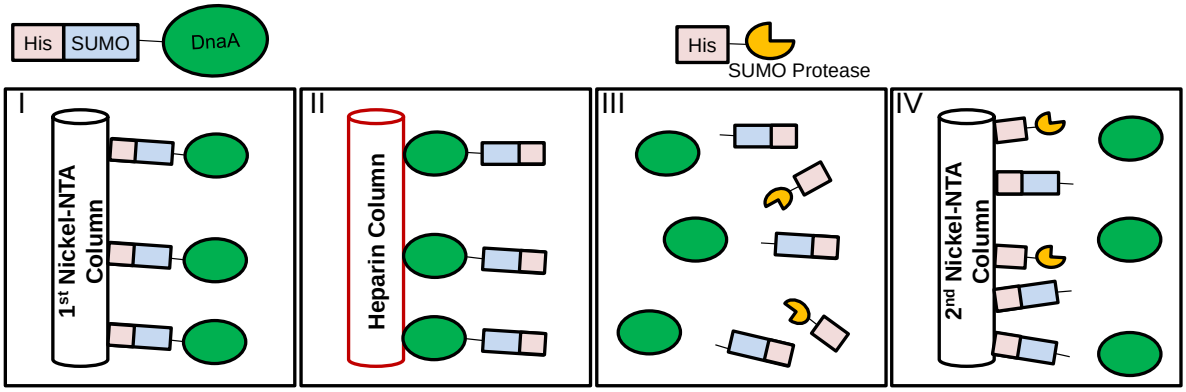
Figure 4.5. Essential residues of the *B. subtilis* DnaA initiator specific motif. (A) The residues of the *B. subtilis* DnaA ISM required for viability mapped onto the crystal structure from *A. aeolicus* (PDB ID 1L8Q) as a cartoon of the helix or a space fill of this region of the AAA+ domain. Residues are labelled according to the *B. subtilis* numbering, with *A. aeolicus* in parentheses. **(B)** The final ISM candidate residues for investigation *in vitro* highlighted in red, mapped onto the DnaA crystal structure as per A. The space fill is of only the ISM and not the wider AAA+ domain for simplicity.

4.3. Use of a His-SUMO tag allows for the specific purification of DnaA

To investigate the function of DnaA variants *in vitro* the proteins needed to be purified. Purification of DnaA was achieved using immobilised metal affinity chromatography (IMAC) and the full protocol is outlined in Section 2.10.

The proteins were genetically fused with a His_{14}-SUMO tag and expressed from a plasmid in BL21 *E. coli*. Whole cell lysates were passed through a HisTrap nickel column followed by Heparin HP column (Figure 4.6.A). Following elution from the heparin column the His_{14}-SUMO tag was cleaved using His_{14}-SUMO protease. The reaction was passed through a second HisTrap nickel column and untagged DnaA proteins were collected in the flow-through (Figure 4.6.A). An SDS-PAGE gel of samples from each step of the purification process for wild type DnaA is displayed in Figure 4.6.B to give an overview of the purification process. The purified DnaA proteins were then used for *in vitro* assays.

A



B

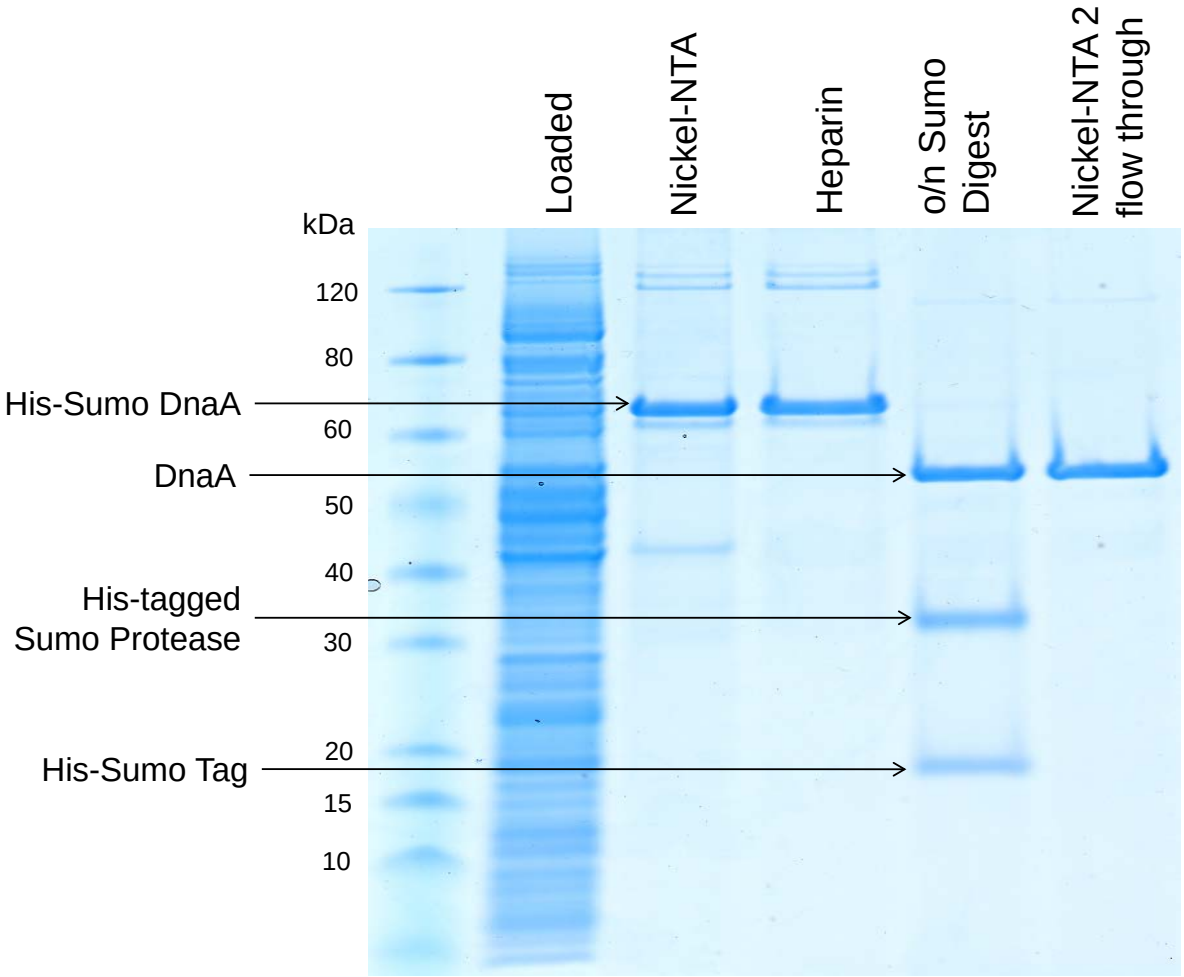


Figure 4.6. Purification of wild type DnaA. (A) Overview of steps for purifying *B. subtilis* DnaA. **(B)** Samples of the protein(s) collected after each stage of the purification process of wild type DnaA. Lane one shows a pre-stained protein ladder used to estimate molecular weight.

4.4. Two of the DnaA ISM variants are capable of forming ATP-dependent oligomers *in vitro*

As has been extensively discussed in previous chapters, DnaA forms ATP-dependent filaments. Therefore, the DnaA variants identified in Section 4.2 were investigated for their ability to oligomerise using an *in vitro* crosslinking assay as described in Section 2.11.1. A pair of cysteine residues has been introduced into the AAA+ domain substituting the amino acid positions 191 and 198. These are the only two cysteines present in the polypeptide and the protein is functional *in vivo*. This variant is referred to as $DnaA^{CC}$ (Scholefield *et al.*, 2012). When $DnaA^{CC}$ forms a filament the C191 residue of one protomer comes in close proximity to the C198 of the adjacent protein as highlighted in Figure 4.7.A. Use of the cysteine specific crosslinker BMOE (Bismaleimido ethane) enables the capturing of oligomeric species. The protein complexes can then be separated via SDS-PAGE and detected through an immunoblot using DnaA specific antibodies.

The candidate amino acid residues determined in Section 4.2 were substituted for alanine within the context of $DnaA^{CC}$ and the mutant proteins purified as described in Section 4.3. They were then used in the BMOE crosslinking assay to test their ability to form ATP-dependent filaments in solution (Figure 4.7.B).

The results show the captured $DnaA^{CC}$ complexes of increasing molecular weight running as a ladder with heavier complexes, and therefore larger oligomers, dependent on ATP. Both the I190A and I193A variants produced similar high molecular weight ATP-dependent complexes compared to the wild-type protein, indicating that they retain the ability to form filaments (Figure 4.7.B). In contrast the F189A variant was clearly defective forming ATP-dependent complexes (Figure 4.7.B). This result suggests the F189 residue is required for oligomerisation, rather than ssDNA binding. The position of the phenylalanine facing into the space between the two α-helices of the ISM (Figure 4.5) suggests that its essential function could be to maintain the structure of this motif.

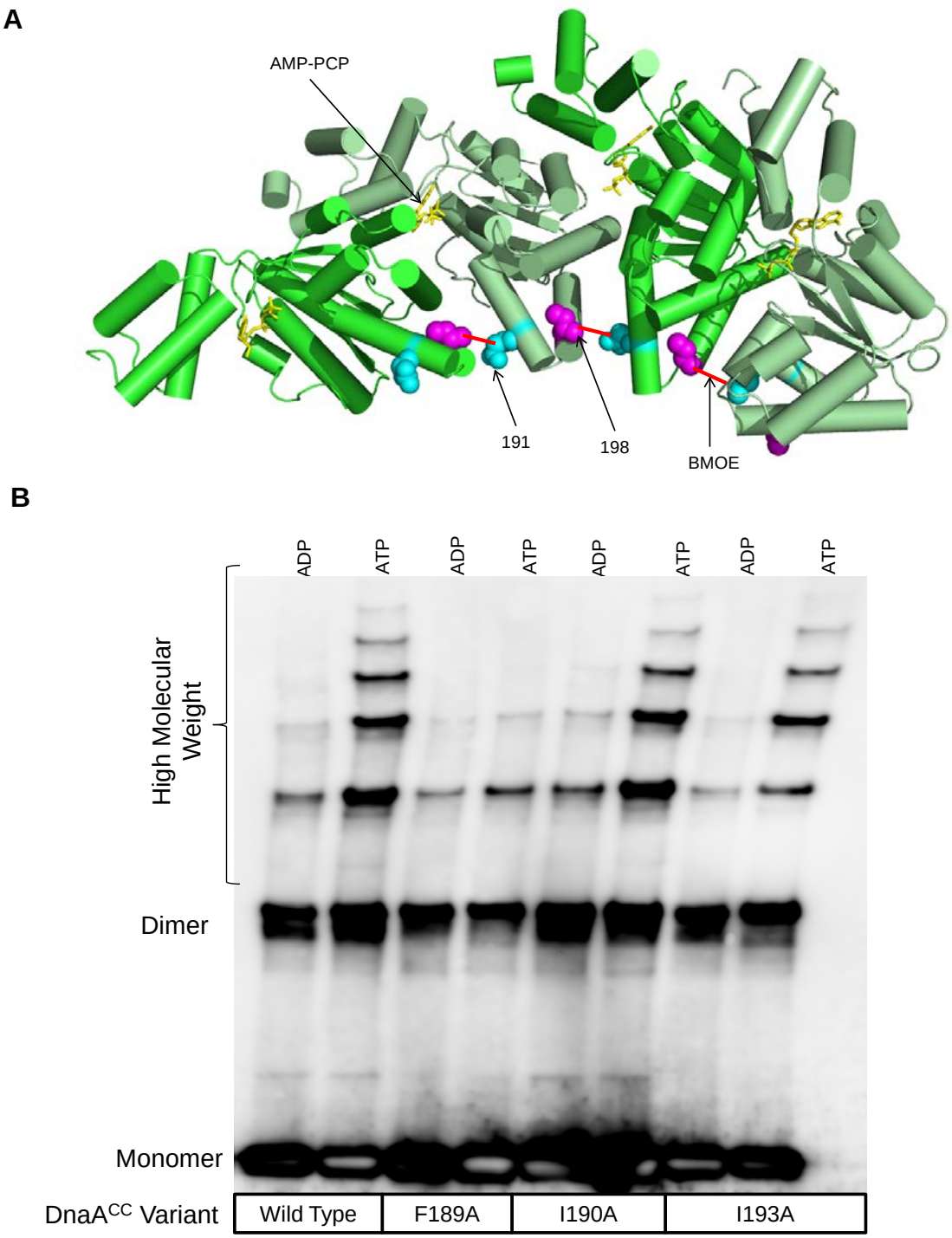


Figure 4.7. ATP-dependent oligomerisation of DnaA variants *in vitro*. (A) Crystal structure (PDB ID 2HCB) of an *A. aeolicus* ATP-dependent filament (AAA+ domains only) of DnaA highlighting the positions of the residues homologues to positions 191 (cyan) and 198 (magenta) in *B. subtilis* that have been cysteine substituted to create the DnaA^{CC} variant. A demonstration of how BMOE (red) captures oligomeric species is also shown. **(B)** DnaA oligomers captured forming in solution using cysteine specific crosslinking. DnaA complexes were resolved by SDS-PAGE and detected via western blot.

4.5. DnaA I190A and I193A are incapable of forming filaments on single-stranded DNA substrates

Since both the I190A and I193A variants are still capable of forming filaments, it was next assessed whether these proteins could form filaments specifically on DNA substrates containing DnaA-trios.

The assay employed again uses BMOE and the $DnaA^{CC}$ variants, but here the proteins were incubated with DNA substrates, referred to as DNA scaffolds, prior to crosslinking (described in Section 2.11.2). This enables the capturing of oligomeric species forming specifically on each scaffold, investigating both dsDNA binding activity on DnaA-boxes and ssDNA binding on DnaA-trios (Richardson *et al.*, 2016). The sequence of the origin region used for constructing the DNA scaffolds is shown in Figure 4.8.A. The DNA scaffolds were constructed from two oligonucleotides (listed in Table 2.5) annealed together. The wild-type scaffold contains the sequences of two DnaA-boxes (#6 and #7) which are double-stranded and a single-stranded tail carrying the DnaA-trio sequence (3′-GATGATAATGAAGATGAT-5′). Various mutant scaffolds were also constructed and the results of a crosslinking assay are displayed in Figure 4.8.B.

Here the wild-type $DnaA^{CC}$ protein will form a long ATP-dependent filament on the 5′ ssDNA tail carrying the DnaA-trio sequence (Richardson *et al.*, 2016). Filaments require the DnaA-boxes and will not form on the complementary DNA strand (the 3′ tail), which is consistent with the finding that DnaA is loaded from dsDNA onto ssDNA and forms a filament with 3′→5′ polarity (Cheng *et al.*, 2015). The filament is also specific for the DnaA-trio sequence as the level of filament formation is reduced when the complimentary sequence is used for the 5′-tail.

The results for the mutant variants show that as expected the I190A mutant protein will not form filaments on any of the tailed scaffolds (Figure 4.8.B). The protein will still form dimers that require the DnaA-boxes, suggesting that it can still specifically bind DnaA-boxes. The I193A mutant protein also appears to be able to specifically bind DnaA-boxes as dimer formation is stronger when the binding sites are present. While the I193A variant can still form a short oligomer on the substrate containing the DnaA-trios, it was clearly defective in filament formation compared to the wild-type protein. Neither I190A nor I193A variants were able to form filaments efficiently on either the 3′-tailed substrate or the 5′-tailed scaffold carrying the complementary sequence.

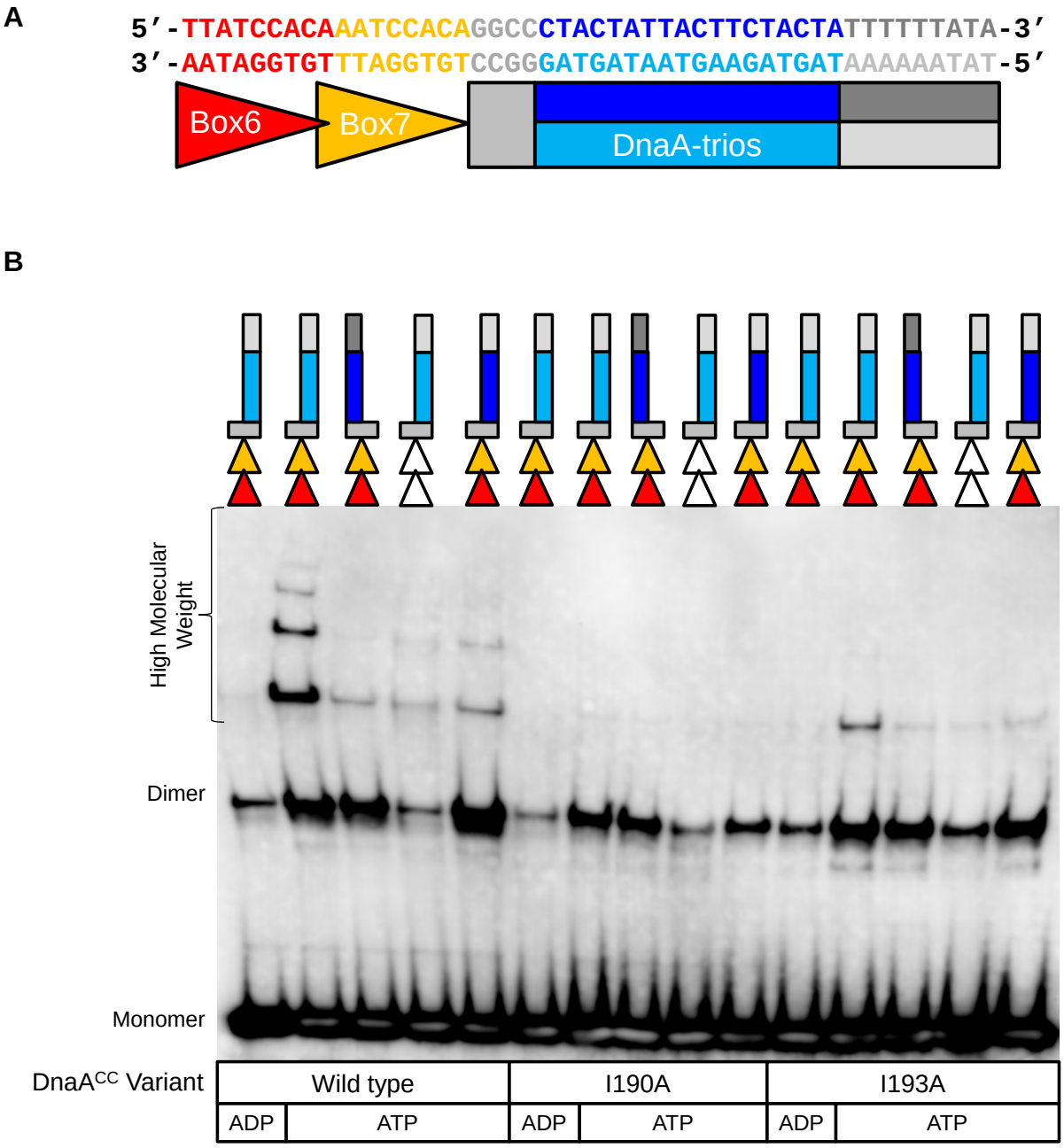


Figure 4.8. ATP-dependent oligomerisation of DnaA variants on DNA scaffolds *in vitro*. (A) Colour coded sequence and schematic of the region of *oirC* used for constructing the DNA scaffolds used in B. **(B)** DnaA oligomers captured using cysteine-specific crosslinking forming on and binding to DNA scaffold. The scaffold schematics are shown above colour coded to match A, white triangles indicate a scrambled sequence.

4.6. DnaA variants are defective in unwinding DNA

Section 4.5 established that both I190A and I193A DnaA variants are defective at forming filaments on single-stranded DNA. However, the I193A variant was more active at filament formation than the known single-stranded binding mutant I190A. Therefore, it was unclear whether this could account for the essential function required by the I193 residue. Therefore, I went on to assess whether this low level of filament formation was sufficient to promote origin unwinding.

To investigate unwinding activity, a DnaA-dependent strand separation assay was employed (described in Section 2.11.3 (Richardson *et al.*, 2019)). The assay utilises a DNA scaffold with a Cy5 fluorescent probe and a Black Hole Quencher (BHQ) (illustrated in Figure 4.9.A). Three oligonucleotides (listed in Table 2.6) were annealed together to form a double-stranded scaffold containing DnaA-box#6/7 and the DnaA-trios. The oligonucleotide complementary to the DnaA-trios was labelled on its 5′ terminus with a Cy5 fluorophore, while the oligonucleotide forming the upper strand of the DnaA-boxes was labelled on its 3′ terminus with a BHQ. The quencher absorbs the fluorescence wavelength emitted by Cy5.

The current model for DNA unwinding by DnaA is that DnaA-ATP forms a filament upon the DnaA-trios, which stretches one DNA strand and prevents complimentary base-pairing. On the DNA scaffold used for the strand separation assay this will result in the separation of the oligonucleotide complementary to the DnaA-trios (Figure 4.9.A). Once separated from the scaffold, fluorescence emitted by the Cy5-probe will no-longer be absorbed by the BHQ and so will become detectable. This allows for the measurement of how much DNA is being separated through an increasing fluorescence signal, an indication of whether, and how fast, the DnaA protein is unwinding DNA.

The purified DnaA^{CC} variants were used in the strand displacement assay to determine if they are capable of unwinding DNA. In the presence of ATP the wild-type DnaA^{CC} protein quickly separates the complementary strand as seen by the rapid increase in the fluorescence signal (plateauing after ~8 minutes), indicating the protein is unwinding the DNA (Figure 4.9.B). In the presence of ADP or if the DnaA-trio sequence is mutated (ΔTrios) then no fluorescence accumulates, suggesting the DNA scaffold is remaining intact and that the protein is not unwinding DNA under these conditions (as expected based on the model).

In the presence of ATP the level of fluorescence accumulating when the DNA scaffolds are incubated with the I190A variant are nearly identical to those seen in the presence of ADP or the ΔTrio sequence, suggesting no DNA strands are being separated and so the I190A protein is unable to unwind DNA (Figure 4.9.B.).

Interestingly, there is a slight accumulation of fluorescence in the presence of ATP when the scaffolds are incubated with the I193A variant compared to ADP and the ΔTrio sequence, but it is significantly slower, and the magnitude is less, compared to wild-type DnaA (Figure 4.9.B.). This also correlates with the filament formation assay where I193A was slightly more active than I190A but far less active than the wild-type protein. Taken together, the data suggests that I193 is an essential residue required for engaging DnaA-trios and unwinding the chromosome origin.

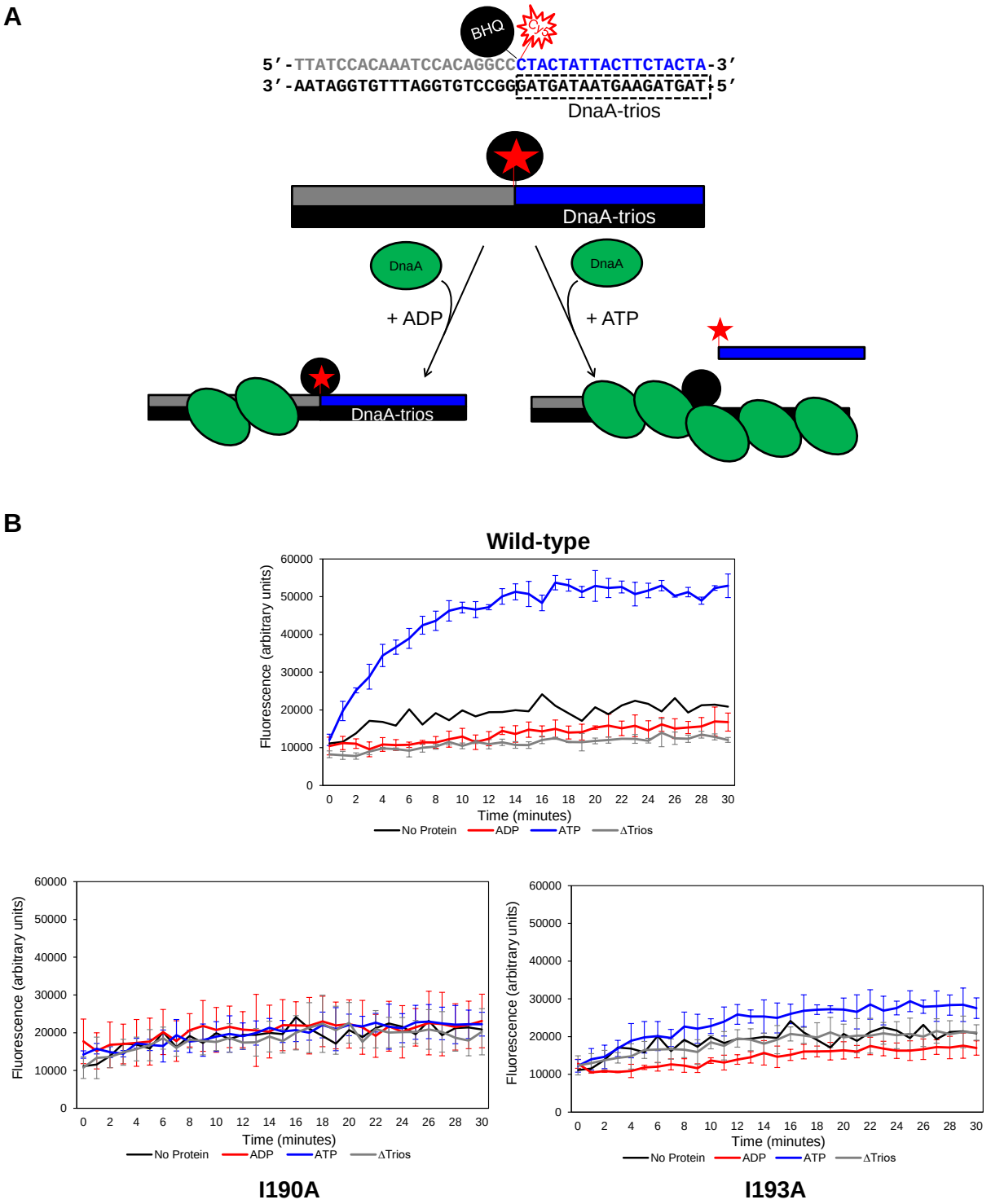


Figure 4.9. DNA unwinding by DnaA protein variants *in vitro*. (A) Outline of the DNA strand displacement assay. The DNA scaffold assembled is shown as a colour coded sequence and schematic. The black hole quencher (BHQ) is shown in black and the cy5 probe in red. **(B)** Fluorescence measured during the DNA strand seperation assay for the DnaA mutant variants incubated with a native substrate in the presence of ADP (blue) or ATP (red), or when incubated with a ΔDnaA-trios substrate in the presence of ATP (grey). Error bars display the standard deviation from the mean for three biological replicates. The fluorescence measurment for the DNA scaffold incubated without protein (black) is shown as a control.

Chapter 4 – Discussion

DnaA-trio specificity and the essential isoleucine residues of the ISM

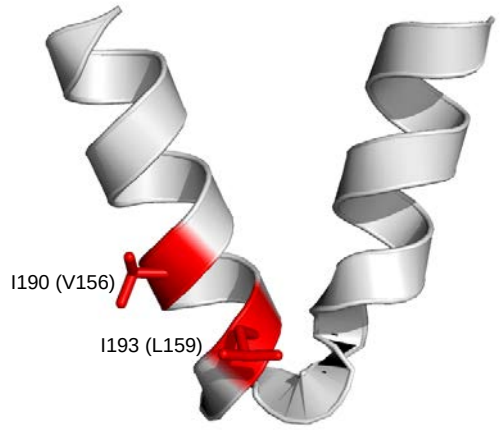
The current model for origin opening is that DnaA-ATP forms a filament originating on the DnaA-boxes from where it is loaded onto a single strand of the unwinding site to promote melting of the duplex. The DnaA-trios have been shown to provide the specific sequence for precisely guiding DnaA filament formation onto a single strand of the origin. However, how DnaA specifically recognises and binds DnaA-trios is not understood.

Using the inducible *oriN* system and recombinant DnaA protein variants two isoleucine residues, I190 and I193 (Figure 4.10.A), have been determined to be essential *in vivo* (Figure 4.2) and required for forming filaments on ssDNA and unwinding the DNA duplex *in vitro* (Figures 4.8 and 4.9). The I190 residue is a conserved hydrophobic residue that was previously shown *in vitro* to be required for binding single-stranded DNA and the DnaA-trio sequence (Duderstadt *et al.*, 2011; Richardson *et al.*, 2019). To explore the conservation of the I193 residue, a comparative alignment was performed for DnaA from a diverse range of bacterial species. The table in Figure 4.10.B shows the result of this alignment for the residues homologous to positions 190 and 193 in *B. subtilis*. It revealed that the I193 residue is also a highly conserved hydrophobic residue. However, the importance of I193 in specific ssDNA binding was not previously appreciated.

The variant proteins with the isoleucines individually substituted are capable both of forming ATP dependent filaments (Figure 4.7) and specifically binding DnaA-boxes (Figure 4.8), suggesting the residues were not required for oligomerisation or proper protein folding. However, both residues appear to be required for forming filaments on single-stranded DnaA-trios and unwinding duplex DNA (Figures 4.8 and 4.9).

Interestingly, the I193A variant is able to form oligomers more readily than the I190A substitution (Figure 4.8). Consistently, the I193A protein showed a low level of DNA unwinding activity in the presence of ATP (Figure 4.9). This result suggests that perhaps the I193A variant could still form a short or unstable ssDNA bound filament specifically on the DnaA-trios capable of unwinding duplex DNA, but that its activity is reduced to a level incompatible with viability.

A



B

			B. subtilis DnaA Residue	
	Phylum	Speices	190	193
Gram Positive	Firmicutes	***B. subtilis***	I	I
		O. picturae	I	I
		L. monocytogenes	I	I
		S. aureus	I	I
		S. pneumoniae	L	L
		L. garvieae	V	T
		C. diffacile	I	I
	Actinobacteria	*C. epidermidicanis*	I	V
		M.tuberculosis	I	L
	Cyanobacteria	*N. punctiforme*	I	I
		S. synechocystis	I	I
		C. thermalis	I	I
Gram Negative	Thermotogae	*T. maritima*	V	M
	Aquificae	*A. aeolicus*	V	L
	Bacteroidetes	*F. johnsoniae*	I	V
	Spirochaetes	*T. denticola*	V	V
		B. hermsii	V	I
		L. interrogans	R	L
	Chlamydiae	*C. trachomatis*	V	L
		C. pneumoniae	V	I
	Proteobacteria	*C. jejuni*	T	L
		H. pylori	L	L
		R. slovaca	V	L
		C. crescentus	V	V
		R. solanacearum	V	Y
		N. meningitidis	M	V
		X. fastidosa	I	L
		P. putida	V	L
		H. influenzae	V	V
		V. chloerae	V	L
		E. coli	V	L

Figure 4.10. Essential isoleucine residues of the DnaA initiator specific motif. (A) Structure showing the location of the essential *B. subtilis* DnaA isoleucine's (red) on the structure of the *A. aeolicus* ISM (PDB ID 1L8Q). Residues are labelled according to the *B. subtilis* numbering, with *A. aeolicus* in parentheses. **(B)** Table highlighting the conservation of hydrophobic side chains for the DnaA amino acids homologous to positions 190 and 193 in *B. subtilis*. Species were selected on protein sequence availability to represent most major phyla and the genus within them and ordered according to phylogenetic relatedness to *B. subtilis*. Hydrophobic residues are highlighted orange with none-hydrophobic in blue. (All DnaA sequences were downloaded from NCBI PubMed, initially aligned using clone manger version 9 and phylogenetic relatedness was determined using (Bern and Goldberg, 2005; Horiike *et al.*, 2009)).

While both I190A and I193A variant proteins were defective in filament formation on scaffolds containing 5′ ssDNA tails, these DNA substrates also contained DnaA-boxes (Figure 4.8). This adds the possibility that the mechanism of loading the DnaA from the DnaA-boxes to the ssDNA tails is being affected in the I190A and I193A proteins. To determine if the two isoleucines are required for general ssDNA binding, assays which investigate the ability of the protein variants to bind non-specific ssDNA substrates need to be utilised, for example performing the crosslinking assay in the presence of single-stranded oligonucleotides or investigating protein-nucleotide interactions via SPR (Surface Plasmon Resonance).

Investigations analysing the combined data from dozens of protein-DNA complexes have identified that, where hydrophobic residues such as leucine and isoleucines are utilised to interact with DNA, they are more likely to be making contacts with the sugar or phosphate group rather than the base (Sathyapriya *et al.*, 2008), such as the single-stranded DNA binding protein of *Bacillus anthracis* (Biswas-Fiss *et al*., 2012). This natural propensity for using hydrophobic residues for non-specific DNA binding supports the hypothesis that the two essential isoleucines are involved in a single-stranded DNA binding mechanism.

If the two isoleucine residues are mainly involved in non-specific ssDNA binding, this then still leaves the mechanism for DnaA-trio recognition and binding unidentified. As discussed previously, DnaA has been shown to bind to a synthetic ssDNA substrate using the ISM and a neighbouring pair of helices (α5 and α6). It is possible that the specific residues required for DnaA-trio recognition are located here rather than the ISM. It is also not impossible that the mechanisms for DnaA-trio recognition and engagement are performed by unique motifs/residues. In this scenario an unidentified region of DnaA could be responsible for the specific recognition of the DnaA-trio motif before or after the protein binds the sequence using the general ssDNA binding mechanism. Determining if any of these hypotheses is correct could be accomplished by utilising the techniques employed here to investigate additional DnaA residues. These residues must be located within the AAA+ or DNA binding domains of the protein as previous investigations have found that DnaA-trios are specifically recognised by DnaA mutants lacking domains I-II (Richardson *et al.*, 2016).

Finally, there is a possibility that the specificity for DnaA-trios does not lie in specific amino-acid residues within DnaA, but rather the specific structure of the DNA itself.

DNA has been shown to vary structurally in a sequence dependent manner and this variation can be used by proteins to recognise specific sequences (Rohs *et al.*, 2010). For example, A-tracts in AT-rich sequences are associated with a narrow minor groove. This conformation enhances the local negative electrostatic potential of the DNA which can provide a specific binding site for a positive amino acid side chain (Rohs *et al.*, 2009). Therefore, it is possible that the DnaA-trio sequence forms a specific 3-dimensional shape that promotes binding by DnaA.

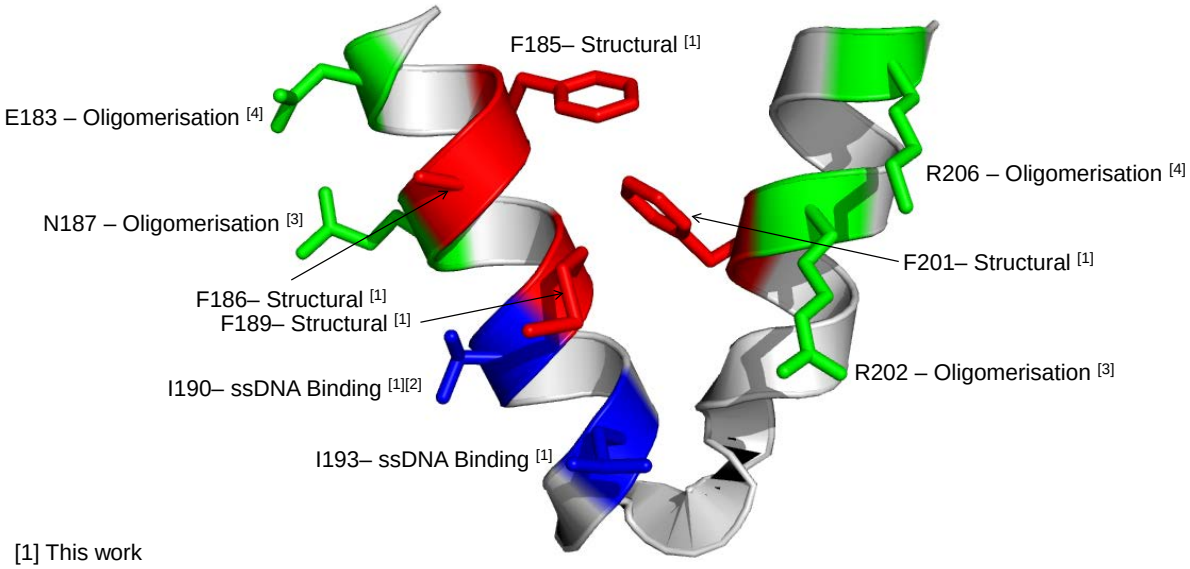
While the two isoleucine residues identified may not be involved in the specific recognition or binding of the DnaA-trio sequence it is possible they are required for unwinding the DNA duplex. It has been established that the DNA double helix is stabilised not only by hydrogen bonding between the complementary bases, but also through stacking of adjacent base pairs (Herskovits, 1962). Hydrophobic cohesion, resulting from this base-pair stacking, is a major contributor to stability and requires abundant water (Feng *et al.*, 2019). It has recently been reported that semihydrophobic molecules can interfere with this cohesion, resulting in unstacking of the bases and leading to transient holes occurring between the base-pairs, destabilising the helix. However no molecular mechanism has been proposed for how this occurs (Feng *et al.*, 2019). It is possible therefore that the two conserved hydrophobic residues within the ISM could be asserting similar effects upon the base-pair stacks. This could result in the destabilisation of the double-stranded DNA, contributing to duplex unwinding.

The structure and function of the *Bacillus subtilis* DnaA initiator specific motif

The identification of the two essential isoleucine residues was achieved through the wider investigation into the amino acid residues which form the DnaA initiator specific motif. For the first time all 26 residues forming this motif in *Bacillus subtilis* were investigated for their requirement for a functional DnaA protein *in vivo*.

It was determined that 10 of the ISM amino acids were required for either viability or optimal growth. Through further investigation, both *in vitro* and by searching the published literature, a functional role has been proposed for each of these residues, highlighted in Figure 4.11. Several of the residues were previously shown to be involved in filament formation (green Figure 4.11) and two further residues now appear to be involved in ssDNA binding (blue Figure 4.11). I speculate that essential residues facing the buried region of the ISM are likely playing a structural role (red Figure 4.11).

The ISM is a unique insertion into the AAA+ domain of proteins associated with the initiation of DNA replication. The investigation performed here provides new insights into both the structure and function of the ISM, providing a clearer understanding of the importance of the amino acids that compose it.



[1] This work
[2] Ozaki *et al*., 2008
[3] Duderstadt *et al*., 2010
[4] Ozaki *et al*., 2012

Figure 4.11. Proposed roles for the amino acid residues of the DnaA initiator specific motif. The residues of the ISM of DnaA that are required for protein function along with the activities they are proposed to function in. Residues are highlighted on the structure from *A. aeolicus* (PDB ID 1L8Q) and labelled according to the *B. subtilis* numbering. The investigations the proposed roles are based on are indicated. Residues proposed to be playing a structural role are highlighted in red, those proposed as functioning in oligomerisation are highlighted green and those proposed as functioning in single-stranded DNA (ssDNA) binding are highlighted blue.

Chapter 5

Molecular mechanisms and regulation of replicative helicase loading in *Bacillus subtilis*

Chapter 5 – Introduction

The process of initiating DNA replication in bacteria can be broken down into several key steps; recruitment of DnaA to the origin, unwinding of the origin, recruitment of the replicative helicases and the coordinated loading of them onto the open DNA strands. The mechanisms underlying DnaA recruitment to *oriC* are well understood and the process by which the protein opens the origin is starting to become clearer. The next unanswered questions of DNA replication initiation are the downstream molecular mechanisms behind the recruitment and loading of the replicative helicase.

As was discussed in Chapter 1 (Section 1.1.1), in bacteria the replicative helicase is loaded around a single strand of the open origin complex. In *E. coli* DnaA directly interacts with the helicase, loading it with the assistance of a helicase loader protein which complexes with the helicase. DnaA domain I ($DnaA^{DI}$) of *E. coli* has an interaction surface which has been proposed to directly mediate helicase binding (most notably via residues Glu21 (Abe *et al.*, 2007) and Phe46 (Kaguni, 2011)) (Section 1.3.5).

In *B. subtilis*, replicative helicase loading involves a pathway encompassing a primosomal complex of several essential origin binding proteins (section 1.5). These proteins are DnaA, the *Firmicute* specific accessory proteins DnaB and DnaD, and the helicase loader DnaI (homologous to the *E. coli* loader DnaC) (Briggs *et al.*, 2012).

Several studies have investigated the essential interactions between these proteins, creating the foundations for a molecular picture of the initiation complex previously discussed in Section 1.5.3. One such study utilised assays to detect the association of the initiator proteins and helicase with *oriC* during initiation *in vivo* (Smits *et al.*, 2010). These assays determined which proteins are required for the temporal association of the others with the origin and these interactions have been confirmed through two-hybrid studies (Ishigo-Oka *et al.*, 2001; Matthews and Simmons, 2019). This has led to the establishment of a model for the hierarchical ordered recruitment of replication initiator proteins re-outlined in Figure 5.1.A.

The exact molecular mechanisms underpinning these interactions still remain to be identified. Investigations, however, have started to determine the residues and interactions required for the key first step in helicase loading, the DnaA-DnaD interaction. Two recent independent studies utilised differing approaches to produce models for the interaction. The first model is from Matthews and Simmons, 2019 who used an *E. coli* two-hybrid approach to investigate protein-protein interactions, and is outlined in Figure 5.1.B (model one). This model proposes that the N-terminal domains of both proteins form the interaction interface. The second model is proposed by Martin *et al.,* 2019 who utilised Nuclear Magnetic Resonance (NMR). This model also proposes that the NTD of both proteins interact although the residues identified vary slightly. The study also proposes there is a second, potentially weaker, interaction occurring between $DnaA^{DI}$ and the C-terminal domain of DnaD (Figure 5.1.B, model two).

Critically however, the *in vivo* relevance of the proposed interactions was not clear. Moreover, both studies these models are based on did not utilise full length proteins (i.e. – they used isolated domains of each protein). This chapter will investigate the mechanisms and regulation of helicase loading in *B. subtilis*, starting with detailed analysis of the DnaA-DnaD interaction before producing a clearer picture of the interactions of the full-length primosomal complex proteins of *B. subtilis*.

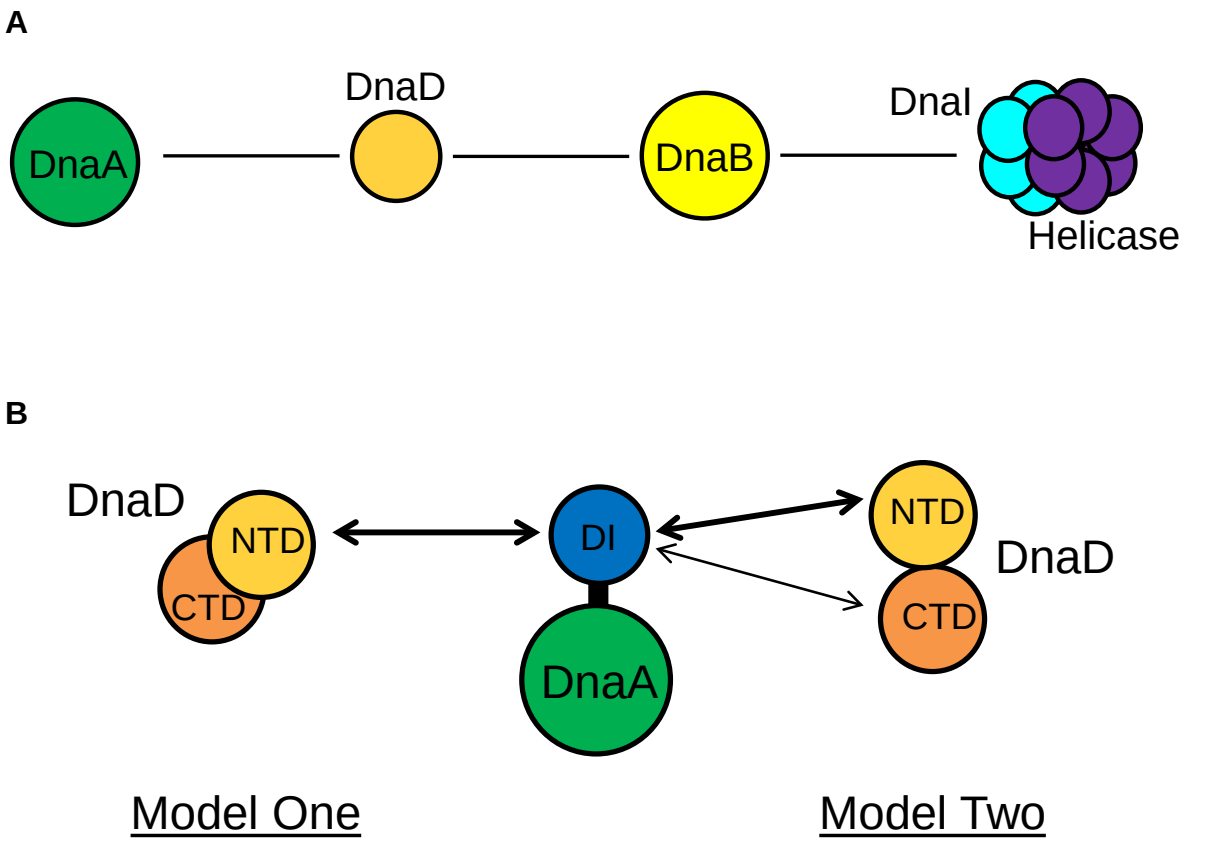


Figure 5.1. Helicase loading in *Bacillus subtilis*. (A) The proposed helicase loading pathway from *B. subtilis* showing the hierarchical order of recruitment to *oriC*. **(B)** Proposed alternative models for the interaction between DnaA and DnaD. Boldness of arrows indicates the relative strength of the interaction.

Chapter 5 – Results

5.1. The DnaD interaction interface of DnaA domain I is physiologically relevant

As discussed in the introduction to this chapter, two independent studies have proposed models explaining the interaction between DnaA and DnaD. Both investigations identified the same surface of $DnaA^{DI}$. However the physiological relevance of this proposed interface had not been determined.

The inducible *oriN* strain from Chapter 3 was employed to examine whether any of the proposed $DnaA^{DI}$ residues are essential *in vivo*. The first residues investigated were the small number proposed by both Matthews and Simmons, 2019 and Martin *et al.,* 2019 (T26, W27, F49) which have been mapped in Figure 5.2.A. As previously, the residues of the endogenous gene were substituted individually for alanine. The results indicated that all three residues are required for a functional DnaA protein (Figure 5.3.B). An immunoblot confirmed that all the mutant proteins are stably expressed (Figure 5.3.C).

While the investigation of Matthews and Simmons, 2019 focused solely on the three residues investigated above, the NMR study performed by Martin *et al.,* 2019 proposed several additional residues. A subset of these, mapped in figure 5.3.A, formed a clear cluster around the F49 residue which was shown above to be both essential and required for the interaction with DnaD. Therefore, these residues (E48, D52, W53) were investigated using the same set of assays to determine their physiological relevance.

The cluster of residues was substituted for alanine within the *oriN* strain. The results of the substitutions indicated that the residues E48 and W53 were both essential as their substitution resulted in a dependency on IPTG, comparable to that of the F49 substitution (Figure 5.3.B). The D52 residue is non-essential. An immunoblot confirmed the mutant proteins were being properly expressed (Figure 5.3.C).

As mentioned these residues were identified via assays utilising truncated proteins. To establish if these residues are required for the interaction of full length DnaA with DnaD an assay allowing the investigation of full length protein-protein interactions was required.

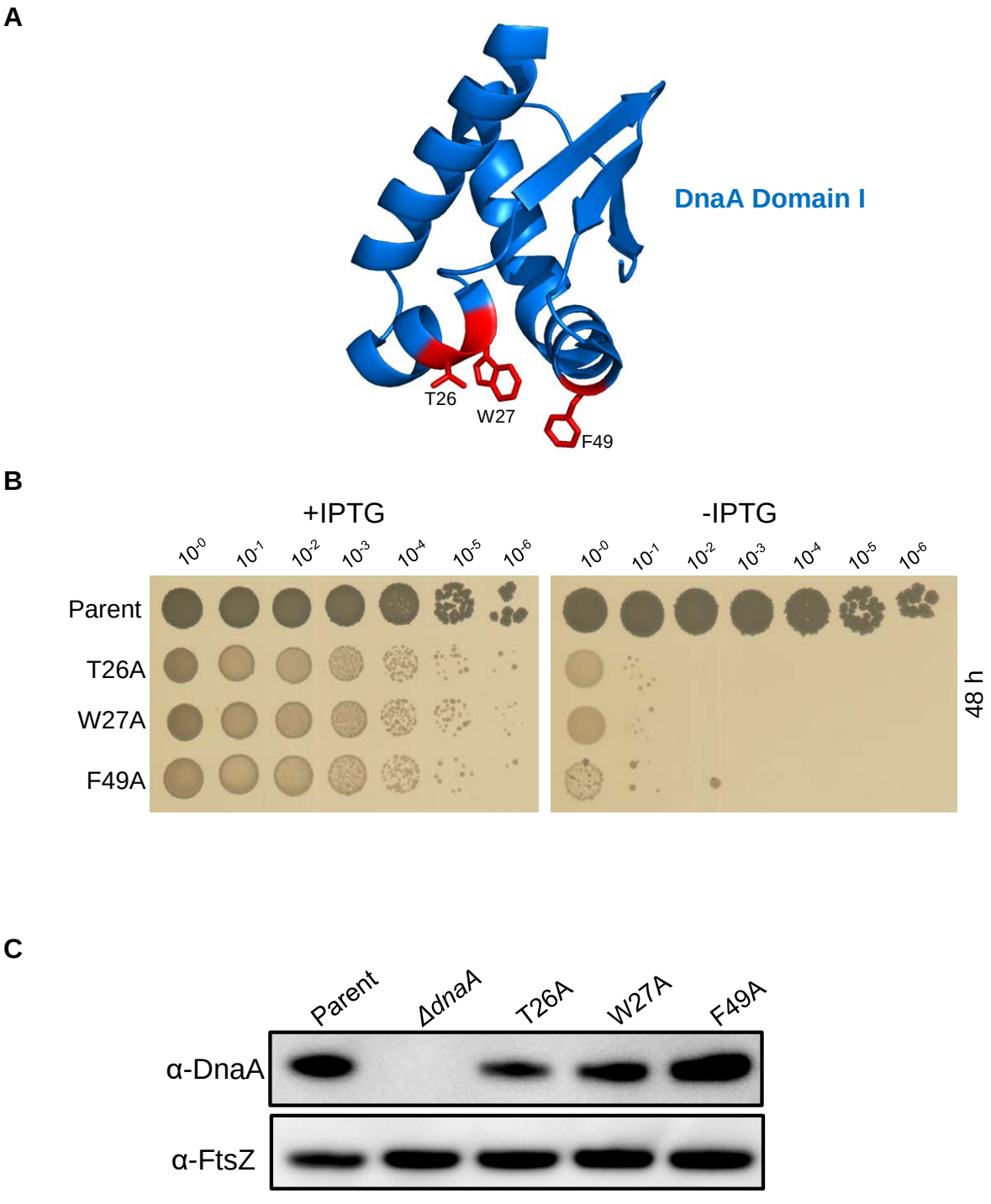


Figure 5.2. *In vivo* analysis of the proposed DnaD interaction interface of DnaA. **(A)** The residues of DnaA domain I proposed as interacting with DnaD mapped onto a *B. subtilis* DnaA domain I crystal structure (PDB ID 4TPS). **(B)** Analysis of *B. subtilis* DnaA substitution mutants following 48 hours incubation at 37°C. **(C)** Immunoblot analysis of the DnaA substitution mutants with an FtsZ loading control.

Parent (HM1108), T26A (HM1540), W27A (HM1541), F49A (DS31), *ΔdnaA* (HM1424).

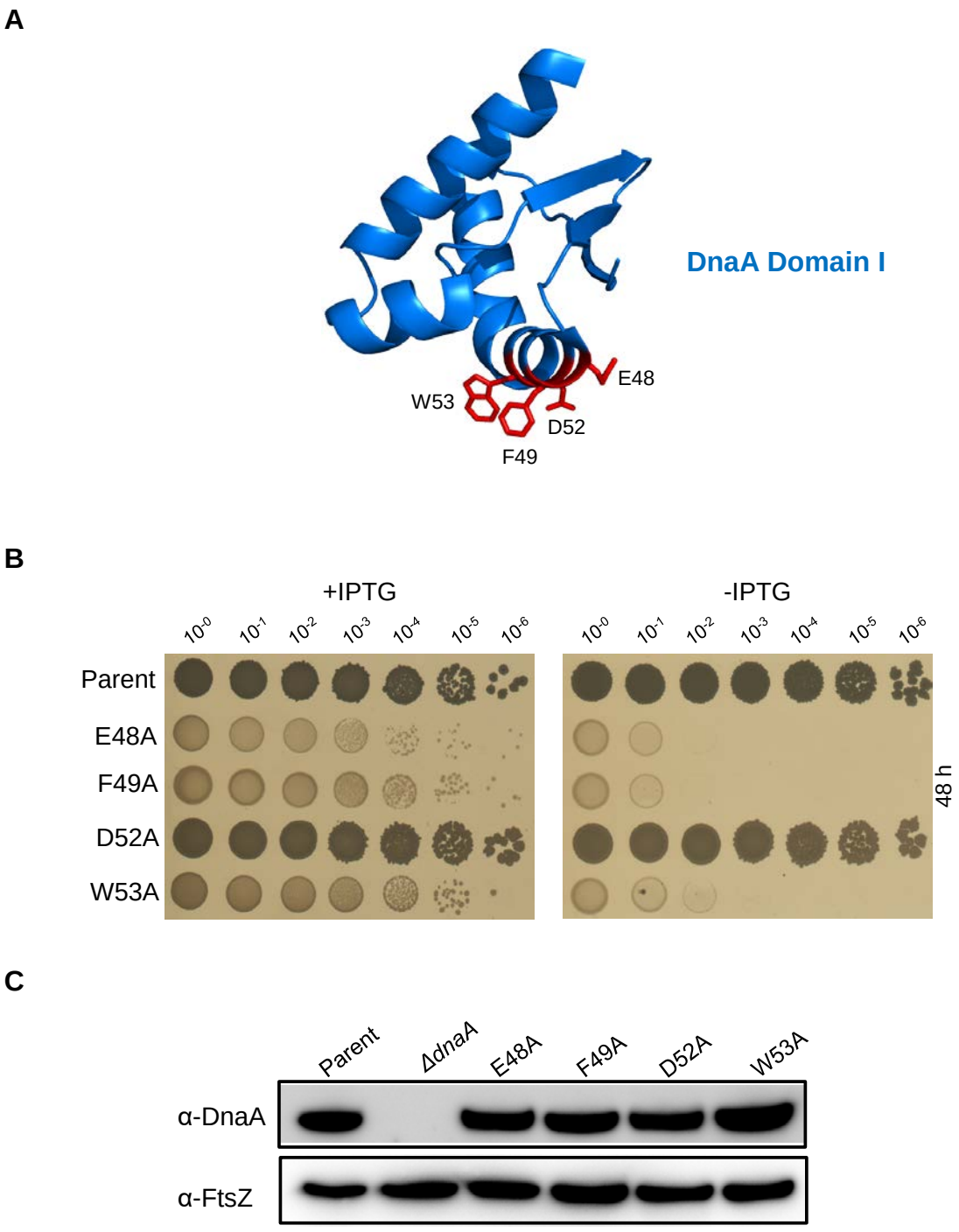


Figure 5.3. *In vivo* analysis of DnaA domain I residues. (A) The residues of DnaA domain I investigated mapped onto a *B. subtilis* DnaA domain I crystal structure (PDB ID 4TPS). **(B)** Analysis of *B. subtilis* DnaA substitution mutants following 48 hours incubation at 37°C. **(C)** Immunoblot analysis of the DnaA substitution mutants with an FtsZ loading control.

Parent (HM1108), E48A (DS35), F49A (DS31), D52A (DS32), W53A (DS33) *ΔdnaA* (HM1424).

5.2. *Bacillus subtilis* initiator protein-protein interactions can be investigated via bacterial two-hybrid

To investigate full length protein-protein interactions an adapted bacterial two-hybrid (BTH) assay was developed that allowed for the investigation of the interactions between full length *B. subtilis* initiator proteins. This BTH is illustrated in Figure 5.4.

The assay is based on the well-established method of fusing the genes of the proteins of interest to the complimentary fragments (T18 and T25) of the *Bordetella pertussis* adenylate cyclase (Karimova *et al.*, 1998). The hybrid fusions are heterologously expressed in *E. coli* where an interaction between the test proteins results in the functional complementation of the cyclase fragments resulting in enzymatic synthesis of cyclic AMP (cAMP) (Figure 5.4). Following synthesis cAMP binds CAP (catabolite activator protein) and the CAP:cAMP complex binds to the P_{lac} promoter, where it activates transcription resulting in the expression of *lacZ*. This expression can be detected through a β-galactosidase activity reporter, in this case the colonies turning blue in the presence of X-Gal (Battesti and Bouveret, 2012).

It has previously been shown that expressing *Bacillus* DnaA in *E. coli* is toxic as it negatively impacts DNA replication, likely by interfering with native *E. coli* DnaA activity at *oriC* (Krause and Messer, 1999), rendering investigation of full length DnaA via BTH impossible (Matthews and Simmons, 2019). To bypass this, a derivative of the two-hybrid reporter strain was utilised which carries a deletion of *rnhA* (Section 1.1.1). This gene encodes for RNase HI which resolves R-loops. R-loops are a structure formed of three nucleic acid strands from the hybridization of an RNA polymer to a complimentary DNA strand during transcription. This displaces the second DNA strand into a displaced loop. The deletion of *rnhA* results in stable R-loop formation, essentially producing RNA primers. These structures can be recognised by helicase/polymerase to initiate new rounds of DNA synthesis. This mode of replication initiation is independent of both *oriC* and DnaA (Kogoma and von Meyenburg, 1983; Lombrana *et al.*, 2015).

The bacterial two-hybrid assay was performed heterologously in an *E. coli* strain undergoing *oriC*-independent DNA replication. It is possible, though unlikely, that these factors could influence the results of the assay, as in a non-native environment under replicative stress the *B. subtilis* proteins could be more likely to associate or localise with one another. Performing the assay in *B. subtilis* could alleviate these concerns but

is unlikely to be possible with the initiator proteins. To perform the assay in *Bacillus* would require deleting the native initiator proteins, so would require a mechanism for *oriC* independent growth. One such mechanism has been discussed in this thesis already, the *oriN* system. This system however could not be used in this context as, although RepN can substitute for DnaA, the other initiator proteins, DnaD and DnaB, are both required for the function of *oriN* (Hassan *et al.*, 1997). As such the bacterial two-hybrid was executed in *E. coli* as the closest analogue to the *B. subtilis in vivo* conditions. Performing the assay in *E. coli* is unlikely to invalidate any results but identified interactions may require further validation through other assays before firm conclusions can be drawn.

The BTH assay was performed as per section 2.9. Plasmids expressing the protein-cyclase fusions were co-transformed into the *ΔrnhA E. coli* reporter strain, incubated overnight before dilution and spotting onto media containing X-gal to detect β-galactosidase activity and appropriate antibiotics to select for the transformed plasmids.

Test proteins can be tagged with the cyclase fragments either N- or C-terminally. The interaction between N-terminally tagged test proteins proved to show the clearest and most consistent interactions and so these were the fusions used for the two-hybrid investigations. Tagging the test proteins at their N-terminus is the same approach taken by previous two-hybrid investigations into the interactions of the *B. subtilis* initiator proteins (Matthews and Simmons, 2019).

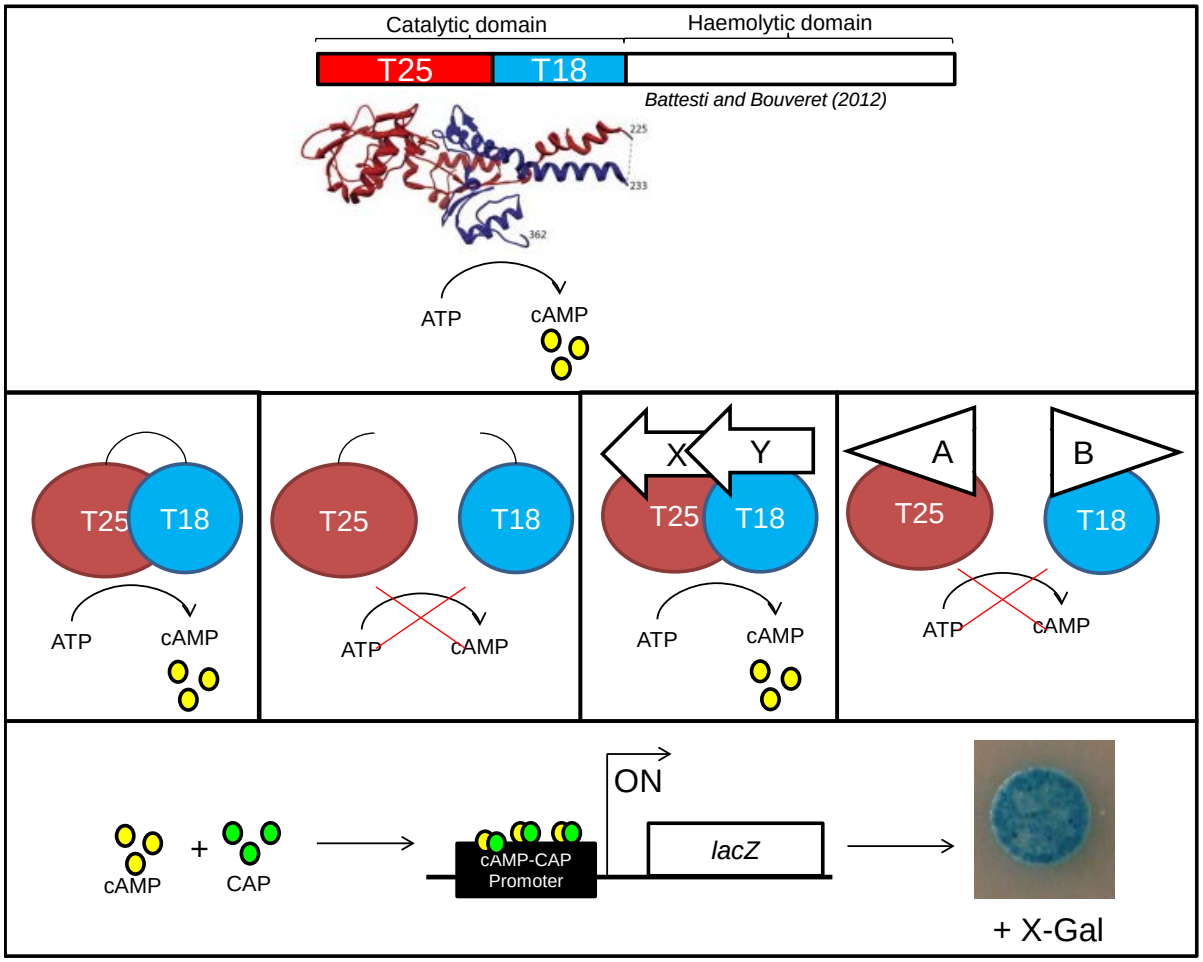


Figure 5.4. Principle of the adapted bacterial two-hybrid system. The catalytic domain of *Bordetella pertussis* adenylate cyclase synthesises cAMP. The two fragments of the domain expressed as a fusion can synthesises cAMP but if expressed separately cannot. Genetic fusion of the fragments to proteins which interact rescues cAMP synthesis, but if the proteins do not interact synthesis is not rescued. cAMP synthesis (and therefore a protein:protein interaction) can be detected through the activation of the *lac* promotor which results in colonies turning blue in the presence of X-Gal. Crystal structure taken from Battesti and Bouveret, 2012.

5.3. The DnaD interaction interface of DnaA domain I is required for the interaction between full length proteins

Section 5.1 established that a number of residues proposed as forming the DnaD interaction interface of $DnaA^{DI}$ by two independent studies are physiologically relevant. These residues were identified via assays utilising truncated proteins, so to establish if these residues are required for the interaction of full length DnaA with DnaD the adapted bacterial two-hybrid (Section 5.2) was used. The lethal substitutions were introduced into the adenylate-cyclase tagged DnaA and investigated for their ability to interact with full length DnaD.

The residues proposed by both earlier studies (T26, W27, F49) were, again, the first to be investigated. The results showed that wild-type DnaA interacted with itself and DnaD, validating that the adapted BTH could be used for investigating the protein-protein interactions of full length DnaA (Figure 5.5.A). The mutant variants of DnaA all maintained the interaction with wild-type DnaA, suggesting that the proteins are functionally expressed in the assay. All three mutant DnaA proteins lost the ability to interact with DnaD (Figure 5.5.A).

The adapted bacterial two-hybrid assay was utilised to determine if the additional essential $DnaA^{DI}$ residues initially proposed by Martin *et al.,* 2019 were required for the interaction with DnaD. The results show the $DnaA^{W53A}$ protein variant interacts with wild-type DnaA, indicating the mutant protein is expressed in the two-hybrid (Figure 5.5.B). The same variant lost the ability to interact with DnaD, suggesting this residue is also required for the interaction with DnaD. The E48A variant did not interact with anything, a result found regardless of which terminal the mutant protein was tagged (Figure 5.5.B and data not shown). As such, the role of this residue for binding to DnaD remains unclear.

The results indicate that a physiologically relevant interaction surface of $DnaA^{DI}$ composed of the residues T26, W27, F49 and W53 is required for the interaction of full length DnaA with DnaD (Figure 5.5.C).

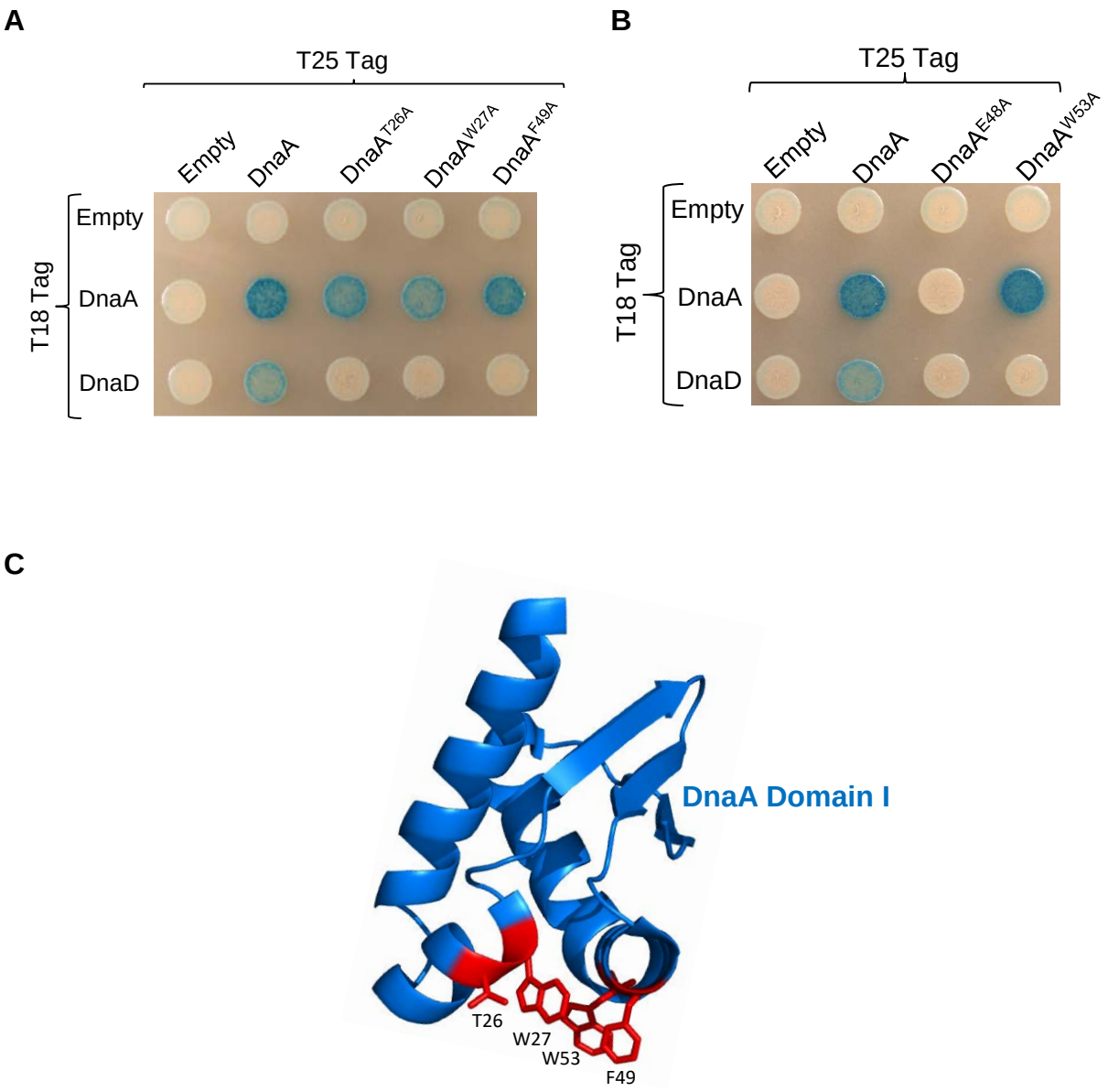


Figure 5.5. Interaction of DnaA with DnaD. Bacterial two-hybrid assays showing the interaction between wild-type DnaA and the DnaA mutant variants **(A)** T26A W27A, F49A and **(B)** E48A, W53A with DnaA and DnaD. All proteins are N-terminally tagged with the adenylate cyclase fragment. **(C)** The residues of DnaA domain I that are physiologically relevant and required for the interaction with DnaD mapped onto a *B. subtilis* DnaA domain I crystal structure (PDB ID 4TPS).

T18 Empty (pUT18C), T18 DnaA (pHM638), T18 DnaD (pHM642), T25 Empty (pST25), T25 DnaA (pHM640), T25 DnaA^{T26A} (pDS119), T25 DnaA^{W27A} (pDS120), T25 DnaA^{F49A} (pDS84), T25 DnaA^{E48A} (pDS87), T25 DnaA^{W53A} (pDS137).

5.4. The N-terminal domain of DnaD contains a physiologically relevant surface required for the interaction with DnaA

Both of the recent studies used to investigate the interaction between DnaA and DnaD proposed residues on the same surface of $DnaA^{DI}$ as interacting with DnaD (Section 5.1). These same studies also proposed residues from DnaD that are involved in the interaction with DnaA. However, bacterial two-hybrid analysis by Matthews and Simmons, 2019, indicated that the N-terminal domain (NTD) of DnaD is the site of the only interaction with DnaA, whereas NMR analysis by Martin *et al*., 2019, suggested that the C-terminal domain (CTD) also interacts with DnaA (Figure 5.1.B). Both studies performed the investigation using only truncated proteins of individual domains.

To investigate how full length DnaA and DnaD interact, the bacterial two-hybrid was employed to determine the interactions between either full length proteins or the individual protein domains. It has already been established that full length DnaD interacts with DnaA (Section 5.3) and previous studies have indicated DnaD interacts with itself and DnaB (Matthews and Simmons, 2019). $DnaA^{DII-IV}$ is a DnaA variant lacking DnaA domain I which presumably will not interact with DnaD based on the results of the previous sections.

The results of the BTH show that full length DnaD interacts with itself, DnaA and DnaB, and that the interaction with DnaA requires $DnaA^{DI}$ (Figure 5.6.A). Identical results were found for the $DnaD^{NTD}$, except interestingly the interaction with DnaA seemed to be stronger than that of full length DnaD with DnaA. The $DnaD^{CTD}$ failed to interact with anything except full length DnaD, although this interaction was heterogeneous and does not conclusively show if the C-terminal domain was being functionally expressed (a result consistent with Matthews and Simmons, 2019). The result provides evidence that there is an interaction surface with DnaA within the $DnaD^{NTD}$ but does not rule out an interaction with the CTD.

To further investigate the interaction between DnaA and DnaD a series of residues within the $DnaD^{NTD}$, highlighted in Figure 5.6.B and implicated in the interaction previously (Matthews and Simmons, 2019), were substituted for alanine. These substitutions were, unlike the previous study, introduced into the full length protein. The result showed that these substitutions completely knocked out the interaction with DnaA, while maintaining the interaction with wild-type DnaD (indicating the mutant proteins are functionally expressed) (Figure 5.6.C). While the $DnaD^{F51}$ and $DnaD^{I83}$

residues had previously been shown to be required for the interaction with DnaA, the DnaD^{E95} residue had only been suggested as being required but had not, until now, been demonstrated as being required.

While these results do not completely dismiss an interaction between DnaA and the C-terminal domain of DnaD, taken together they do suggest that the strongest interaction is between DnaA and DnaD is mediated through an interface on the N-terminal domain.

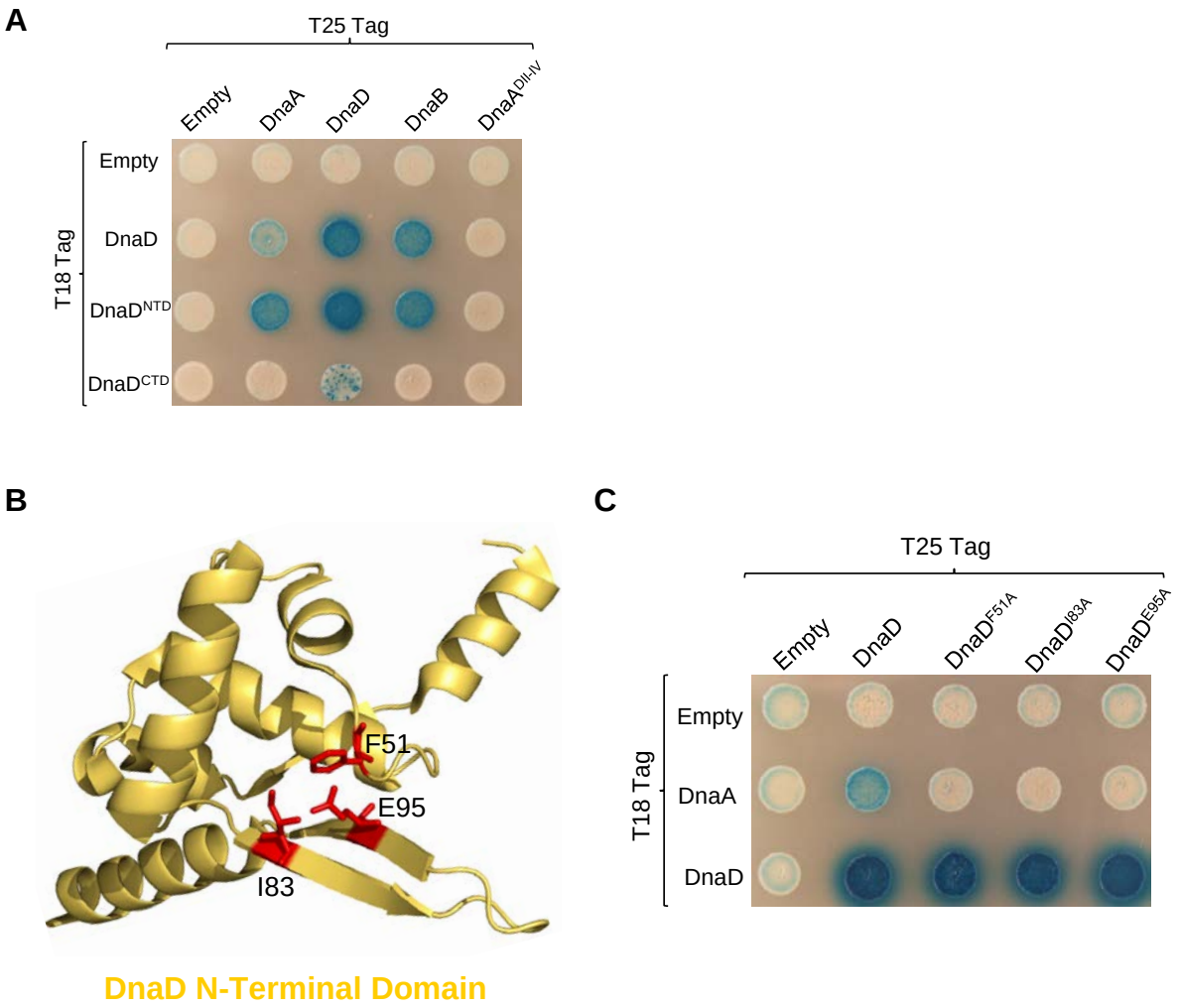


Figure 5.6. The interactions of the domains of DnaD with DnaA. (A) Bacterial two-hybrid assay showing the interaction between full length DnaD and the individual protein domains with DnaA, DnaD, DnaB and DnaA lacking domain I (DnaA^{DII-IV}). **(B)** The residues of the DnaD^{NTD} implicated in the interaction with DnaA mapped onto a crystal structure of the domain (PDB ID). **(C)** Bacterial two-hybrid assay showing the interaction between wild-type DnaD and DnaD mutant variants with DnaA and DnaD.

T18 Empty (pUT18C), T18 DnaD (pHM642), T18 DnaD^{NTD} (pDS121), T18 DnaD^{CTD} (pDS125), T18 DnaA (pHM638), T25 Empty (pST25), T25 DnaA (pHM640), T25 DnaD (pHM644), T25 DnaB (pHM652), T25 DnaA^{DII-IV} (pHM648), T25 DnaD^{F51A} (pDS126), T25 DnaD^{I83A} (pDS127), T25 DnaD^{E95A} (pDS132).

To investigate the physiological relevance of the interaction surface of DnaD a tool that allows the investigation of potentially lethal *dnaD* mutants *in vivo* was utilised (Figure 5.7.A). This tool is a strain of *B. subtilis* with an ectopic copy of *dnaD* introduced under the control of an IPTG-inducible promoter. Expression of the inducible copy of *dnaD* is sufficient to sustain growth if the native *dnaD* is deleted (Figure 5.7.B). Basal expression of the ectopic *dnaD* was reduced by fusing an *ssrA* degradation tag which targets the protein for degradation by the protease ClpXP (Wiegert and Schumann, 2001). This modification produced a complementation system dependent upon induction of *dnaD-ssrA*.

The $DnaD^{NTD}$ residues forming the interaction interface with DnaA were substituted for alanine in the endogenous *dnaD* within the inducible *dnaD-ssrA* strain (data collected by Charles Winterhalter, unpublished). The results of the substitutions revealed that all three DnaD residues are required for a functional protein *in vivo* (Figure 5.7.B). An immunoblot confirmed that the mutant proteins were being expressed (Figure 5.7.C). These results suggest that the $DnaD^{NTD}$-DnaA interaction surface is physiologically relevant.

A potential model for the interaction interface between DnaA and DnaD, incorporating all the results obtained so far is shown in Figure 5.7.D.

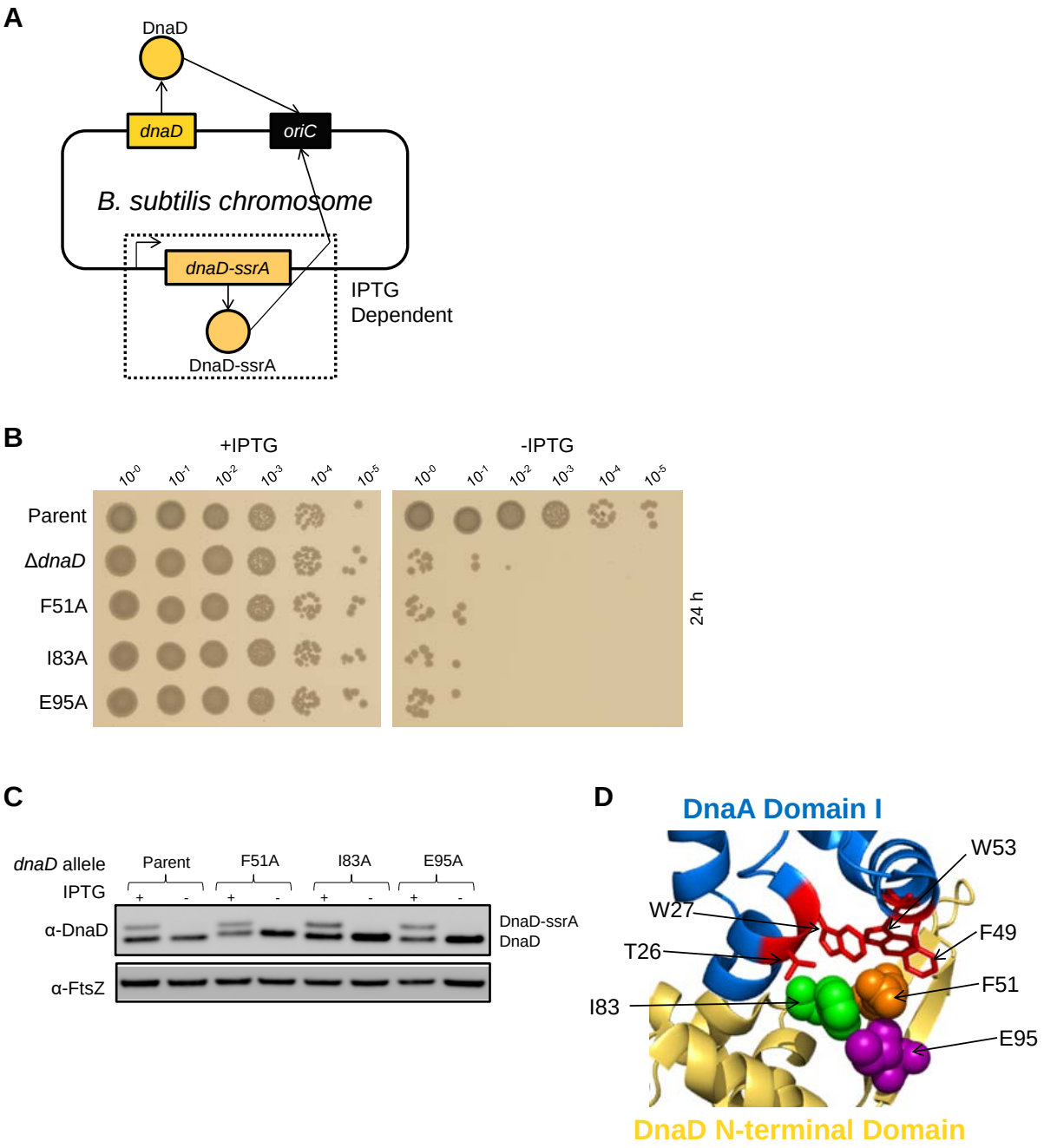


Figure 5.7. *In vivo* analysis of the DnaA interaction interface of DnaD. (A) Schematic of the chromosome of the *B. subtilis* strain capable of tolerating lethal mutations to the native *dnaD*. **(B)** Analysis of *B. subtilis* DnaD substitution mutants using the *dnaD-ssrA* strain. Growth is shown after 24 hours incubation at 37°C. **(C)** Immunoblot analysis of the DnaD substitution mutants with an FtsZ loading control. **(D)** Model for the interaction between DnaA domain I and the N-terminal domain of DnaD highlighting the implicated residues (DnaA domain I PDB ID 4TPS, DnaD^{NTD} PDB ID 2V79). Data collected by Charles Winterhalter (unpublished).

Parent (CW162), *ΔdnaD* (CW164), F51A (CW174), I83A (CW170), E95A (CW166).

5.5. SirA specifically interacts with DnaA and inhibits the DnaA-DnaD interaction

During times of low nutrient availability certain Gram-positive bacteria such as *B. subtilis* can undergo a process of differentiation known as sporulation, which ultimately results in the formation of a highly resistant endospore, a state these cells can remain in until conditions improve (Veening *et al.*, 2009). Spore development requires just two chromosomes and re-ination of DNA replication needs to be prevented (Section 1.6). To prevent this in *B. subtilis*, after committing to sporulation, the negative regulator of DNA replication initiation SirA is expressed (Rahn-Lee *et al.*, 2009).

It has been established how SirA binds to domain I of DnaA. Figure 5.8.A shows a comparison between the model for the DnaA-DnaD interaction from section 5.4 and the co-crystal structure of SirA bound to $DnaA^{DI}$ (Jameson *et al.*, 2014). As the comparison reveals, the key core DnaA residues of both interactions are practically identical. It has previously been speculated that SirA inhibits replication initiation by blocking the DnaA-DnaD interaction thereby preventing loading of the replicative helicase (Jameson *et al.*, 2014; Matthews and Simmons, 2019).

To determine whether SirA specifically interacts with only DnaA, a BTH was performed to investigate the interactions with all of the initiation proteins required for helicase recruitment and loading. The results revealed that SirA only interacts with DnaA and none of the other proteins of the *B. subtilis* helicase loading pathway (DnaD, DnaB, DnaI) suggesting SirA specifically targets DnaA (Figure 5.8.B).

To investigate if the mechanism for SirA's function is indeed to inhibit the DnaA-DnaD interaction, the construct shown in Figure 5.8.C was assembled to test for inhibition via bacterial two-hybrid. Here untagged *sirA* was placed downstream of the *dnaD*-adenylate cyclase fragment fusion, allowing for the expression of both the hybrid protein and SirA from the same plasmid using the same promoter. This vector was used to determine the effect of the presence of SirA on the DnaA-DnaD interaction. The result indicates that DnaA interacts with itself and SirA, as previously shown (Figure 5.8.B) and interacts with DnaD but only in the absence of SirA, as this interaction is lost when SirA is present (Figure 5.8.D). DnaD interacts with itself in the absence and presence of SirA, showing DnaD is still being expressed in the presence of SirA. This result suggests that SirA is inhibiting the interaction between DnaA and DnaD, presumably by binding to DnaA.

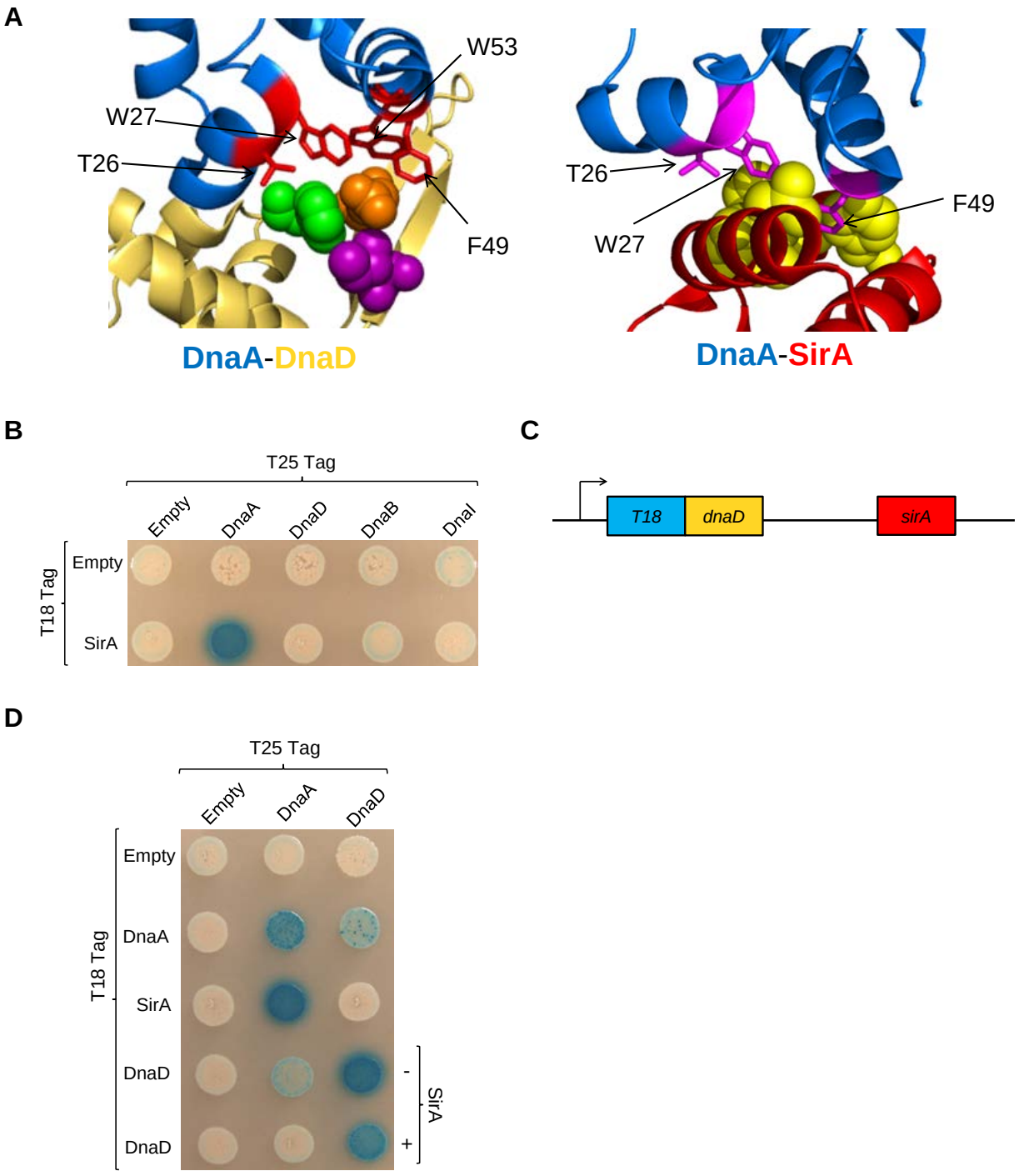


Figure 5.8. SirA inhibits the interaction between DnaA and DnaD. (A) Comparison between the model for the DnaA-DnaD interaction (from figure 5.5) with the co-crystal structure of $DnaA^{DI}$-SirA (PDB ID 4TPS). Only the residues forming the core of the DnaA-SirA interaction are highlighted. **(B)** Bacterial two-hybrid of SirA against DnaA, DnaD, DnaB and DnaI. **(C)** The construct created for expressing SirA and adenylate cyclase fragment fused DnaD from the same plasmid. **(D)** Bacterial two-hybrid showing DnaA and DnaD interacting with DnaA and SirA and also with DnaD in the presence and absence of untagged SirA. All adenylate cyclase fragment tagged proteins are done so N-terminally. T18 Empty (pUT18C), T18 SirA (pHM359), T18 DnaA (pHM638), T18 DnaD (pHM642), T25 Empty (pST25), T25 DnaA (pHM640), T25 DnaD (pHM644), T25 DnaB (pHM652), T25 DnaI (pHM656).

5.6. DnaA interacts with DnaB but does not interact with DnaI

As mentioned in the introduction to this chapter and section 1.5 it has been established that the proteins forming the helicase loading pathway in *B. subtilis* are recruited to the origin in a hierarchical order DnaA→DnaD→DnaB→DnaI (Figure 5.1.A). Previous bacterial two-hybrid assays have confirmed these interactions utilising individual protein domains. To investigate the interactions between full length proteins, the adapted BTH assay (Section 5.2) was employed. Genetic protein-adenylate cyclase fragment fusions were generated for each of the *B. subtilis* initiator proteins.

Evidence suggests that all *B. subtilis* initiator proteins assemble into homo-oligomeric complexes therefore each protein was initially assessed for its ability to self-interact. The results show that all the proteins are capable of self-interaction suggesting that all the proteins are functionally expressed in the two-hybrid system (Figure 5.9).

Beyond the self-interactions, the majority of the protein-protein interactions detected were as anticipated based on previous investigations (Matthews and Simmons, 2019). DnaA and DnaD interact, as demonstrated previously (Sections 5.3). DnaD and DnaB interact with each other as do DnaB and DnaI. Some of the interactions however, appear to be weaker than others, most notably the DnaA-DnaD interaction, and the DnaB-DnaI interaction (Figure 5.9).

Interestingly the two-hybrid assays detected a novel interaction not previously reported between DnaA and DnaB (stared yellow in Figure 5.9). Also of interest was that no interaction was detected between DnaA and the helicase loader DnaI (stared red in Figure 5.9). As discussed in section 1.3.5 investigations into the helicase loader of other bacteria appeared to show that the AAA+ domains of DnaA and the loader interacted in an ATP-dependent manner, but this either is not the case in *B. subtilis* or is undetectable by BTH.

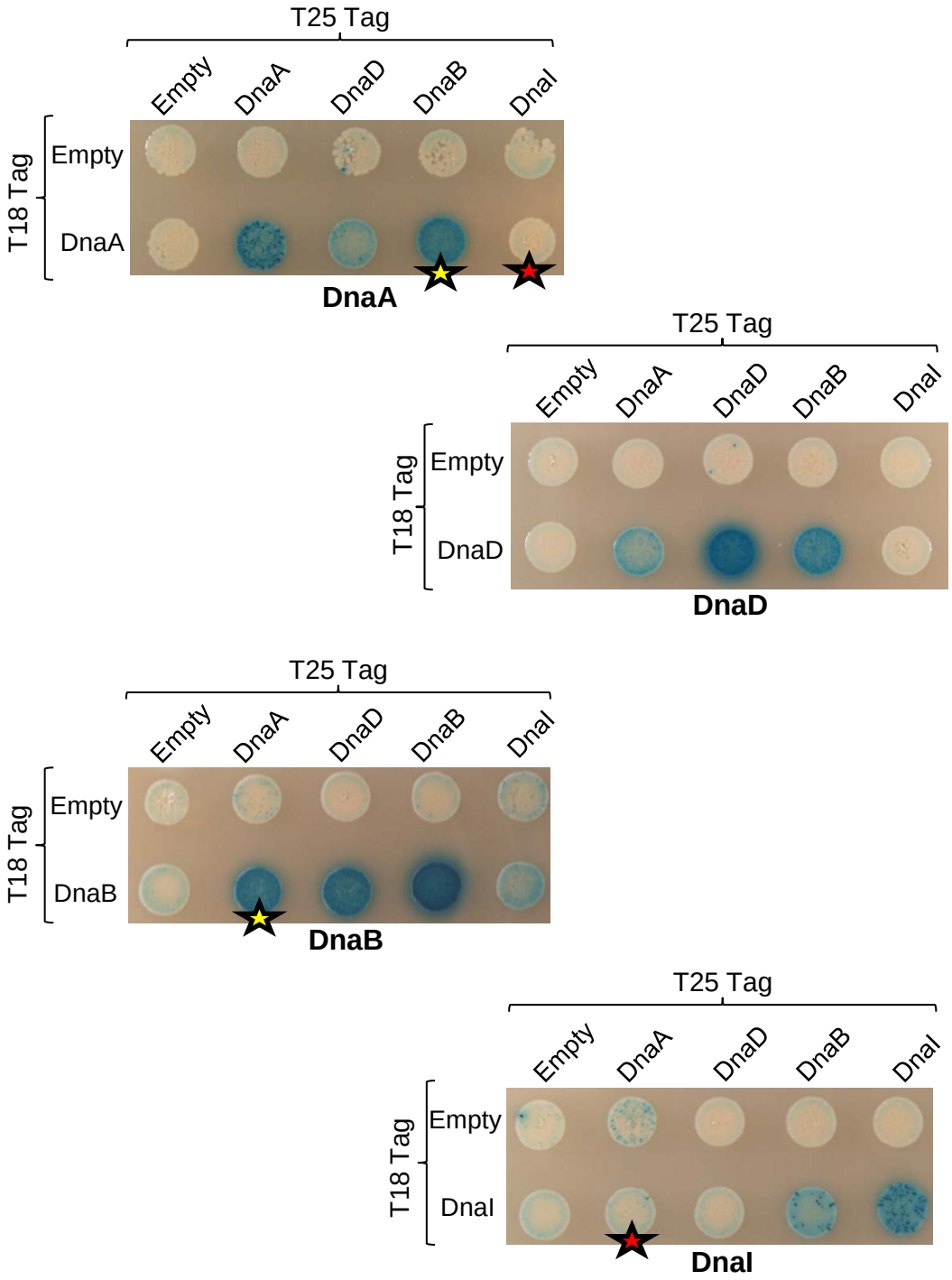


Figure 5.9. Interactions of the *B. subtilis* initiator proteins. Bacterial two-hybrid assays showing the interactions of full length DnaA, DnaD, DnaB and DnaI with each other. Interactions with empty vectors are shown as negative controls. All adenylate cyclase fragment tagged proteins are done so N-terminally. The DnaA-DnaB interaction is highlighted by a yellow star, with the DnaA-DnaI interaction indicated with a red star. T18 Empty (pUT18C), T18 DnaA (pHM638), T18 DnaD (pHM642), T18 DnaB (pHM650), T18 DnaI (pHM654), T25 Empty (pST25), T25 DnaA (pHM640), T25 DnaD (pHM644), T25 DnaB (pHM652), T25 DnaI (pHM656).

5.7. DnaB interacts with domain I of DnaA but uses a binding site distinct from the DnaD interface.

The results of the investigation into the interactions of the full length *B. subtilis* initiator proteins highlighted an unanticipated interaction between DnaA and DnaB (section 5.6). This interaction has presumably gone unappreciated as a sequential knockout, such as that performed by Smits *et al*., 2010, reveals only which proteins are no-longer associating with *oriC* as each protein is removed not individual protein:protein interactions.

As discussed throughout this chapter domain I of DnaA has been established as carrying the interaction surface for DnaD. $DnaA^{DI}$ has also been shown to be a protein interaction hub (section 1.3.5). To determine whether binding of DnaB to DnaA requires domain I, a BTH was performed using a truncated version of DnaA lacking domain I ($DnaA^{II-IV}$). $DnaA^{II-IV}$ interacts with both itself and full length DnaA, indicating it is functionally expressed (Figure 5.10.A). As shown previously, without domain I DnaA does not interact with DnaD. $DnaA^{II-IV}$ also does not interact with DnaB, suggesting that $DnaA^{DI}$ is required for the interaction of DnaA with DnaB (Figure 5.10.A).

Further BTHs were performed to determine if any of the essential residues identified within $DnaA^{DI}$ for the interaction with DnaD are also required for the interaction with DnaB. None of the residues required for the interaction with DnaD appear to be required for the interaction with DnaB, as the results of the BTH showed none of the protein variants, confirmed to be being expressed, lost the ability to interact with DnaB (Figure 5.10.B).

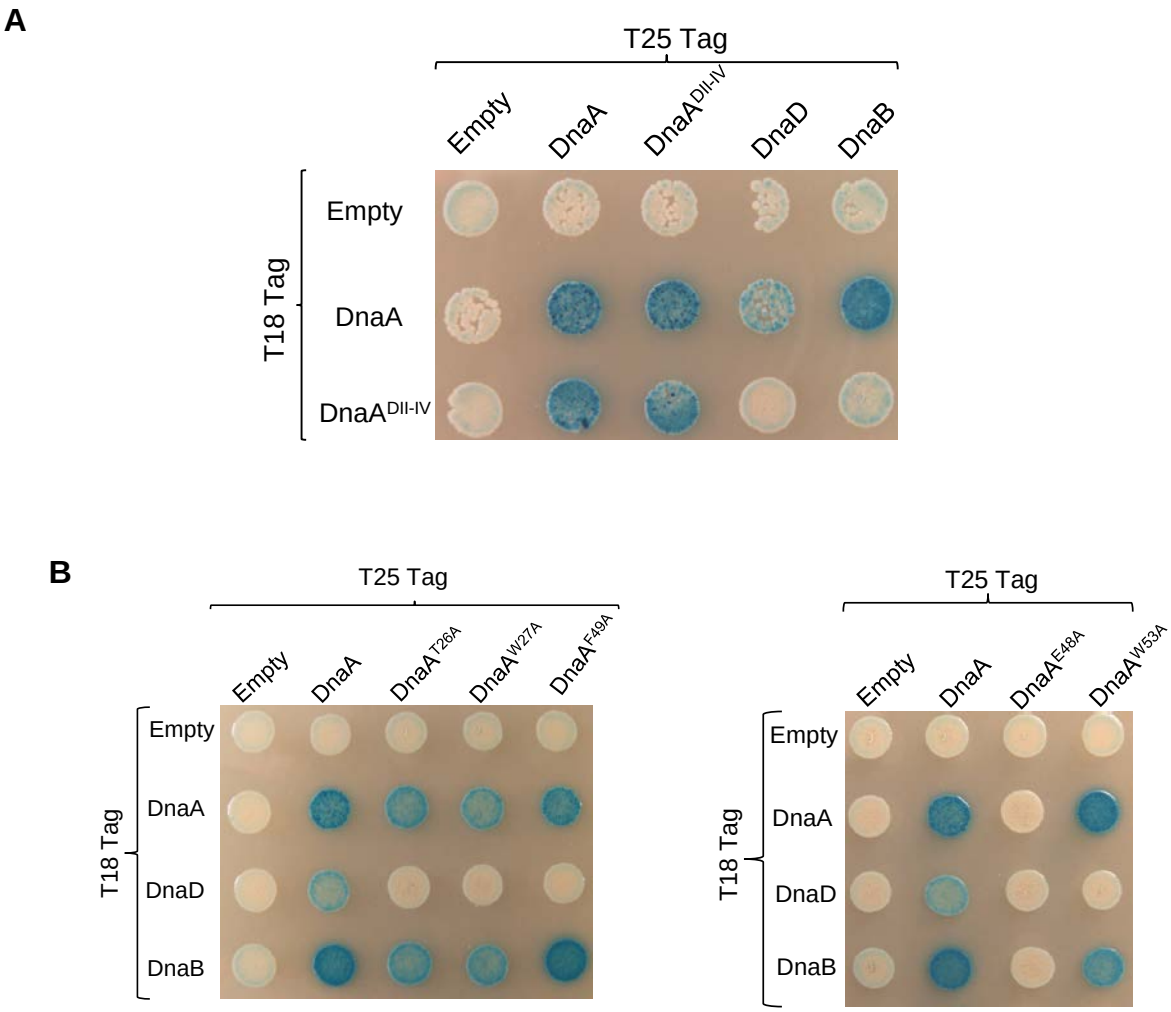


Figure 5.10. Interaction of DnaA with DnaB. (A) Bacterial two-hybrid of the interactions of full length DnaA, or DnaA lacking domain I (DnaA^{DII-IV}) with each other, DnaD and DnaB. **(B)** Bacterial two-hybrid assay showing the interaction between wild-type DnaA and the DnaA mutant variants with DnaA, DnaD and DnaB. All proteins are N-terminally tagged with the adenylate cyclase fragment.

T18 Empty (pUT18C), T18 DnaA (pHM638), T18 DnaA^{DII-IV} (pHM646), T18 DnaD (pHM642), T18 DnaB (pHM650), T25 Empty (pST25), T25 DnaA (pHM640), T25 DnaA^{DII-IV} (pHM648), T25 DnaD (pHM644), T25 DnaB (pHM652), T25 DnaA^{T26A} (pDS119), T25 DnaA^{W27A} (pDS120), T25 DnaA^{F49A} (pDS84), T25 DnaA^{E48A} (pDS87), T25 DnaA^{W53A} (pDS137).

Chapter 5 – Discussion

The DnaA-DnaD interaction

The interaction between DnaA and DnaD is the essential first step towards recruiting the replicative helicase during the initiation of DNA replication in the *Firmicutes* phylum, including *B. subtilis*. The molecular mechanism underpinning this key interaction is only starting to be understood and a couple of recent investigations (Martin *et al.*, 2019; Matthews and Simmons, 2019) proposed alternate mechanisms for this interaction.

The investigation performed here (Sections 5.1-5.5) reaffirms that residues within domain I of DnaA are required for the interaction with DnaD. For the first time this investigation shows that the DnaA residues involved with the interaction with DnaD are essential *in vivo* and required for the interaction between full length proteins. Interestingly the helicase interaction surface of *E. coli* DnaA domain I appears similar to the interaction surface identified in *B. subtilis* for the interaction with DnaD (*B. subtilis* F49 is homologous to *E. coli* F46 (Section 1.3.5). This suggests some level of conservation in the use of these residues as an interaction site for the mechanism of helicase recruitment, even with no direct conservation of the interaction partner.

While there was agreement that $DnaA^{DI}$ was the site of the interaction surface with DnaD there were alternate proposals for where the interaction surface was within DnaD. It was proposed that either the NTD of DnaD was the only domain interacting with DnaA, or that both domains interacted but with one interaction being weaker than the other (Figure 5.1.B). These proposals were investigated in Section 5.3 and the results presented there showed that the NTD of DnaD appeared to interact with DnaA stronger than the full length protein and this interaction required $DnaA^{DI}$. Furthermore, substitutions to N-terminal domain residues in the full length protein were sufficient to knockout the DnaA-DnaD interaction. These $DnaD^{NTD}$ residues were also all essential for a functional DnaD *in vivo*.

These results support the findings that the NTD of DnaD contains a surface that interacts with DnaA. Unfortunately the C-terminal domain of DnaD does not appear to be sufficiently expressed in the two-hybrid to draw unambiguous conclusions about its interactions. However, the result of the NTD residue substitutions within the BTH and *in vivo* supports the proposal that, at the very least, the N-terminal domain interaction is the strongest with DnaA, since the presence of the CTD was not sufficient to restore

the DnaA-DnaD interaction or retain viability when the NTD residues were substituted. If the CTD does interact with DnaA it likely plays a role in supporting or stabilising DnaD binding to DnaA.

SirA and the regulation of helicase loading during DNA replication initiation in *B. subtilis*

During sporulation the developmentally expressed inhibitor SirA binds to DnaA and inhibits initiation to prevent re-initiation in cells commited to sporulation (Section 1.6). As described in Section 5.5, SirA was previously shown to bind to domain I of DnaA, but the exact molecular mechanism behind its function remained unknown (Rahn-Lee *et al.*, 2011; Jameson *et al.*, 2014).

The interface between $DnaA^{DI}$ and the DnaD NTD is required for DnaD recruitment to *oriC*, the initial steps of *B. subtilis* helicase loading (Figure 5.11.A). Essential residues within $DnaA^{DI}$ have now been shown to be required for the interaction with DnaD (Section 5.3). These residues have also been shown to be required for the interaction with SirA (Jameson *et al.*, 2014; Matthews and Simmons, 2019) leading to the suggestion that SirA could function by inhibiting the DnaA-DnaD interaction.

The results in Section 5.5 were able to show for the first time that the expression of SirA inhibits the interaction of DnaA and DnaD. This result suggests that SirA inhibits initiation by blocking helicase loading. This blocking would be achieved through SirA binding domain I of DnaA occluding DnaD, thereby preventing its recruitment to *oriC* and consequently stopping the helicase recruitment pathway. These findings provide evidence for the first time for a mechanism by which SirA inhibits initiation, a model for which is outlined in Figure 5.11.B.

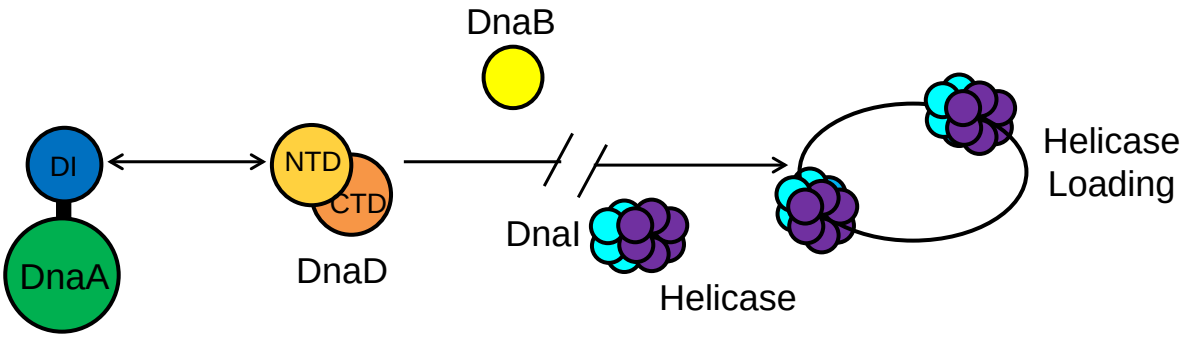
The results provide a mechanism for SirA function as well as revealing a system for developmentally regulating helicase loading in *B. subtilis.* In eukaryotes, as discussed in Section 1.1.1, the process of loading and activating the replicative helicase is distinct from the bacterial process. The mechanisms of regulating replication in eukaryotes is also distinct, involving regulating ORC proteins binding to the origin and also directly regulating helicase recruitment, loading and activation (Parker *et al.*, 2017). As discussed in section 1.6, the strategies bacteria employ to regulate initiation, and ensure replication occurs once per cell cycle, involve regulating DnaA *oriC* binding (e.g. SeqA), controlling DnaA protein levels in the cell or at the origin (e.g. the *datA*

locus), or through direct activation/inactivation of DnaA (e.g. DiaA, Hda) (Keyamura *et al.*, 2009; Katayama *et al.*, 2010; Nakamura and Katayama, 2010). Thus far direct regulation of helicase recruitment or loading has not been a strategy identified as being employed by bacteria.

The mechanism of initiation inhibition by SirA represents a first endogenous system identified for regulating bacterial helicase loading, during a specific developmental pathway. As outlined in the models in Figure 5.11 during normal growth conditions the DnaA-DnaD interaction occurs ultimately leading to helicase loading (Figure 5.11.A). However during differentiation for endospore formation and with the expression of SirA the DnaA-DnaD interaction would be inhibited resulting in the prevention of helicase loading (Figure 5.11.B).

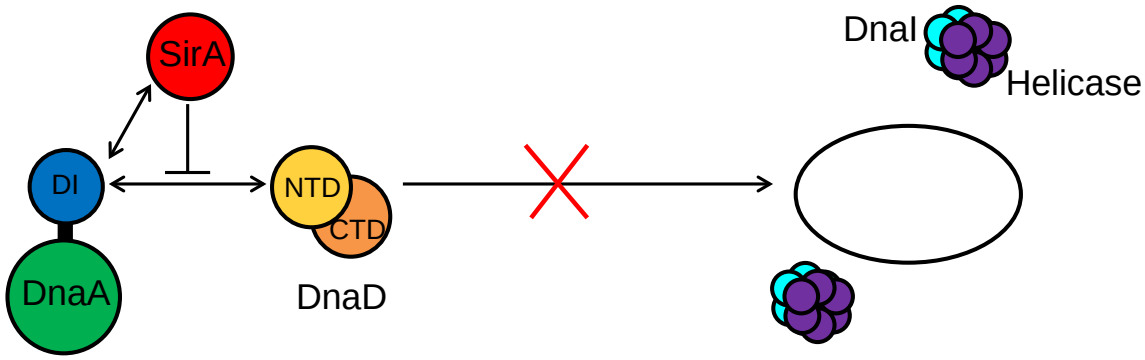
A

***Bacillus subtilis* helicase loading**



B

Regulation of helicase loading during *Bacillus subtilis* sporulation



5.11. Regulation of helicase loading in *Bacillus subtilis*. (A) Simplified schematic of the helicase loading pathway of *Bacillus subtilis* highlighting the key first interaction between domain I of DnaA and the N-terminal domain of DnaD. **(B)** Model for how the developmentally expressed SirA regulates helicase loading during *B. subtilis* DNA replication initiation.

The interactions of the *B. subtilis* initiator proteins

Previous investigations have established a helicase loading pathway in *B. subtilis* with the initiator proteins recruited to the origin in a hierarchical order DnaA→DnaD→ DnaB→DnaI (Figure 5.12.A). These interactions were confirmed using a bacterial two-hybrid assay using truncated proteins. Here an improved BTH was employed to investigate the interactions of the full length *B. subtilis* initiator proteins (section 5.6) and identified an unexpected interaction between DnaA and DnaB. If protein association requires multiple interaction partners, for example if DnaB is recruited by DnaA and DnaD, inactivation of either would lead to no observable recruitment.

Further investigation (Section 5.7) revealed that $DnaA^{DI}$ was required for the interaction with DnaB. None of the essential DnaA residues, involved in the interaction with DnaD, appear to be required for interacting with DnaB (Section 5.7). This result leaves open a few possibilities for the interaction between DnaA and DnaB which will require further investigation to determine the exact mechanism:

I) DnaB interacts with a different surface of $DnaA^{DI}$ than interacts with DnaD. If this is the case then $DnaA^{DI}$ presumably contains two separate surfaces for interacting with either protein.

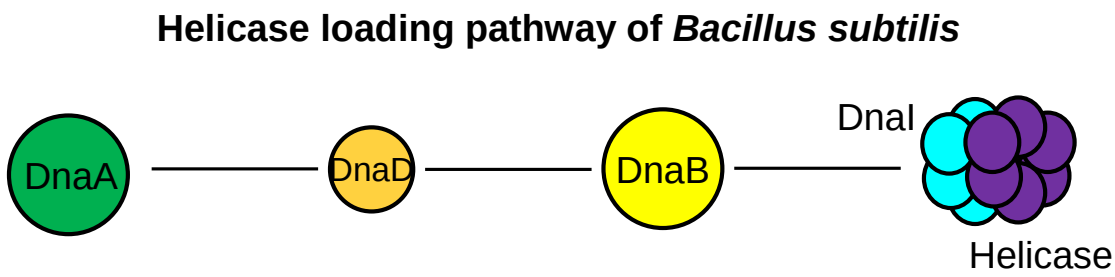
II) DnaB interacts with the NTD of DnaD using two domains (DnaB NTD and CTD) (Matthews and Simmons, 2019). DnaB could therefore be interacting with DnaA using two domains, each with an equally strong interaction. If any of the investigated residues are required for the interaction with DnaB they may only interact with one domain. The other domain may maintain the interaction making the requirement of the investigated residues undetectable.

If the DnaB binding to DnaA is essential, then the interaction could be serving a similar function to the speculated function of the interaction between DnaD and DnaB; recruiting the initiation complex at *oriC* to the membrane. A proposed alteration to the helicase loading pathway in *B. subtilis* incorporating this newly appreciated interaction is shown if Figure 5.12.B.

The BTH investigating the interactions of the *B. subtilis* initiator proteins was unable to detect an interaction between DnaA and the helicase loader DnaI. As discussed in section 1.3.5 it has been proposed that the AAA+ domains of DnaA and the helicase

loader interact, with it speculated that this interaction is utilised by the loader to regulate the recruitment and spatial positioning of the helicase onto the origin (Mott *et al.*, 2008). The initiation machinery of the bacteria from which this proposal originates (*A. aeolicus*) is formed of DnaA, the helicase and the helicase loader only. In such bacteria it has been established that DnaA directly binds the helicase (Kaguni, 2011), something that has not been demonstrated for *B. subtilis* DnaA. Therefore, it is possible that *Bacillus* DnaA does not interact with DnaI and that DnaB (which does interact with the helicase loader) fulfils the function of the DnaA-loader interaction. It is also possible that the DnaA-DnaI interaction is not detectable through heterologous expression. The speculated DnaA-helicase loader interaction requires ATP-bound DnaA suggesting the loader interacts with a DnaA filament. It is possible the BTH expressed DnaA is not adopting the correct conformation for interacting with DnaI, either through the presence of the adenylate-cyclase tag or by being unable to filament in *E. coli*, and therefore any possible DnaA-DnaI interaction is either not occurring or not detectable. It should also be noted that the proposed DnaA-helicase loader interaction is based on *in vitro* observations using a truncated version of loader lacking the N-terminal domain. Therefore, it is possible that the interaction is either not physiological or is inhibited by the presence of the N-terminal region of the full length protein.

A



B

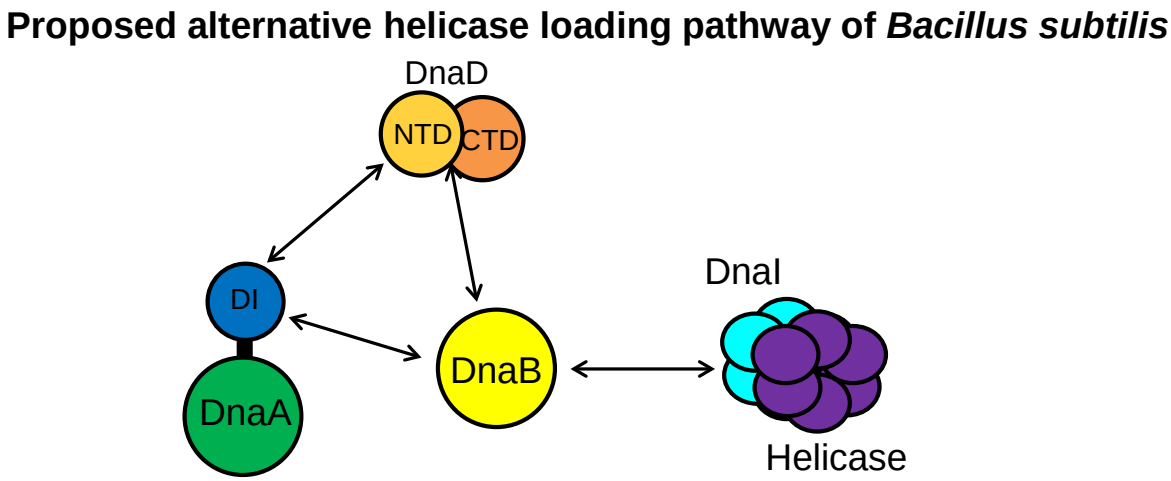


Figure 5.12. The helicase loading pathway of *Bacillus subtilis*. (A) The previously established helicase loading pathway of *B. subtilis*. **(B)** A proposed alternative helicase loading pathway for *B. subtilis* incorporating the newly appreciated DnaA-DnaB interaction and the interacting domains of DnaA and DnaD.

Chapter 6

General Discussion and Future Work

DNA replication is fundamental for all life and the work presented in this thesis has investigated how this process is initiated in the bacterium *Bacillus subtilis* to fill gaps in our understanding of these events.

Here I have specifically investigated the molecular mechanisms underpinning some of the essential activities of the master initiator protein DnaA, furthering our appreciation of some of the functions of DnaA and adding significantly to existing models. However, more work is required to answer the several outstanding questions.

6.1. Investigation of the upstream *incC* subregion and the proposed DNA loop

Previous investigation into the unwinding region of the *B. subtilis oriC* (*incC*) identified the minimal origin architecture required to support growth. This model includes two sub regions of DnaA-boxes, one proximal to the site of unwinding and one further upstream. The activities required by DnaA binding specifically to the distal subregion were determined *in vivo* in Chapter 3 utilising a chimeric DnaA system. Several activities required for origin opening were identified as essential, suggesting the protein binding here is directly involved in unwinding the DNA duplex. This result, combined with previous investigation, leads to the proposal that the role of the upstream region is to increase the local DnaA concentration at the site of unwinding. A DNA loop is the proposed mechanism of delivery (Richardson *et al.*, 2019).

There are a number of techniques that could be used to further investigate this model and provide further evidence of a DNA loop and the delivery of DnaA to the site of unwinding. One such technique would be electron microscopy which could be used to detect if the two subregions are required to form a DNA loop, in much the same way it was shown that both parts of the bipartite *B. subtilis* origin interact (Krause *et al.*, 1997). Another technique is 3C (Chromatin Conformation Capture) which can be used to investigate DNA looping. For this method cells are crosslinked to bind DNA segments in close proximity together. The DNA is fragmented, and crosslinked fragments ligated to form DNA hybrids. PCR is used to determine if specific DNA regions have been captured together (Dekker *et al.*, 2002). If the regions are usually distal to one another then being captured as a DNA hybrid would indicate a DNA loop is forming.

6.2. Investigating the residues implicated for a compact DnaA filament

It has been proposed that for DnaA to be competent for binding ssDNA the DnaA filament undergoes a conformational change with the DBD of one protomer docking against the AAA+ domain of another (Duderstadt *et al.*, 2010). The residues implicated in this DBD/AAA+ interaction were investigated in Chapter 3 where it was established the cells could tolerate alanine substitutions to any of these positions, although substitution with more dramatic amino acid changes was much less tolerable and caused severe growth defects.

The consideration from these findings was that the substituted residues causing the growth defects could be affecting duplex DNA binding (Fujikawa *et al.*, 2003) or an interaction with the helicase loader protein (Mott *et al.*, 2008). The investigation performed here was purely *in vivo* and could not differentiate between whether the growth defects are due to a disruption of a DBD/AAA+ interaction or due to disrupting another DnaA activity. Full differentiation between these activities could be achieved by investigation *in vitro*.

For investigation *in vitro* the variant proteins would first need to be purified, which could be achieved by use of the His-SUMO tag utilised in Chapter 4. To determine if duplex DNA or DnaA-box binding is being disrupted a similar crosslinking assay to that demonstrated in Chapter 4 (Sections 4.5 and 4.6) could be employed. Incubating the variant proteins with double-stranded or DnaA-box scaffolds would allow for the capturing of oligomeric species capable of binding such substrates. This could help to identify if any of the deleterious substitutions are affecting dsDNA or DnaA-box binding. Another technique would be an EMSA (electrophoretic mobility shift assay), widely used for investigating protein-DNA interactions. EMSA works on the principle that a protein-DNA complex will migrate slower than free DNA during gel electrophoresis. This would determine if the variant proteins are capable of binding a dsDNA substrate or not and such a technique has been used previously for investigating DnaA binding to origin DNA sequences (Richardson *et al.*, 2019).

Purified proteins could also be used to determine if the residues are involved in a DnaA-helicase loader interaction. This could be achieved by fusing the variant protein with a His-tag and using a pull down assay. If DnaA and the loader interact then when incubated together they should bind. Passing the complex through a HisTrap nickel column would capture the tagged DnaA and anything bound to it. If the substitutions

affect such an interaction then a DnaA-loader complex will not be captured, demonstrating if any of the non-tolerated substitutions are affecting a protein-protein interaction. Similar assays have been used for investigating DnaA interactions with various proteins including Hda (Keyamura and Katayama, 2011) and SirA (Jameson *et al.*, 2014).

If the variant proteins are still capable of binding DNA and are not affecting a protein-protein interaction then it leaves open the possibility that a DBD/AAA+ interaction is being disrupted. A final consideration was that the DBD/AAA+ interaction could be conditional and so not observed under the methods investigated here. This can be achieved by altering the growth conditions of the cells carrying the DnaA variants, such as growth at higher or lower temperatures and a range of environmental stresses (e.g. - salt conditions, sub-lethal antibiotic concentrations).

6.3. Investigation into specificity for the DnaA-trio motif

To unwind *oriC* DnaA engages and stretches a specific DNA strand, with the DnaA-trios providing the specific sequence to guide filament formation (Richardson *et al.*, 2016). Chapter 4 outlined how two isoleucine residues were determined to be required for forming filaments on ssDNA and unwinding the DNA duplex.

It remains to be concluded if these residues are required specifically for binding the DnaA-trio motif or for non-specific ssDNA binding. A number of assays could be performed to determine which the case is. One such assay would be to perform the crosslinking assay from sections 4.5 and 4.6 with single-stranded non-specific oligos. Capturing the oligomeric species capable of binding such substrates could help to identify if the two isoleucine residues are required for non-specific ssDNA binding. Other assays for investigating the interactions between various interaction partners including proteins and nucleic acids would be SPR (Surface Plasmon Resonance) or FP (Fluorescence Polarisation) (Anderson *et al.*, 2008; Wang *et al.*, 2015). Either assay would identify if the essential isoleucine residues are required for specific/non-specific ssDNA binding. Indeed fluorescence polarisation has previously been used to investigate DNA binding by variants of *A. aeolicus* DnaA (Duderstadt *et al.*, 2010).

If the two isoleucine residues are involved in non-specific ssDNA binding, then as discussed in Chapter 4, the mechanism for DnaA-trio recognition remains unidentified. It has been shown that only the AAA+ and DBD domains of DnaA are required for

recognising the DnaA-trio sequence (Richardson *et al.*, 2016) so further investigation into the residues located here could help identify whether others are potentially involved in recognising/binding DnaA-trios. Starting with alpha helices 5 and 6 of the AAA+ domain is logical as residues here have previously been implicated in ssDNA binding (Section 1.3.4) (Duderstadt *et al.*, 2011). Producing a crystal structure of DnaA bound specifically to DnaA-trios could also help identify a selection of residues. Any potential residues could be substituted using the *oriN* strain (Section 3.1) to determine physiological relevance and then investigated further using the *in vitro* assays utilised in Chapter 4 and those described above.

As touched upon in Chapter 4, it is possible specificity for the DnaA-trios arises from the specific structure of the DNA itself. Producing a crystal structure, as suggested above, could also reveal a distinctive structure for the Trios as specificity arising from DNA structure has been identified this way previously (Rohs *et al.*, 2009). If the DnaA:DnaA-Trios structural approach is unsuccessful NMR (Nuclear Magnetic Resonance) could be used to determine a high-resolution structure of the DnaA-trio repeats (Campagne *et al.*, 2011). This could help identify if the DnaA-trios are organising into a specific structure that could be recognised by DnaA.

Finally as discussed in chapter 4 it is possible that the two isoleucine residues are required for unwinding the DNA duplex by interfering with the hydrophobic cohesion stabilising the double helix through base-pair stacking. This proposal could be investigated in a number of ways. Firstly if hydrophobic residues are required at these positions could be investigated by making further amino acid substitutions to the isoleucine residues. Replacing them with similar sized hydrophobic (for example valine or leucine) or non-hydrophobic amino acids (such as serine or threonine) and determining viability for reveal the requirement for hydrophobic residues at these positions. If hydrophobic residues are all that are required it would provide support for the hypothesis. Lastly the DnaA-trio sequence itself could be investigated to see if base-stacking interference is the mechanism for unwinding. Looking bioinformatically at existing DnaA-trio sequences and the relative strength of the base-pairs in the stack would be one approach. Identifying if the base-pairs forming the end of the tri-nucleotides are the consistently weakest and so most easily separated would support the proposal. Also it could be investigated biochemically by altering the trio-sequence such that the stacking pattern is preserved but the nucleotide sequence is changed. If

such a substrate could still be unwound then it may provide further evidence that the mechanism of origin unwinding is via hydrophobic interactions destabilising the duplex.

6.4. Investigation of the proposed system for regulating helicase loading in *B. subtilis*

The work performed in Chapter 5 determined that DnaD and the developmentally expressed replication inhibitor SirA share a binding surface of $DnaA^{DI}$. It was also shown that the expression of SirA results in the loss of an interaction between DnaA and DnaD in a bacterial two-hybrid. This led to a proposed mechanism for SirA function to developmentally regulate helicase loading in *B. subtilis*. The model states that during sporulation the expression SirA leads to the inhibition of the DnaA-DnaD interaction, preventing recruitment of the loading complex and blocking helicase recruitment/deposition.

Further investigation is required to test this proposal. One such route of investigation would be to demonstrate the ability of SirA to inhibit DnaD recruitment to the origin and prevent helicase loading. This could be achieved through ChIP (chromatin immunoprecipitation) with an inducible copy of *sirA* integrated into the *B. subtilis* genome. ChIP would enable the demonstration that the initiator proteins and helicase are blocked from associating with the origin during the expression of SirA. Introducing suppressor mutations into either SirA or DnaA and seeing the effect this has on initiator recruitment to the origin during SirA expression could further support any findings.

The demonstration that SirA competes with DnaD and inhibits the interaction with DnaA used proteins expressed in a heterologous organism. Another route of investigation would be to test this *in vivo*. Expressing SirA in normally growing cells is lethal (Jameson *et al.*, 2014) and this lethality could potentially be rescued by overexpressing DnaD. If this is the case then it would demonstrate that SirA functions by outcompeting DnaD for DnaA binding. It may also be possible to demonstrate that SirA inhibits the DnaA-DnaD interaction *in vitro* using a similar pull down assay as described in section 6.2. Demonstrating that a DnaA-DnaD complex is captured using a His-tagged DnaA only when SirA is absent could be another way of showing that SirA inhibits the DnaA-DnaD interaction.

6.5. Investigating the interaction of DnaA with DnaB and DnaI

The helicase loading complex of *B. subtilis* has been demonstrated through previous ChIP investigations as being recruited in a hierarchical order DnaA→DnaD→ DnaB→DnaI. Chapter 5 reaffirmed these interactions but also identified an unexpected interaction between DnaA and DnaB. Also surprisingly, an interaction between DnaA and the helicase loader DnaI, which had been proposed to occur in homologs (Mott *et al.*, 2008), was not detected.

It was shown that $DnaA^{DI}$ is required for the interaction with DnaB (Section 5.7) and that the DnaA residues required for binding DnaD are not involved. To investigate DnaA-DnaB further, substitutions could be made to surface exposed residues within $DnaA^{DI}$ using the bacterial two-hybrid assay to identify any surfaces required for the interaction with DnaB. Any amino acids identified could be substituted *in vivo* using the *oriN* strain to determine the physiological relevance of a DnaA-DnaB interaction.

For DnaB, individual domains could be tested to see which are involved in the DnaA interaction. Surface exposed residues could then be substituted and used in the bacterial two-hybrid assay. Molecular modelling using already published structures could be a way of selecting residues which could be potentially involved in the interaction from either $DnaA^{DI}$ or DnaB. To determine the physiological relevance of any identified DnaB residues, a similar tool to that used for DnaD in section Figure 5.7 could be developed utilising an inducible ectopic copy of *dnaB*. As described for investigating protein-protein interactions in other sections, pull down assays and ChIP could both be utilised to further investigate the DnaA-DnaB interaction. Formaldehyde cross-linking followed by mass spectrometry would be another approach to identify the residues forming the DnaA-DnaB interaction interface.

A strategy to further investigate the putative DnaA-DnaI interaction was proposed in section 6.2 where it was suggested a pull down assay could be utilised. The DnaA-DnaI interaction was investigated through BTH in section 5.6 where no such interaction was observed. It was speculated that the interaction was not detectable as evidence suggests the loader interacts with a DnaA filament (Mott *et al.*, 2008), a conformation DnaA expressed in the BTH may not be adopting. Based on these observations investigating the DnaA-DnaI interaction further, potentially via a pull down assay (as proposed in section 6.2), may require DnaA oligomerisation to be promoted.

7. References

7.1. Structures

1J1V - (Fujikawa *et al.*, 2003)

1L8Q - (Erzberger *et al.*, 2002)

2E0G - (Abe *et al.*, 2007)

2HCB - (Erzberger *et al.*, 2006)

2K7R - (Loscha *et al.*, 2009)

2QBY - (Dueber *et al.*, 2007)

2V1U - (Gaudier *et al.*, 2007)

2V79 - (Schneider *et al.*, 2008)

2W58 - (Tsai *et al.*, 2009)

4TPS - (Jameson *et al.*, 2014)

5WTN - (Li *et al.*, 2017)

7.2. References

Abe, Y., Jo, T., Matsuda, Y., Matsunaga, C., Katayama, T. and Ueda, T. (2007) 'Structure and function of DnaA N-terminal domains: specific sites and mechanisms in inter-DnaA interaction and in DnaB helicase loading on oriC', *J Biol Chem*, 282(24), pp. 17816-27.

Anderson, B.J., Larkin, C., Guja, K. and Schildbach, J.F. (2008) 'Using fluorophore-labeled oligonucleotides to measure affinities of protein-DNA interactions', *Methods in enzymology*, 450, pp. 253-272.

Arias-Palomo, E., Puri, N., O'Shea Murray, V.L., Yan, Q. and Berger, J.M. (2019) 'Physical Basis for the Loading of a Bacterial Replicative Helicase onto DNA', *Mol Cell*, 74(1), pp. 173-184.e4.

Ausiannikava, D., Mitchell, L., Marriott, H., Smith, V., Hawkins, M., Makarova, K.S., Koonin, E.V., Nieduszynski, C.A. and Allers, T. (2018) 'Evolution of Genome Architecture in Archaea: Spontaneous Generation of a New Chromosome in Haloferax volcanii', *Molecular Biology and Evolution*, 35(8), pp. 1855-1868.

Badrinarayanan, A., Le, T.B.K. and Laub, M.T. (2015) 'Bacterial chromosome organization and segregation', *Annual review of cell and developmental biology*, 31, pp. 171-199.

Battesti, A. and Bouveret, E. (2012) 'The bacterial two-hybrid system based on adenylate cyclase reconstitution in Escherichia coli', *Methods*, 58(4), pp. 325-34.

Bell, S.P. and Kaguni, J.M. (2013) 'Helicase loading at chromosomal origins of replication', *Cold Spring Harbor perspectives in biology*, 5(6), p. a010124.

Berghuis, B.A., Raducanu, V.-S., Elshenawy, M.M., Jergic, S., Depken, M., Dixon, N.E., Hamdan, S.M. and Dekker, N.H. (2018) 'What is all this fuss about Tus? Comparison of recent findings from biophysical and biochemical experiments', *Critical Reviews in Biochemistry and Molecular Biology*, 53(1), pp. 49-63.

Bern, M. and Goldberg, D. (2005) 'Automatic selection of representative proteins for bacterial phylogeny', *BMC Evol Biol*, 5, p. 34.

Biswas-Fiss, E.E., Kukiratirat, J. and Biswas, S.B. (2012) 'Thermodynamic analysis of DNA binding by a Bacillus single stranded DNA binding protein', *BMC biochemistry*, 13, pp. 10-10.

Bleichert, F., Leitner, A., Aebersold, R., Botchan, M.R. and Berger, J.M. (2018) 'Conformational control and DNA-binding mechanism of the metazoan origin recognition complex', *Proceedings of the National Academy of Sciences*, 115(26), pp. E5906-E5915.

Briggs, G.S., Smits, W.K. and Soultanas, P. (2012) 'Chromosomal replication initiation machinery of low-G+C-content Firmicutes', *J Bacteriol*, 194(19), pp. 5162-70.

Burkholder, W.F., Kurtser, I. and Grossman, A.D. (2001) 'Replication Initiation Proteins Regulate a Developmental Checkpoint in Bacillus subtilis', *Cell*, 104(2), pp. 269-279.

Campagne, S., Gervais, V. and Milon, A. (2011) 'Nuclear magnetic resonance analysis of protein-DNA interactions', *Journal of the Royal Society, Interface*, 8(61), pp. 1065-1078.

Castilla-Llorente, V., Munoz-Espin, D., Villar, L., Salas, M. and Meijer, W.J. (2006) 'Spo0A, the key transcriptional regulator for entrance into sporulation, is an inhibitor of DNA replication', *Embo j*, 25(16), pp. 3890-9.

Cayrou, C., Coulombe, P., Puy, A., Rialle, S., Kaplan, N., Segal, E. and Méchali, M. (2012) 'New insights into replication origin characteristics in metazoans', *Cell Cycle*, 11(4), pp. 658-667.

Cheng, H.M., Groger, P., Hartmann, A. and Schlierf, M. (2015) 'Bacterial initiators form dynamic filaments on single-stranded DNA monomer by monomer', *Nucleic Acids Res*, 43(1), pp. 396-405.

Cho, E., Ogasawara, N. and Ishikawa, S. (2008) 'The functional analysis of YabA, which interacts with DnaA and regulates initiation of chromosome replication in Bacillus subtils', *Genes Genet Syst*, 83(2), pp. 111-25.

Chodavarapu, S., Felczak, M.M., Yaniv, J.R. and Kaguni, J.M. (2008a) 'Escherichia coli DnaA interacts with HU in initiation at the E. coli replication origin', *Mol Microbiol*, 67(4), pp. 781-92.

Chodavarapu, S., Gomez, R., Vicente, M. and Kaguni, J.M. (2008b) 'Escherichia coli Dps interacts with DnaA protein to impede initiation: a model of adaptive mutation', *Mol Microbiol*, 67(6), pp. 1331-46.

Coman, D. and Russu, I.M. (2005) 'Base pair opening in three DNA-unwinding elements', *J Biol Chem*, 280(21), pp. 20216-21.

Creager, R.L., Li, Y. and MacAlpine, D.M. (2015) 'SnapShot: Origins of DNA replication', *Cell*, 161(2), pp. 418-418.e1.

Dao, F.-Y., Lv, H., Wang, F., Feng, C.-Q., Ding, H., Chen, W. and Lin, H. (2018) 'Identify origin of replication in Saccharomyces cerevisiae using two-step feature selection technique', *Bioinformatics*, 35(12), pp. 2075-2083.

Dekker, J., Rippe, K., Dekker, M. and Kleckner, N. (2002) 'Capturing chromosome conformation', *Science*, 295(5558), pp. 1306-11.

Dewar, J.M. and Walter, J.C. (2017) 'Mechanisms of DNA replication termination', *Nature reviews. Molecular cell biology*, 18(8), pp. 507-516.

Dhingra, N. and Kaplan, D.L. (2016) 'Introduction to Eukaryotic DNA Replication Initiation', in Kaplan, D.L. (ed.) *The Initiation of DNA Replication in Eukaryotes*. Cham: Springer International Publishing, pp. 1-21.

diCenzo, G.C. and Finan, T.M. (2017) 'The Divided Bacterial Genome: Structure, Function, and Evolution', *Microbiol Mol Biol Rev*, 81(3).

Dimude, J.U., Midgley-Smith, S.L., Stein, M. and Rudolph, C.J. (2016) 'Replication Termination: Containing Fork Fusion-Mediated Pathologies in Escherichia coli', *Genes*, 7(8), p. 40.

Duderstadt, K.E. and Berger, J.M. (2013) 'A structural framework for replication origin opening by AAA+ initiation factors', *Curr Opin Struct Biol*, 23(1), pp. 144-53.

Duderstadt, K.E., Chuang, K. and Berger, J.M. (2011) 'DNA stretching by bacterial initiators promotes replication origin opening', *Nature*, 478(7368), pp. 209-13.

Duderstadt, K.E., Mott, M.L., Crisona, N.J., Chuang, K., Yang, H. and Berger, J.M. (2010) 'Origin remodeling and opening in bacteria rely on distinct assembly states of the DnaA initiator', *J Biol Chem*, 285(36), pp. 28229-39.

Dueber, E.L., Corn, J.E., Bell, S.D. and Berger, J.M. (2007) 'Replication origin recognition and deformation by a heterodimeric archaeal Orc1 complex', *Science*, 317(5842), pp. 1210-3.

Duggin, I.G., Wake, R.G., Bell, S.D. and Hill, T.M. (2008) 'The replication fork trap and termination of chromosome replication', *Mol Microbiol*, 70(6), pp. 1323-33.

Duigou, S., Yamaichi, Y. and Waldor, M.K. (2008) 'ATP negatively regulates the initiator protein of Vibrio cholerae chromosome II replication', *Proc Natl Acad Sci U S A*, 105(30), pp. 10577-82.

Ekundayo, B. and Bleichert, F. (2019) 'Origins of DNA replication', *PLOS Genetics*, 15(9), p. e1008320.

Erzberger, J.P. and Berger, J.M. (2006) 'Evolutionary relationships and structural mechanisms of AAA+ proteins', *Annu Rev Biophys Biomol Struct*, 35, pp. 93-114.

Erzberger, J.P., Mott, M.L. and Berger, J.M. (2006) 'Structural basis for ATP-dependent DnaA assembly and replication-origin remodeling', *Nat Struct Mol Biol*, 13(8), pp. 676-83.

Erzberger, J.P., Pirruccello, M.M. and Berger, J.M. (2002) 'The structure of bacterial DnaA: implications for general mechanisms underlying DNA replication initiation', *Embo j*, 21(18), pp. 4763-73.

Feng, B., Sosa, R.P., Martensson, A.K.F., Jiang, K., Tong, A., Dorfman, K.D., Takahashi, M., Lincoln, P., Bustamante, C.J., Westerlund, F. and Norden, B. (2019) 'Hydrophobic catalysis and a potential biological role of DNA unstacking induced by environment effects', *Proc Natl Acad Sci U S A*, 116(35), pp. 17169-17174.

Fournes, F., Val, M.-E., Skovgaard, O. and Mazel, D. (2018) 'Replicate Once Per Cell Cycle: Replication Control of Secondary Chromosomes', *Frontiers in microbiology*, 9, pp. 1833-1833.

Frandi, A. and Collier, J. (2019) 'Multilayered control of chromosome replication in Caulobacter crescentus', *Biochemical Society Transactions*, p. BST20180460.

Fujikawa, N., Kurumizaka, H., Nureki, O., Terada, T., Shirouzu, M., Katayama, T. and Yokoyama, S. (2003) 'Structural basis of replication origin recognition by the DnaA protein', *Nucleic Acids Res*, 31(8), pp. 2077-86.

Fuller, R.S., Funnell, B.E. and Kornberg, A. (1984) 'The dnaA protein complex with the E. coli chromosomal replication origin (oriC) and other DNA sites', *Cell*, 38(3), pp. 889-900.

Gaudier, M., Schuwirth, B.S., Westcott, S.L. and Wigley, D.B. (2007) 'Structural basis of DNA replication origin recognition by an ORC protein', *Science*, 317(5842), pp. 1213-6.

Gowrishankar, J. (2015) 'End of the beginning: elongation and termination features of alternative modes of chromosomal replication initiation in bacteria', *PLoS Genet*, 11(1), p. e1004909.

Hansen, F.G. and Atlung, T. (2018) 'The DnaA Tale', *Frontiers in Microbiology*, 9, p. 319.

Hassan, A.K., Moriya, S., Ogura, M., Tanaka, T., Kawamura, F. and Ogasawara, N. (1997) 'Suppression of initiation defects of chromosome replication in Bacillus subtilis dnaA and oriC-deleted mutants by integration of a plasmid replicon into the chromosomes', *J Bacteriol*, 179(8), pp. 2494-502.

Hawkins, M., Malla, S., Blythe, M.J., Nieduszynski, C.A. and Allers, T. (2013) 'Accelerated growth in the absence of DNA replication origins', *Nature*, 503(7477), pp. 544-547.

Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwinn, M.L., Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Umayam, L., Gill, S.R., Nelson, K.E., Read, T.D., Tettelin, H., Richardson, D., Ermolaeva, M.D., Vamathevan, J., Bass, S., Qin, H., Dragoi, I., Sellers, P., McDonald, L., Utterback, T., Fleishmann, R.D., Nierman, W.C., White, O., Salzberg, S.L., Smith, H.O., Colwell, R.R., Mekalanos, J.J., Venter, J.C. and Fraser, C.M. (2000) 'DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae', *Nature*, 406(6795), pp. 477-83.

Herskovits, T.T. (1962) 'Nonaqueous solutions of DNA: Factors determining the stability of the helical configuration in solution', *Archives of Biochemistry and Biophysics*, 97(3), pp. 474-484.

Higgins, D. and Dworkin, J. (2012) 'Recent progress in Bacillus subtilis sporulation', *FEMS microbiology reviews*, 36(1), pp. 131-148.

Horiike, T., Miyata, D., Hamada, K., Saruhashi, S., Shinozawa, T., Kumar, S., Chakraborty, R., Komiyama, T. and Tateno, Y. (2009) 'Phylogenetic construction of 17 bacterial phyla by new method and carefully selected orthologs', *Gene*, 429(1-2), pp. 59-64.

Iismaa, T.P. and Wake, R.G. (1987) 'The normal replication terminus of the Bacillus subtilis chromosome, terC, is dispensable for vegetative growth and sporulation', *J Mol Biol*, 195(2), pp. 299-310.

Ioannou, C., Schaeffer, P.M., Dixon, N.E. and Soultanas, P. (2006) 'Helicase binding to DnaI exposes a cryptic DNA-binding site during helicase loading in Bacillus subtilis', *Nucleic Acids Res*, 34(18), pp. 5247-58.

Ishigo-Oka, D., Ogasawara, N. and Moriya, S. (2001) 'DnaD protein of Bacillus subtilis interacts with DnaA, the initiator protein of replication', *Journal of bacteriology*, 183(6), pp. 2148-2150.

Jameson, K.H., Rostami, N., Fogg, M.J., Turkenburg, J.P., Grahl, A., Murray, H. and Wilkinson, A.J. (2014) 'Structure and interactions of the Bacillus subtilis sporulation inhibitor of DNA replication, SirA, with domain I of DnaA', *Mol Microbiol*, 93(5), pp. 975-91.

Jha, J.K., Ramachandran, R. and Chattoraj, D.K. (2016) 'Opening the Strands of Replication Origins-Still an Open Question', *Front Mol Biosci*, 3, p. 62.

Jonas, K., Liu, J., Chien, P. and Laub, M.T. (2013) 'Proteotoxic stress induces a cell-cycle arrest by stimulating Lon to degrade the replication initiator DnaA', *Cell*, 154(3), pp. 623-36.

Kaguni, J.M. (2011) 'Replication initiation at the Escherichia coli chromosomal origin', *Curr Opin Chem Biol*, 15(5), pp. 606-13.

Kang, S., Lee, H., Han, J.S. and Hwang, D.S. (1999) 'Interaction of SeqA and Dam methylase on the hemimethylated origin of Escherichia coli chromosomal DNA replication', *J Biol Chem*, 274(17), pp. 11463-8.

Karimova, G., Pidoux, J., Ullmann, A. and Ladant, D. (1998) 'A bacterial two-hybrid system based on a reconstituted signal transduction pathway', *Proceedings of the National Academy of Sciences of the United States of America*, 95(10), pp. 5752-5756.

Katayama, T., Ozaki, S., Keyamura, K. and Fujimitsu, K. (2010) 'Regulation of the replication cycle: conserved and diverse regulatory systems for DnaA and oriC', *Nat Rev Microbiol*, 8(3), pp. 163-70.

Kawakami, H., Keyamura, K. and Katayama, T. (2005) 'Formation of an ATP-DnaA-specific initiation complex requires DnaA Arginine 285, a conserved motif in the AAA+ protein family', *J Biol Chem*, 280(29), pp. 27420-30.

Keyamura, K., Abe, Y., Higashi, M., Ueda, T. and Katayama, T. (2009) 'DiaA dynamics are coupled with changes in initial origin complexes leading to helicase loading', *J Biol Chem*, 284(37), pp. 25038-50.

Keyamura, K., Fujikawa, N., Ishida, T., Ozaki, S., Su'etsugu, M., Fujimitsu, K., Kagawa, W., Yokoyama, S., Kurumizaka, H. and Katayama, T. (2007) 'The interaction of DiaA and DnaA regulates the replication cycle in E. coli by directly promoting ATP DnaA-specific initiation complexes', *Genes Dev*, 21(16), pp. 2083-99.

Keyamura, K. and Katayama, T. (2011) 'DnaA protein DNA-binding domain binds to Hda protein to promote inter-AAA+ domain interaction involved in regulatory inactivation of DnaA', *The Journal of biological chemistry*, 286(33), pp. 29336-29346.

Kitagawa, R., Ozaki, T., Moriya, S. and Ogawa, T. (1998) 'Negative control of replication initiation by a novel chromosomal locus exhibiting exceptional affinity for Escherichia coli DnaA protein', *Genes Dev*, 12(19), pp. 3032-43.

Kogoma, T. and von Meyenburg, K. (1983) 'The origin of replication, oriC, and the dnaA protein are dispensable in stable DNA replication (sdrA) mutants of Escherichia coli K-12', *The EMBO journal*, 2(3), pp. 463-468.

Krause, M. and Messer, W. (1999) 'DnaA proteins of Escherichia coli and Bacillus subtilis: coordinate actions with single-stranded DNA-binding protein and interspecies inhibition during open complex formation at the replication origins', *Gene*, 228(1-2), pp. 123-32.

Krause, M., Ruckert, B., Lurz, R. and Messer, W. (1997) 'Complexes at the replication origin of Bacillus subtilis with homologous and heterologous DnaA protein', *J Mol Biol*, 274(3), pp. 365-80.

Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G., Azevedo, V., Bertero, M.G., Bessieres, P., Bolotin, A., Borchert, S., Borriss, R., Boursier, L., Brans, A., Braun, M., Brignell, S.C., Bron, S., Brouillet, S., Bruschi, C.V., Caldwell, B., Capuano, V., Carter, N.M., Choi, S.K., Cordani, J.J., Connerton, I.F., Cummings, N.J., Daniel, R.A., Denziot, F., Devine, K.M., Dusterhoft, A., Ehrlich, S.D., Emmerson, P.T., Entian, K.D., Errington, J., Fabret, C., Ferrari, E., Foulger, D., Fritz, C., Fujita, M., Fujita, Y., Fuma, S., Galizzi, A., Galleron, N., Ghim, S.Y., Glaser, P., Goffeau, A., Golightly, E.J., Grandi, G., Guiseppi, G., Guy, B.J., Haga, K., Haiech, J., Harwood, C.R., Henaut, A., Hilbert, H., Holsappel, S., Hosono, S., Hullo, M.F., Itaya, M., Jones, L., Joris, B., Karamata, D., Kasahara, Y., Klaerr-Blanchard, M., Klein, C., Kobayashi, Y., Koetter, P., Koningstein, G., Krogh, S., Kumano, M., Kurita, K., Lapidus, A., Lardinois, S., Lauber, J., Lazarevic, V., Lee, S.M., Levine, A., Liu, H., Masuda, S., Mauel, C., Medigue, C., Medina, N., Mellado, R.P., Mizuno, M., Moestl, D., Nakai, S., Noback, M., Noone, D., O'Reilly, M., Ogawa, K., Ogiwara, A., Oudega, B., Park, S.H., Parro, V., Pohl, T.M., Portelle, D., Porwollik, S., Prescott, A.M., Presecan, E., Pujic, P., Purnelle, B., et al. (1997) 'The complete genome sequence of the gram-positive bacterium Bacillus subtilis', *Nature*, 390(6657), pp. 249-56.

Lee, Y.-S., Kennedy, W.D. and Yin, Y.W. (2009) 'Structural Insight into Processive Human Mitochondrial DNA Synthesis and Disease-Related Polymerase Mutations', *Cell*, 139(2), pp. 312-324.

Leonard, A.C. and Méchali, M. (2013) 'DNA replication origins', *Cold Spring Harb Perspect Biol*, 5(10), p. a010116.

Li, N., Lam, W.H., Zhai, Y., Cheng, J., Cheng, E., Zhao, Y., Gao, N. and Tye, B.K. (2018) 'Structure of the origin recognition complex bound to DNA replication origin', *Nature*, 559(7713), pp. 217-222.

Li, Y.C., Naveen, V., Lin, M.G. and Hsiao, C.D. (2017) 'Structural analyses of the bacterial primosomal protein DnaB reveal that it is a tetramer and forms a complex with a primosomal re-initiation protein', *J Biol Chem*, 292(38), pp. 15744-15757.

Lindås, A.-C. and Bernander, R. (2013) 'The cell cycle of archaea', *Nature Reviews Microbiology*, 11, p. 627.

Lombrana, R., Almeida, R., Alvarez, A. and Gomez, M. (2015) 'R-loops and initiation of DNA replication in human cells: a missing link?', *Front Genet*, 6, p. 158.

Loscha, K.V., Jaudzems, K., Ioannou, C., Su, X.C., Hill, F.R., Otting, G., Dixon, N.E. and Liepinsh, E. (2009) 'A novel zinc-binding fold in the helicase interaction domain of the Bacillus subtilis DnaI helicase loader', *Nucleic Acids Res*, 37(7), pp. 2395-404.

Luscombe, N.M., Laskowski, R.A. and Thornton, J.M. (2001) 'Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level', *Nucleic Acids Res*, 29(13), pp. 2860-74.

Manna, A.C., Pai, K.S., Bussiere, D.E., White, S.W. and Bastia, D. (1996) 'The dimer-dimer interaction surface of the replication terminator protein of Bacillus subtilis and termination of DNA replication', *Proceedings of the National Academy of Sciences*, 93(8), pp. 3253-3258.

Marston, F.Y., Grainger, W.H., Smits, W.K., Hopcroft, N.H., Green, M., Hounslow, A.M., Grossman, A.D., Craven, C.J. and Soultanas, P. (2010) 'When simple sequence comparison fails: the cryptic case of the shared domains of the bacterial replication initiation proteins DnaB and DnaD', *Nucleic Acids Res*, 38(20), pp. 6930-42.

Martin, E., Williams, H.E.L., Pitoulias, M., Stevens, D., Winterhalter, C., Craggs, T.D., Murray, H., Searle, M.S. and Soultanas, P. (2019) 'DNA replication initiation in Bacillus subtilis: structural and functional characterization of the essential DnaA-DnaD interaction', *Nucleic Acids Res*, 47(4), pp. 2101-2112.

Matthews, L.A. and Simmons, L.A. (2019) 'Cryptic protein interactions regulate DNA replication initiation', *Mol Microbiol*, 111(1), pp. 118-130.

Miller, D.T., Grimwade, J.E., Betteridge, T., Rozgaja, T., Torgue, J.J. and Leonard, A.C. (2009) 'Bacterial origin recognition complexes direct assembly of higher-order DnaA oligomeric structures', *Proc Natl Acad Sci U S A*, 106(44), pp. 18479-84.

Mirkin, E.V. and Mirkin, S.M. (2007) 'Replication fork stalling at natural impediments', *Microbiology and molecular biology reviews : MMBR*, 71(1), pp. 13-35.

Mott, M.L. and Berger, J.M. (2007) 'DNA replication initiation: mechanisms and regulation in bacteria', *Nat Rev Microbiol*, 5(5), pp. 343-54.

Mott, M.L., Erzberger, J.P., Coons, M.M. and Berger, J.M. (2008) 'Structural synergy and molecular crosstalk between bacterial helicase loaders and replication initiators', *Cell*, 135(4), pp. 623-34.

Mulcair, M.D., Schaeffer, P.M., Oakley, A.J., Cross, H.F., Neylon, C., Hill, T.M. and Dixon, N.E. (2006) 'A molecular mousetrap determines polarity of termination of DNA replication in E. coli', *Cell*, 125(7), pp. 1309-19.

Mulugu, S., Potnis, A., Shamsuzzaman, Taylor, J., Alexander, K. and Bastia, D. (2001) 'Mechanism of termination of DNA replication of Escherichia coli involves helicase–contrahelicase interaction', *Proceedings of the National Academy of Sciences*, 98(17), pp. 9569-9574.

Murray, H. and Koh, A. (2014) 'Multiple regulatory systems coordinate DNA replication with cell growth in Bacillus subtilis', *PLoS Genet*, 10(10), p. e1004731.

Nakamura, K. and Katayama, T. (2010) 'Novel essential residues of Hda for interaction with DnaA in the regulatory inactivation of DnaA: unique roles for Hda AAA Box VI and VII motifs', *Mol Microbiol*, 76(2), pp. 302-17.

Neylon, C., Kralicek, A.V., Hill, T.M. and Dixon, N.E. (2005) 'Replication Termination in Escherichia coli: Structure and Antihelicase Activity of the Tus-Ter Complex', *Microbiology and Molecular Biology Reviews*, 69(3), pp. 501-526.

Noguchi, Y., Sakiyama, Y., Kawakami, H. and Katayama, T. (2015) 'The Arg Fingers of Key DnaA Protomers Are Oriented Inward within the Replication Origin oriC and Stimulate DnaA Subcomplexes in the Initiation Complex', *J Biol Chem*, 290(33), pp. 20295-312.

O'Donnell, M., Langston, L. and Stillman, B. (2013) 'Principles and concepts of DNA replication in bacteria, archaea, and eukarya', *Cold Spring Harb Perspect Biol*, 5(7).

Ohbayashi, R., Watanabe, S., Ehira, S., Kanesaki, Y., Chibazakura, T. and Yoshikawa, H. (2015) 'Diversification of DnaA dependency for DNA replication in cyanobacterial evolution', *The Isme Journal*, 10, p. 1113.

Okumura, H., Yoshimura, M., Ueki, M., Oshima, T., Ogasawara, N. and Ishikawa, S. (2012) 'Regulation of chromosomal replication initiation by oriC-proximal DnaA-box clusters in Bacillus subtilis', *Nucleic acids research*, 40(1), pp. 220-234.

Ouellette, S.P., Gauliard, E., Antosova, Z. and Ladant, D. (2014) 'A Gateway((R)) -compatible bacterial adenylate cyclase-based two-hybrid system', *Environ Microbiol Rep*, 6(3), pp. 259-67.

Ozaki, S., Kawakami, H., Nakamura, K., Fujikawa, N., Kagawa, W., Park, S.Y., Yokoyama, S., Kurumizaka, H. and Katayama, T. (2008) 'A common mechanism for the ATP-DnaA-dependent formation of open complexes at the replication origin', *J Biol Chem*, 283(13), pp. 8351-62.

Ozaki, S., Noguchi, Y., Hayashi, Y., Miyazaki, E. and Katayama, T. (2012) 'Differentiation of the DnaA-oriC subcomplex for DNA unwinding in a replication initiation complex', *J Biol Chem*, 287(44), pp. 37458-71.

Parker, M.W., Botchan, M.R. and Berger, J.M. (2017) 'Mechanisms and regulation of DNA replication initiation in eukaryotes', *Crit Rev Biochem Mol Biol*, 52(2), pp. 107-144.

Pinto, U.M., Pappas, K.M. and Winans, S.C. (2012) 'The ABCs of plasmid replication and segregation', *Nature Reviews Microbiology*, 10(11), pp. 755-765.

Rahn-Lee, L., Gorbatyuk, B., Skovgaard, O. and Losick, R. (2009) 'The conserved sporulation protein YneE inhibits DNA replication in Bacillus subtilis', *J Bacteriol*, 191(11), pp. 3736-9.

Rahn-Lee, L., Merrikh, H., Grossman, A.D. and Losick, R. (2011) 'The sporulation protein SirA inhibits the binding of DnaA to the origin of replication by contacting a patch of clustered amino acids', *J Bacteriol*, 193(6), pp. 1302-7.

Raia, P., Delarue, M. and Sauguet, L. (2019) 'An updated structural classification of replicative DNA polymerases', *Biochem Soc Trans*, 47(1), pp. 239-249.

Rajewska, M., Wegrzyn, K. and Konieczny, I. (2012) 'AT-rich region and repeated sequences - the essential elements of replication origins of bacterial replicons', *FEMS Microbiol Rev*, 36(2), pp. 408-34.

Richardson, T.T., Harran, O. and Murray, H. (2016) 'The bacterial DnaA-trio replication origin element specifies single-stranded DNA initiator binding', *Nature*, 534(7607), pp. 412-6.

Richardson, T.T., Stevens, D., Pelliciari, S., Harran, O., Sperlea, T. and Murray, H. (2019) 'Identification of a basal system for unwinding a bacterial chromosome origin', *Embo j*, p. e101649.

Roecklein, B., Pelletier, A. and Kuempel, P. (1991) 'The tus gene of Escherichia coli: autoregulation, analysis of flanking sequences and identification of a complementary system in Salmonella typhimurium', *Res Microbiol*, 142(2-3), pp. 169-75.

Rohs, R., Jin, X., West, S.M., Joshi, R., Honig, B. and Mann, R.S. (2010) 'Origins of specificity in protein-DNA recognition', *Annu Rev Biochem*, 79, pp. 233-69.

Rohs, R., West, S.M., Sosinsky, A., Liu, P., Mann, R.S. and Honig, B. (2009) 'The role of DNA shape in protein-DNA recognition', *Nature*, 461(7268), pp. 1248-53.

Rudolph, C.J., Upton, A.L., Stockum, A., Nieduszynski, C.A. and Lloyd, R.G. (2013) 'Avoiding chromosome pathology when replication forks collide', *Nature*, 500(7464), pp. 608-611.

Russell, D.W. and Zinder, N.D. (1987) 'Hemimethylation prevents DNA replication in E. coli', *Cell*, 50(7), pp. 1071-9.

Sakiyama, Y., Kasho, K., Noguchi, Y., Kawakami, H. and Katayama, T. (2017) 'Regulatory dynamics in the ternary DnaA complex for initiation of chromosomal replication in Escherichia coli', *Nucleic Acids Res*, 45(21), pp. 12354-12373.

Sathyapriya, R., Vijayabaskar, M.S. and Vishveshwara, S. (2008) 'Insights into protein-DNA interactions through structure network analysis', *PLoS Comput Biol*, 4(9), p. e1000170.

Schneider, S., Zhang, W., Soultanas, P. and Paoli, M. (2008) 'Structure of the N-terminal oligomerization domain of DnaD reveals a unique tetramerization motif and provides insights into scaffold formation', *J Mol Biol*, 376(5), pp. 1237-50.

Scholefield, G., Errington, J. and Murray, H. (2012) 'Soj/ParA stalls DNA replication by inhibiting helix formation of the initiator protein DnaA', *Embo j*, 31(6), pp. 1542-55.

Scholefield, G. and Murray, H. (2013) 'YabA and DnaD inhibit helix assembly of the DNA replication initiation protein DnaA', *Mol Microbiol*, 90(1), pp. 147-59.

Sequeira-Mendes, J. and Gómez, M. (2012) 'On the opportunistic nature of transcription and replication initiation in the metazoan genome', *BioEssays*, 34(2), pp. 119-125.

Shimizu, M., Noguchi, Y., Sakiyama, Y., Kawakami, H., Katayama, T. and Takada, S. (2016) 'Near-atomic structural model for bacterial DNA replication initiation complex and its functional insights', *Proc Natl Acad Sci U S A*, 113(50), pp. E8021-e8030.

Simmons, L.A., Felczak, M. and Kaguni, J.M. (2003) 'DnaA Protein of Escherichia coli: oligomerization at the E. coli chromosomal origin is required for initiation and involves specific N-terminal amino acids', *Molecular Microbiology*, 49(3), pp. 849-858.

Simon, A.C., Zhou, J.C., Perera, R.L., van Deursen, F., Evrin, C., Ivanova, M.E., Kilkenny, M.L., Renault, L., Kjaer, S., Matak-Vinković, D., Labib, K., Costa, A. and Pellegrini, L. (2014) 'A Ctf4 trimer couples the CMG helicase to DNA polymerase α in the eukaryotic replisome', *Nature*, 510, p. 293.

Smits, W.K., Goranov, A.I. and Grossman, A.D. (2010) 'Ordered association of helicase loader proteins with the Bacillus subtilis origin of replication in vivo', *Mol Microbiol*, 75(2), pp. 452-61.

Smulczyk-Krawczyszyn, A., Jakimowicz, D., Ruban-Ośmiałowska, B., Zawilak-Pawlik, A., Majka, J., Chater, K. and Zakrzewska-Czerwińska, J. (2006) 'Cluster of DnaA Boxes Involved in Regulation of Streptomyces Chromosome Replication: from In Silico to In Vivo Studies', *Journal of Bacteriology*, 188(17), pp. 6184-6194.

Sueoka, N. (1998) 'Cell membrane and chromosome replication in Bacillus subtilis', *Prog Nucleic Acid Res Mol Biol*, 59, pp. 35-53.

Tan, I.S. and Ramamurthi, K.S. (2014) 'Spore formation in Bacillus subtilis', *Environmental microbiology reports*, 6(3), pp. 212-225.

Taylor, J.A., Ouimet, M.C., Wargachuk, R. and Marczynski, G.T. (2011) 'The Caulobacter crescentus chromosome replication origin evolved two classes of weak DnaA binding sites', *Mol Microbiol*, 82(2), pp. 312-26.

Tsai, K.L., Lo, Y.H., Sun, Y.J. and Hsiao, C.D. (2009) 'Molecular interplay between the replicative helicase DnaC and its loader protein DnaI from Geobacillus kaustophilus', *J Mol Biol*, 393(5), pp. 1056-69.

Tsodikov, O.V. and Biswas, T. (2011) 'Structural and thermodynamic signatures of DNA recognition by Mycobacterium tuberculosis DnaA', *J Mol Biol*, 410(3), pp. 461-76.

Usongo, V. and Drolet, M. (2014) 'Roles of Type 1A Topoisomerases in Genome Maintenance in Escherichia coli', *PLOS Genetics*, 10(8), p. e1004543.

Veening, J.W., Murray, H. and Errington, J. (2009) 'A mechanism for cell cycle regulation of sporulation initiation in Bacillus subtilis', *Genes Dev*, 23(16), pp. 1959-70.

Villa, F., Simon, A.C., Ortiz Bazan, M.A., Kilkenny, M.L., Wirthensohn, D., Wightman, M., Matak-Vinkovic, D., Pellegrini, L. and Labib, K. (2016) 'Ctf4 Is a Hub in the Eukaryotic Replisome that Links Multiple CIP-Box Proteins to the CMG Helicase', *Mol Cell*, 63(3), pp. 385-96.

Waldminghaus, T. and Skarstad, K. (2009) 'The Escherichia coli SeqA protein', *Plasmid*, 61(3), pp. 141-50.

Wang, S., Poon, G.M.K. and Wilson, W.D. (2015) 'Quantitative Investigation of Protein-Nucleic Acid Interactions by Biosensor Surface Plasmon Resonance', *Methods in molecular biology (Clifton, N.J.)*, 1334, pp. 313-332.

Warren, E.M., Huang, H., Fanning, E., Chazin, W.J. and Eichman, B.F. (2009) 'Physical interactions between Mcm10, DNA, and DNA polymerase alpha', *J Biol Chem*, 284(36), pp. 24662-72.

Weigel, C., Schmidt, A., Seitz, H., Tungler, D., Welzeck, M. and Messer, W. (1999) 'The N-terminus promotes oligomerization of the Escherichia coli initiator protein DnaA', *Mol Microbiol*, 34(1), pp. 53-66.

Wiegert, T. and Schumann, W. (2001) 'SsrA-mediated tagging in Bacillus subtilis', *Journal of bacteriology*, 183(13), pp. 3885-3889.

Wigley, D.B. (2009) 'ORC proteins: marking the start', *Curr Opin Struct Biol*, 19(1), pp. 72-8.

Wolański, M., Donczew, R., Zawilak-Pawlik, A. and Zakrzewska-Czerwińska, J. (2015) 'oriC-encoded instructions for the initiation of bacterial chromosome replication', *Frontiers in Microbiology*, 5(735).

Wolanski, M., Jakimowicz, D. and Zakrzewska-Czerwinska, J. (2012) 'AdpA, key regulator for morphological differentiation regulates bacterial chromosome replication', *Open Biol*, 2(7), p. 120097.

Wu, Z., Liu, J., Yang, H. and Xiang, H. (2014) 'DNA replication origins in archaea', *Frontiers in microbiology*, 5, pp. 179-179.

Yamtich, J. and Sweasy, J.B. (2010) 'DNA polymerase Family X: Function, structure, and cellular roles', *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1804(5), pp. 1136-1150.

Yang, W. (2014) 'An overview of Y-Family DNA polymerases and a case study of human DNA polymerase η', *Biochemistry*, 53(17), pp. 2793-2803.

Yao, N.Y. and O'Donnell, M. (2010) 'SnapShot: The replisome', *Cell*, 141(6), pp. 1088, 1088.e1.

Yao, N.Y. and O'Donnell, M.E. (2016) 'Evolution of replication machines', *Crit Rev Biochem Mol Biol*, 51(3), pp. 135-49.

Yuan, Z., Georgescu, R., Bai, L., Zhang, D., Li, H. and O'Donnell, M. (2019) 'DNA unwinding mechanism of a eukaryotic replicative CMG helicase', *bioRxiv*, p. 737015.

Zakrzewska-Czerwinska, J., Jakimowicz, D., Zawilak-Pawlik, A. and Messer, W. (2007) 'Regulation of the initiation of chromosomal replication in bacteria', *FEMS Microbiol Rev*, 31(4), pp. 378-87.

Zawilak-Pawlik, A., Donczew, R., Szafranski, S., Mackiewicz, P., Terradot, L. and Zakrzewska-Czerwinska, J. (2011) 'DiaA/HobA and DnaA: a pair of proteins co-evolved to cooperate during bacterial orisome assembly', *J Mol Biol*, 408(2), pp. 238-51.

Zawilak-Pawlik, A., Kois, A., Stingl, K., Boneca, I.G., Skrobuk, P., Piotr, J., Lurz, R., Zakrzewska-Czerwinska, J. and Labigne, A. (2007) 'HobA--a novel protein involved in initiation of chromosomal replication in Helicobacter pylori', *Mol Microbiol*, 65(4), pp. 979-94.