ResearchOnline@JCU

This file is part of the following reference:

Moreau, Morgane Julie Jackie (2013*) DNA replication in Escherichia coli: a comprehensive study of the Tus-Ter complex***. PhD thesis, James Cook University.**

Access to this file is available from:

http://researchonline.jcu.edu.au/31903/

JAMES COOK
UNIVERSITY
AUSTRALIA

DNA replication in *Escherichia coli*:
A comprehensive study of the Tus-*Ter* complex

Morgane Julie Jackie Moreau BSc (Hons)

JAMES COOK
UNIVERSITY
AUSTRALIA

for the degree of Doctor of Philosophy
School of Pharmacy and Molecular Sciences
James Cook University
March 2013

Statement on the contribution of others

Nature of Assistance	Contribution	Names, Titles and Affiliation of Co-contributors
Financial support	Stipend	JCU IPRS
	Research grant (Project title: Genome wide, temporal and spatial distribution of replisomal proteins in *E. coli)*	Graduate Research Scheme, Faculty of Medicine, Health and Molecular Sciences of James Cook University, Townsville (Australia)
Research assistance	Construction of strains carrying ectopic *Ter* sites and growth rates analysis (Chapter 7)	Peter J Enyart (BS), Savitri Mandapati, Jiri Perutka, Institute for Cellular & Molecular Biology. University of Texas at Austin (USA),

This thesis is composed of my original work, and contains material that has been previously published. The published manuscripts have been adapted for consistency and the work contributed by my co-authors in these publications has not been included in my thesis as they were not directly related to the study of the Tus-*Ter* complex. The study on the effect of ectopic *Ter* sites was carried out in collaboration with Prof. Ellington's group as mentioned in the above table. For this experiment, I selected the *Ter* DNA sequences to be inserted in the *E. coli* genome and assisted in identifying the optimal region for their insertion. The construction of the strains and the growth curve analysis have been carried out by Jiri Perutka (*in silico* design), Savitri Mandapati (strain development) and Peter Enyeart (project management) at the Institute for Cellular & Molecular Biology, University of Texas in Austin, USA.

Publications included in this thesis:

- **Moreau, M. J.**, Morin, I. & Schaeffer, P. M. 2010. Quantitative determination of protein stability and ligand binding using a green fluorescent protein reporter system. Mol biosyst, 6, 1285-92.

- **Moreau, M. J**. & Schaeffer, P. M. 2012. A polyplex qPCR-based binding assay for protein-DNA interactions. Analyst, 137, 4111-3.

- **Moreau, M. J.** & Schaeffer, P. M. 2012. Differential Tus-*Ter* binding and lock formation: implications for DNA replication termination in *Escherichia coli*. Mol biosyst, 8, 2783-91.

- **Moreau, M. J.**, Morin, I., Askin, S. P., Cooper, A., Moreland, N. J., Vasudevan, S. G. & Schaeffer, P. M. 2012. Rapid determination of protein stability and ligand binding by differential scanning fluorimetry of GFP-tagged proteins. Rsc advances, 2, 11892-11900.

Declaration on ethics

The research presented and reported in this thesis was conducted in compliance with the National Health and Medical Research Council (NHMRC) Australian Code of Practice for the Care and Use of Animals for Scientific Purposes, 7th Edition, 2004 and the Qld Animal Care and Protection Act, 2001. The proposed research study received animal ethics approval from the JCU Animal Ethics Committee Approval Number # A1642.

Copyright declaration

I acknowledge that copyright of all material contained in my thesis resides with the copyright holder(s) of that material. Every reasonable effort has been made to gain permission and acknowledge the owners of copyright material. I would be pleased to hear from any copyright owner who has been omitted or incorrectly acknowledged.

Acknowledgments

I would like to thank many people who have helped me along the way up to the completion of this dissertation. First, I am indebted to my supervisor, Ass. Prof. Patrick Schaeffer of James Cook University, whose training, guidance and patience has been an invaluable support. I have benefited from his direct academic experience and I have grown immensely thank to this chance. He continuously and convincingly conveyed a spirit of adventure in research, scholarship and teaching. His mentorship did and will always inspire me. I am also very grateful to my co-supervisor Prof. James Burnell for his support and giving me access to his facilities. If one day I teach like Prof. Burnell, I will have reached one goal. I would like to thank Prof. Alan Baxter, a captivating, intelligent and insightful adviser; he is also a true inspiration. Additionally, I would like to thank the Graduate Research School for funding my time at James Cook University.

I am honored to have worked with Dr Isabelle Morin and Samuel Askin, dynamic and intelligent, they provided a stimulating and fun environment in which to learn and grow. They have always been a great source of advice and feedbacks and always with a smile. Sam, you are so genuine, resourceful and fun, I loved working with you. Isa, same as for Sam but I have to add a big thank you for supporting me when experiments failed as much as helping me celebrating the successes. I also would like to mention Dr Alanna Cooper, for a great training in chicken catching.

Next and above all, I would like to thank Yannick Zimmermann for always believing in me and encouraging me for the last 15 years. Il y avait toujours ta petite voie quand j'avais besoin de force, merci d'avoir toujours cru en moi, c'est aussi grâce à toi que je suis la.

Abstract

The circular chromosome of bacteria is replicated by two replisomes assembled at the unique origin and moving in opposite direction until they meet at specific termination sites. The process of DNA replication termination is the stage of replication that is the least understood, both in prokaryotes and eukaryotes. In *E. coli*, the termination protein Tus binds to 14 termination sites (*TerA-J*, *TerK*, *L*, *TerY*, *Z*) spread throughout the genome. The intriguing organization and symmetry of *Ter* sites has puzzled scientists for decades. The Tus-*Ter* complex is polar and blocks replication forks approaching from one direction but not from the other. Most *Ter* sites are oriented to form a fork trap so that convergent forks can enter and merge in the terminus region but not exit. However, the significance of having maintained such a wide fork trap remains unclear. The mechanism responsible for the polarity of the Tus-*Ter* complex is still being debated. A protein-protein interaction between the DnaB helicase at the forefront of the replisome and Tus bound to *Ter* has been proposed (Bastia et al., 2008, Mulugu et al., 2001). The alternative mechanism involves the formation of the Tus-*Ter*-lock (TT-lock) where Tus captures the cytosine at position 6 in the *Ter* core sequence upon duplex unwinding by DnaB and becomes locked on *Ter* thereby preventing DnaB translocation (Mulcair et al., 2006). Since the discovery of the TT-lock, there has been no further investigation on its formation in the remaining Tus-*Ter* complexes. However, the proportion of fork pausing at each *Ter* sites has previously been determined *in vivo* and was detected at seven *Ter* sites (*TerA-D*, *TerG*, *TerH* and *TerI*). The remaining *Ter* sites were classified as *pseudo-Ter* (Duggin and Bell, 2009). Nevertheless, all *Ter* were able to arrest forks in an artificial context, yet with varying efficiencies (Duggin and Bell, 2009). This prompted the question of whether or not the outer *Ter* sites maintained their biological function. This work provides the first comparative study of the ten primary *Ter* sites (*TerA-J*) in terms of their affinity and specificity

for Tus and whether they are all able to form a TT-lock. The variation in affinity and TT-lock forming ability of *Ter* sites was compared to their intrinsic efficiency in arresting a replisome and to the *in vivo* distribution of Tus on *Ter* sites. Finally, ectopic *Ter* sites were inserted into the *E. coli* genome to determine the effect of TT-lock formation on cell growth. Several new methods were developed during this thesis for the characterization of Tus-*Ter* and Tus-*Ter*-lock complexes in a time and cost-effective manner.

The *Ter* sites were shown to be different both in terms of their affinity for Tus and in their ability to form a TT-lock. Six strong Tus binding sites (*TerA-E* and *TerG*) were identified and the outermost *TerH*, *TerI* and *TerJ* were classified as moderate binders. The binding of Tus to *TerF* was only marginally stronger than a non-specific DNA region of the *oriC*. The strong binders were all able to form a strong TT-lock whereas moderate binders varied in their TT-lock forming efficiencies. *TerF* and *TerH* were unable to form significant locks. The affinity and TT-lock forming efficiencies of the *Ter* sites correlated well with their intrinsic pausing efficiency determined by Duggin and Bell (2009). In the cell, Tus was distributed onto *Ter* sites according to their intrinsic affinity. It was demonstrated that only the strong *Ter* sites are able to cause significant fork arrest suggesting that replication forks are unlikely to break through the innermost *Ter* sites and that the outer *Ter* sites may be used to prevent non-*oriC* initiated forks to travel towards the origin. A new paradigm is being proposed to explain the multiplicity of *Ter* sites and the advantage in maintaining such a wide fork trap. Finally, the three new assays developed in this study, GFP-Basta, DSF-GTP and the qPCR-based DNA binding assay, proved to be invaluable tools for the detailed characterization of protein-DNA complexes. These news techniques have considerable applications in both genomic and proteomic programs.

Table of Content

Table of Figures

List of Tables

Chapter 1: General introduction

Bacteria are ideal model organisms for studying fundamental life processes; much of the present knowledge about biological systems stems from studies of the bacteria *Escherichia coli* (Lee and Lee, 2003). They are also implicated in many processes related to human health and diseases (Reyes-Lamothe et al., 2008b). Bacterial and viral DNA replication systems are obvious targets for the search of novel anti-bacterial and anti-viral agents (Berdis, 2008, Lange et al., 2007, Tanner et al., 2009). Our understanding of DNA replication comes from investigation of the prokaryotic replication machinery of the T4 and T7 bacteriophages, and from the gram-positive and gram-negative bacteria *Bacillus subtilis* and *Escherichia coli*. Knowledge obtained from these model organisms gave hints to understand this process in higher organisms (i.e. eukaryotes) that have a more complex genome in terms of chromosome structure and number of proteins involved. It is important to fully understand DNA replication not only to increase our knowledge on this fundamental process but also to learn the mechanisms by which proteins and DNA interact. DNA replication is mediated and coordinated by a large number of proteins that form highly dynamic protein-DNA and/or protein-protein interactions to achieve the duplication of the chromosome prior to cell division.

Replisomal proteins have been extensively studied in different organisms, most of them are characterized in terms of structure and function but for others the role is uncertain. Indeed,

the mechanism of replication termination in *E.coli* is still a matter of debate and the presence of a large "fork trap" remains to be understood.

This chapter describes our current understanding of the molecular basis of DNA replication in *E. coli* with a particular emphasis on its replication termination system. DNA replication termination has been the focus of the work presented in this thesis which aimed at further characterizing the function and role of the replication fork trap.

1.1 DNA replication and cell cycle in *E. coli*

The average generation time of *E. coli* varies between ~20 and ~40 minutes depending on growth conditions. The time necessary to complete the replication of the chromosome is estimated at 40 minutes and is longer than the time required to divide under optimal growth conditions (Ferullo et al., 2009). Therefore a second round of initiation of replication must start before completion of the first round, resulting in new cells born with partially replicated chromosome (up to 16-ploid for the origin region of the chromosome; Ferullo et al., 2009). Under poor growth conditions, DNA replication is completed before the cells divide and therefore contain a maximum of two chromosomes.

Ferullo et al. (2009) synchronized cultures of *E. coli* and obtained mainly 4 N cells (73 % of cells with 4 chromosomes) in nutrient-rich minimal media and replication of *oriC* was detected as early as 6-10 minutes after release of the DNA replication inhibitor serine hydroxamate which prevent replication initiation through protein synthesis inhibition (Tosa and Pizer, 1971). DNA content shifted from 4 N to 8 N between 30 and 44 minutes after release, a second round of replication seemed to start at 44 min and complete cell division occurred at 60 minutes (Ferullo et al., 2009). Interestingly, Bates and Kleckner (2005) studied the effect of three growth rates in *E. coli* (T_d= 90, 125 and 300 minutes) on the timing of replication events (G1, S, G2 phases) and found that G1 and S phases were unaffected. The length of the G2

phase, between the end of bulk replication and cell division, accommodated differences in growth rates, suggesting that growth conditions affect septation and cell division but not DNA replication dynamics.

1.2 DNA synthesis in *E. coli*

1.2.1 The *E. coli* chromosome and fundamental aspects of replication

The *E. coli* chromosome is circular and is 4.6 Mbp long. As in eukaryotes, bacterial replication is semi-conservative with each daughter cell receiving a new DNA double helix made of one parental and one newly synthesized strand. The process of DNA replication involves the coordinated activity of many proteins that collectively compose a molecular machinery called the replisome.

In *E. coli*, two replisomes are assembled at the single origin of replication, *oriC*, and initiate DNA unwinding, primer synthesis and strand extension (detailed below). The two replisomes proceed bidirectionally until they meet in the termination region opposite to the origin (Figure 1A). The whole chromosome is replicated as a single replicon of 4.6 Mbp by two replisomes. It was suggested earlier that the two replisomes were both anchored in a specific location in the cell and formed a fixed replication factory where DNA was pulled through to be replicated (Adachi et al., 2005, den Blaauwen et al., 2006, Lemon and Grossman, 1998, Lemon and Grossman, 2000, Molina and Skarstad, 2004). However more recent evidence established that forks move independently as each replisome follow the path of compacted DNA (Bates and Kleckner, 2005, Reyes-Lamothe et al., 2008a, Wang et al., 2011).
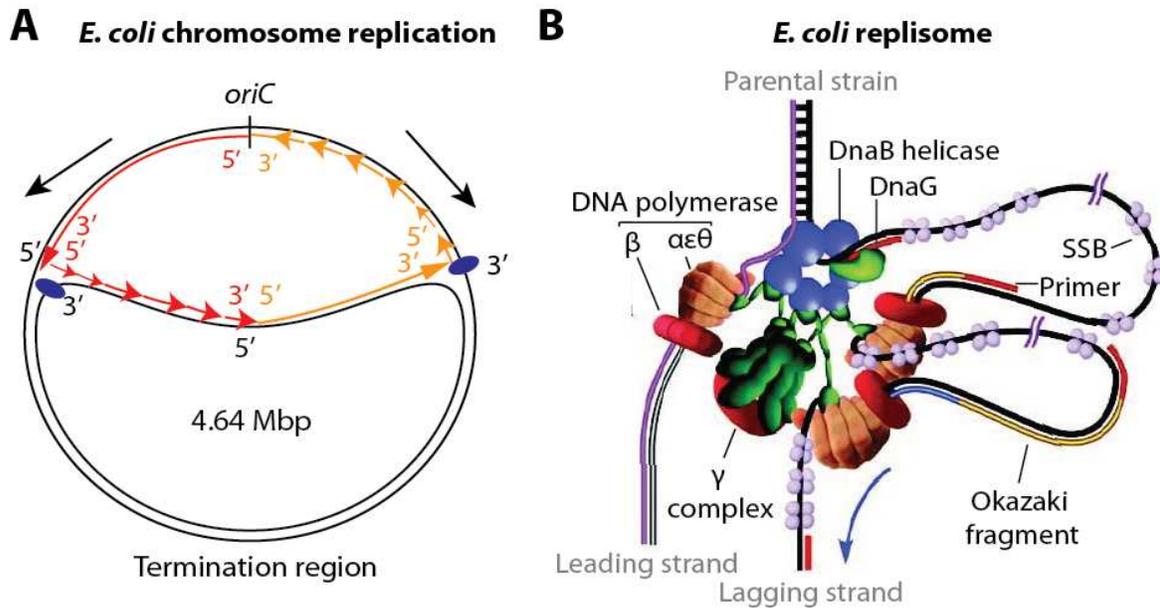
Figure 1: Semi-discontinuous replication and replisomal proteins in *E. coli*. (A) Replisomes (blue circles) are assembled at *oriC* and move in opposite directions towards the termination region. Parental DNA is in black and newly synthesized DNA is in orange (clockwise moving fork) and in red (anti-clockwise moving fork). The lagging strand is initially made of a series of Okazaki fragments (short arrows) with those nearest to the fork being the most recently synthesized. (B) Arrangement of replisomal proteins in the TriPol replisome with one polymerase on the leading strand and two on the lagging strand, one synthetizing the Ozakazi fragments and the other filling in ssDNA gaps between two Okazaki fragments. The loop enables the co-directionality of the leading and lagging strand polymerases (Adapted from Georgescu et al., 2012).

Within a replisome, the replicative DNA polymerase III holoenzyme (pol III HE) of *E. coli* synthesizes DNA unidirectionally in the 5'-to-3' direction from a previously synthesized RNA primer. Due to this polarity (Figure 1A), the leading strand is synthesized continuously and the lagging strand is made of a series of Okazaki fragments of about 1 to 3 kb in length (Rowen and Kornberg, 1978). Pol III HE catalyses DNA elongation at about 1 kb/s with high fidelity (Bloom, 2006, Furukohri et al., 2008).

1.2.2 Replisomal proteins: structure and function

Each replisome consists of the assembly of numerous replisomal proteins that all together coordinate the faithful copying of the parental chromosome. Each of these proteins has one or more specific role(s) and is a candidate reporter of the replisome dynamics *in vivo* (i.e. leading or lagging strand synthesis dynamics, initiation or termination of replication). This section introduces the main replisomal proteins, their functions and their interactions. The sequence of events leading to the replication of the chromosome is detailed in the following sections (*cf* sections 1.2.3, 1.2.4, 1.2.5).

The *E. coli* replisome is relatively complex in terms of the number of proteins involved compared to other model systems such as the T4 and T7 replisome. The replisome in *E. coli* consists of the primosome (DNA helicase and DNA primase activities) and the DNA polymerase III and accessory factors. The primosome is first recruited at the *oriC* to initiate DNA replication (Mott and Berger, 2007, Zakrzewska-Czerwinska et al., 2007). It comprises the initiator protein DnaA bound to *oriC*, the helicase-loader DnaC, the DnaB helicase unwinding DNA and the DnaG primase priming ssDNA (Ozaki and Katayama, 2009).

The initiator protein DnaA (~ 53 kDa) has four functional domains and its structure has recently been solved (Ozaki and Katayama, 2009). The critical residues involved in DNA binding (domain IV), inter-DnaA interactions (AAA+ domain III and domain I), regulatory interactions with ATP and ssDNA (AAA+ domain III) and DnaB helicase interaction (N-terminal domain and domain III) are now well characterized. DnaA binds sequentially at *oriC* to several copies of a 9 bp DNA consensus sequence (TTATNCACA) which include DnaA boxes R1 to R5 (Davey et al., 2002, Erzberger et al., 2002, Margulies and Kaguni, 1996, Messer, 2002). DnaA is a member of the AAA+ ATPase family (Miller et al., 2009, Neuwald et al., 1999) and the AAA+ domain is responsible for the cooperative self-assembly of subunits in high-order ring-shaped hexameric complexes on *oriC* (Erzberger et al., 2006, Kaguni, 2006,

Duderstadt et al., 2011, Ozaki and Katayama, 2012). DnaA can take either an ATP-bound (ATP-DnaA) or an ADP-bound form (ADP-DnaA) but only the ATP-DnaA multimer can initiate replication (Bramhill and Kornberg, 1988, Ozaki and Katayama, 2012, Sekimizu et al., 1987). An ATP-dependent interaction between ATP-DnaA complexes leads to DNA duplex unwinding and opening within *oriC* (Erzberger et al., 2006, Kawakami et al., 2005, Ozaki and Katayama, 2009, Ozaki et al., 2008) for the sudsequent recruitment of DnaB helicase via domain I and III (Ozaki and Katayama, 2009).

The DnaC helicase loader (~27 kDa) is also part of the AAA+ ATPase family and regulates the activity and loading of DnaB helicase on the chromosome with the help of DnaA (Seufert and Messer, 1987). ATP molecules bind at the interface between neighboring DnaC subunits and act as a molecular switch to load the DnaB helicase (Davey et al., 2002, Erzberger et al., 2006).

The DnaB replicative helicase (~ 52 kDa) is a ring-shaped homohexameric complex (DnaB$_6$) that encircles the lagging strand and uses ATP to fuel its translocation in the 5'-3' direction to separate the parental duplex DNA at the front of the replication fork (Delagoutte and von Hippel, 2003, Patel and Picha, 2000, Schaeffer et al., 2005). DnaB$_6$ forms a preprimosome complex with DnaC in the form of a (DnaBC)$_6$ heterohexamer before it is loaded onto the ssDNA (Lanka and Schuster, 1983). ATP-bound DnaC helps loading DnaB$_6$ onto ssDNA who also interacts with the N-terminal domain of DnaA (Seitz et al., 2000). DnaC dissociates from DnaB after ATP hydrolysis activating the DnaB helicase (Kobori and Kornberg, 1982, Marszalek and Kaguni, 1994, Marszalek et al., 1996, Seitz et al., 2000, Wahle et al., 1989, Wickner and Hurwitz, 1975). DnaB translocates 65 nucleotides and recruits DnaG primase (Fang et al., 1999).

The DnaG primase (~ 64 kDa) is a DNA-dependent RNA polymerase that lays a short RNA primer (pRNA) on the newly separated strands for the pol III HE to recognize a nascent

DNA strand. DnaG primase is composed of three domains: the N-terminal DNA binding domain, the oligonucleotide synthesis domain and the C-terminal domain which interacts with the N–terminal of DnaB (Mitkova et al., 2003, Oakley et al., 2005, Rodina and Godson, 2006). DnaG requires the presence of DnaB at the origin or the single stranded binding proteins (SSB) on ssDNA to synthesize an RNA primer (8-12 nucleotides) on the leading strand, and one every ~ 1 kb at a 5'-CTG preferential recognition site to initiate the synthesis of Okazaki fragments on the lagging strand (Frick and Richardson, 2001, Kitani et al., 1985, Yoda et al., 1988).

The single-stranded DNA binding protein (SSB) (~ 20 kDa) from *E. coli* binds tightly to helicase-produced ssDNA regions independently of the sequence (Meyer and Laine, 1990). SSB stabilizes ssDNA and prevents or removes secondary structures such as hairpins or cruciforms (Kuznetsov et al., 2006). SSB forms a tetramer with each subunit binding to ssDNA. The subunits interact with each other's N-terminal binding domain (Raghunathan et al., 2000, Raghunathan et al., 1997) and interact with other DNA-handling enzymes, i.e. nuclease (Genschel et al., 2000), PriA helicase (i.e. for replisome restart; Cadman and McGlynn, 2004) and polymerases (Witte et al., 2003) through the more flexible C-terminal domain (Kunzelmann et al., 2010, Roy et al., 2007). The tetramer has multiple binding modes (i.e. conformations) to ssDNA (two major modes: SSB_{35} and SSB_{65}) that coordinate the function of many DNA processing enzymes, either by protein-protein interactions or by controlling the accessibility to ssDNA at replication forks (Lohman and Ferrari, 1994, Meyer and Laine, 1990, Roy et al., 2007, Sun and Godson, 1998). These binding modes are salt-dependent; only one of the two modes is significantly populated (more stable) outside a narrow range of salt concentration (Roy et al., 2007). However, the properties of each binding mode are not yet defined (Kunzelmann et al., 2010).

    E. coli polymerase III holoenzyme (pol III HE) is a large assembly of several subunits that make the HE a very efficient enzyme that incorporates complementary oligonucleotides with almost perfect fidelity (Schaeffer et al., 2005). These properties require the cooperation of ten different subunits arranged in three sub-assemblies: the core $(\alpha\epsilon\theta)_2$, the β_2 sliding clamp and the γ complex or clamp loader. The last two subassemblies enable the Pol III core to have the highest processivity of any *E. coli* DNA polymerase (Johnson and O'Donnell, 2005 and reference therein). The isolable core polymerase $(\alpha\epsilon\theta)$ is responsible for the DNA polymerase and proofreading exonucleases activities (El Houry Mignan et al., 2011, Scheuermann et al., 1983) and is composed of three polypeptides: the catalytic subunit α (~ 130 kDa), the 3'-5' exonuclease ϵ (~ 27 kDa) and the accessory unit θ (~ 9 kDa) that binds to ϵ to stimulate its editing function (Benkovic et al., 2001). Until recently, the HE was thought to comprise two cores but it is now known to comprise three cores, one on the leading strand and two on the lagging strand (Figure 1B; Georgescu et al., 2012, Lia et al., 2012, Reyes-Lamothe et al., 2010). The second lagging strand PolIII core (Pol_{lag}) is thought to finish Okazaki fragment (OF) synthesis when the initial Pol_{lag} is released before reaching the 5'-end of the following OF as per the signal release model (Lia et al., 2012). The processivity of the Pol_{lag} is considerably improved in tripolymerase replisomes due to an increased gap filling efficiency between OF compared to dipolymerase replisomes (Georgescu et al., 2012). The second subassembly of Pol III HE is the β_2 sliding clamp (~ 82 kDa), a dimer that encircles dsDNA at the primer terminus and tethers Pol III core on DNA via interactions with α. The β_2 sliding clamp enables the Pol III to synthesize DNA continuously without dissociation (Kim and McHenry, 1996). The third subassembly is the γ clamp loader complex, an AAA+ ATP-ase which loads the β_2 sliding clamp onto the 3'-end of a primed DNA. The clamp loader is composed of six different subunits: δ', γ/τ, δ, ψ and χ (~ 37, 47, 71, 37, 15 and 16 kDa respectively; Bloom, 2006, Jeruzalmi et al., 2001, McHenry, 2003, O'Donnell, 2006, Schaeffer et al., 2005) arranged in the

following stoichiometry: [τ/γ]$_3$.δ. δ'.ψ.χ (Lia et al., 2012, Reyes-Lamothe et al., 2010). The τ and γ subunits can interchangeably be components of the clamp loader but only τ can recruit the PolIII cores (Johnson and O'Donnell, 2005, Reyes-Lamothe et al., 2010). The γ subunit is a non-essential truncated form of τ produced by a programmed frameshift during *dnaX* gene translation (Blinkowa and Walker, 1990, Flower and McHenry, 1990, Tsuchihashi and Kornberg, 1990). This subunit is not associated with the clamp loader in replisomes and was proposed to replace τ in post-replication repair-associated events (Reyes-Lamothe et al., 2010). τ also interacts with DnaB helicase (Gao and McHenry, 2001, Jergic et al., 2007, Kim and McHenry, 1996, Studwell-Vaughan and O'Donnell, 1991). δ interacts directly with the β-clamp and is sufficient to open and recycle it on the lagging strand. The δ' subunit helps modulating this δ-β interaction (O'Donnell et al., 2001, Turner et al., 1999). The ψ·χ heterodimer is not essential to clamp loading (El Houry Mignan et al., 2011, Xiao et al., 1993) but greatly enhances clamp loader stability. χ is the only direct link between PolIII and SSB (Kelman et al., 1998, Witte et al., 2003) and is necessary for the primase-to-polymerase switch as it competes with DnaG for SSB binding (Yuzhakov et al., 1999). ψ bridges χ to the (τ/γ)$_3$δδ' complex, stabilizing the clamp loader (Olson et al., 1995) and increasing its affinity for the β-clamp (Anderson et al., 2007).

The replication termination protein Tus (~ 36 kDa) is a monomeric protein that recognizes multiple DNA sequences called termination (*Ter*) sites. These sites are about 21-bp in length and are mainly scattered in the bottom half of the chromosome opposite to *oriC* (Kamada et al., 1996, Neylon et al., 2000). These sites are polar and can block replication forks moving in one direction but not in the other. They are arranged to create a fork trap constraining converging forks in the terminus region of the chromosome (Duggin and Bell, 2009, Mulcair et al., 2006, Neylon et al., 2005, Schaeffer et al., 2005). The structure of the

termination region and the mechanism of the fork trap polarity are reviewed in details in section 1.2.6.

1.2.3 Initiation of replication

The initiation of replication requires the precisely timed formation of a nucleoprotein complex composed of several DnaA initiator proteins at *oriC* (Kaguni, 2006, Miller et al., 2009, Mott and Berger, 2007, Mott et al., 2008). Only ATP-DnaA molecules can multimerise on the DnaA-assembly region (DAR) and form an active complex able to unwind the duplex unwinding elements (DUE, AT-rich repeats, Figure 2; Messer, 2002). DAR includes DnaA boxes R1-R5, I1-3 and τ1-2 (Ozaki and Katayama, 2012). The firing of chromosomal replication initiation is triggered by an increase in cellular level of ATP-DnaA (Fujimitsu et al., 2009, McGarry et al., 2004, Sekimizu et al., 1987, Speck et al., 1999). The recruitment of ATP-DnaA multimers on *oriC* is coordinated by the DnaA-initiator association protein DiaA (DnaA-binding protein; Keyamura et al., 2007, Ozaki and Katayama, 2009). ATP-DnaA binds to three high-affinity DnaA boxes R1, R2 and R4 (K_D < 200 nM) within *oriC* and accumulates additional ATP-DnaAs that interact with the lower affinity sites to form the pre-replication complex (Davey et al., 2002, Erzberger et al., 2006, Erzberger et al., 2002, Margulies and Kaguni, 1996, Messer, 2002, Miller et al., 2009, Nievera et al., 2006). Several histone-like proteins (HU, Fis, and/or IHF) also bind to the DAR region and regulate the formation of the pre-replication complex (reviewed in detail in Kaguni, 2011, Ozaki and Katayama, 2009, Ozaki and Katayama, 2012). An ATP-dependent interaction between specific DnaA AAA+ domains causes the 9 bp DUE to be wrapped around the DnaA molecules and the melting of their AT-rich region (Figure 2; Bramhill and Kornberg, 1988, Davey et al., 2002, Duderstadt et al., 2011, Erzberger et al., 2006, Fujimitsu et al., 2009, Messer, 2002, Ozaki and Katayama, 2009, Ozaki et al., 2008).
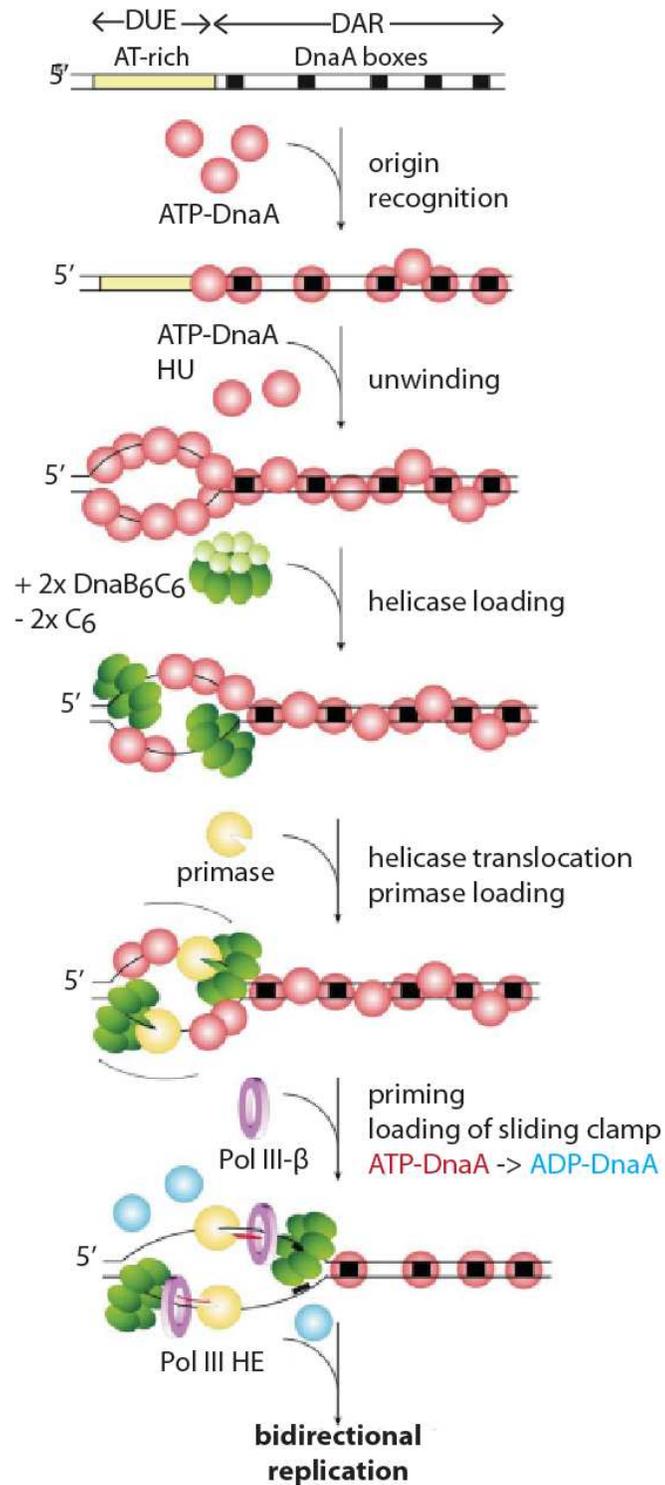
Figure 2: Steps of replication initiation in *E. coli*. ATP-DnaA molecules multimerise at DAR for additional ATP-DnaAs to bind and unwind DUEs. The (DnaBC)₆ complexe enters the openned DNA duplex, DnaC molecules are released activating DnaB. DnaG primase is then recruited and the pol III HE is loaded onto a primed template Adapted form Messer (2002).

DnaA binding to high affinity sites in *oriC* persists throughout most of the cell cycle (Cassler et al., 1995, Nievera et al., 2006, Samitt et al., 1989). The active initiation complex contains 20-30 DnaA monomers, as determined by electron microscopy *in vitro* (Crooke et al., 1993, Funnell et al., 1987). Most recently, the crystal structure of the AAA+ and duplex-DNA domains of *Aquifex aeolicus* DnaA bound to ssDNA revealed a DnaA:ssDNA stochiometry of 4:1 (Duderstadt et al., 2011). Once DnaA is recruited and activated at *oriC*, interactions between DnaC molecules within the pre-primosome complex $(DnaBC)_6$ and DnaA promote the loading of $DnaB_6$ on each strand designated to become the lagging strand (Figure 2; Konieczny, 2003, Marszalek and Kaguni, 1994, Seitz et al., 2000, reviewed in Kaguni, 2011). DnaC leaves the complex after or during DnaB loading, accompanied by ATP-hydrolysis which activates the helicase activity of DnaB (Wahle et al., 1989). The bubble is extended to about 65 nucleotides (Fang et al., 1999) by translocation of the $DnaB_6$ helicase in the 5'-3' direction on the lagging strands of the two replication forks (Schaeffer et al., 2005). Then DnaG primase can enter into the replication complex through the interaction of its C-terminal domain with DnaB. Once recruited, DnaG initiates DNA replication through the synthesis of a short RNA primer. The sliding clamp β_2 of Pol III is then loaded onto each primed template by the γ complex (Kelman and O'Donnell, 1995). The replicative polymerase Pol III is then loaded at the primer termini through interaction with the γ complex and the β_2 sliding clamp (Kaguni, 2011, Naktinis et al., 1996, Pomerantz and O'Donnell, 2007, Stukenberg et al., 1991). β_2 also mediates the hydrolysis of ATP-DnaA to ADP-DnaA, which promotes the sequestration of the newly replicated *oriC* and prevents re-initiation of replication (Katayama and Sekimizu, 1999). Once the replisome is complete Pol III starts to replicate the template strands.

1.2.4 Elongation phase of DNA replication

Due the the fact that a DNA polymerase processes in the 5'-3' direction, the coordination of both polymerases (one on the leading strand and two on the lagging strand) requires the presence of a lagging strand loop or trombone loop that allows the two polymerases to colocalize and point in the same direction (Alberts et al., 1983, Breier et al., 2005, Yao and O'Donnell, 2008). The T7 replication model illustrates this principle in Figure 3. Despite the additional steps required for lagging strand synthesis (i.e. release of Okazaki fragment, primer synthesis and hand-off to PolIII core), the leading and lagging strand manage to have identical apparent synthesis rates (Lee et al., 1998, Salinas and Benkovic, 2000, Wu et al., 1992).
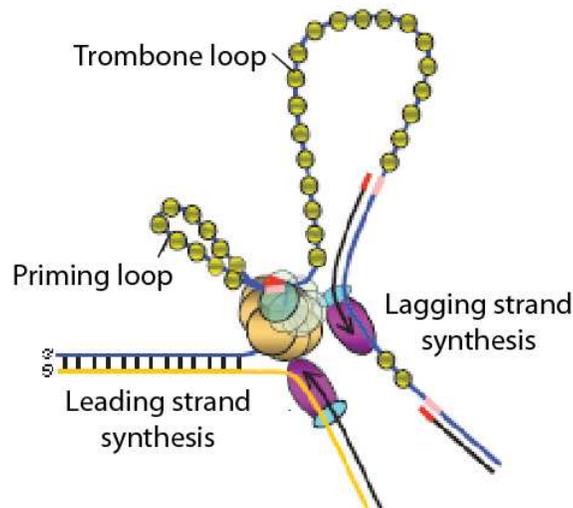
Figure 3: Model of the T7 DNA replication priming system. The leading-template is shown in yellow; the lagging-strand is shown in blue and is coated with gp2.5 (T7 SSB homologue). The trombone loop forms through interactions between T7gp4 (helicase/primase) and gp2.5 upon coupling of the leading and lagging strand synthesis. The priming loop is created between the physically linked primase and helicase domains of the T7 gp4 as a result of ongoing DNA synthesis during primer synthesis (Adapted from Pandey et al., 2009).

This was recently explained by the discovery of a third polymerase on the lagging strand which significantly increases the efficiency of Okazaki fragments synthesis (Figure 1B; Georgescu et al., 2012, Lia et al., 2012, Reyes-Lamothe et al., 2010). The synchronicity is also maintained by the interaction between the DnaG primase and the DnaB helicase which acts as a molecular

break to prevent leading-strand synthesis from outpacing lagging-strand synthesis (Lee et al., 2006, Tanner et al., 2008, Stano et al., 2005). The DnaG primase primes a rapidly moving ssDNA template every 1-2 s (Mitkova et al., 2003) thanks to a ssDNA binding site close to the active site that can take different conformations to allow either DNA to slide across the DnaG surface or capture the template for priming (Corn et al., 2008). It was recently shown using the T7 replication proteins (where primase and helicase activity are coupled in one protein), that RNA primers are made 'on the fly' during ongoing DNA synthesis by forming a priming loop (Figure 3) that keeps the nascent primer within physical reach of the lagging-strand polymerase, thereby avoiding replisome pausing during primer synthesis (Pandey et al., 2009). When the primer is synthesized, it is handed-off to the lagging-strand polymerase and the priming loop becomes part of the trombone loop.

While Pol III never dissociates from the leading strand, synthesis of the Okazaki fragments requires dissociation and exchange of the Pol III HE core and the β_2 sliding clamp on the lagging strand (Lia et al., 2012, Tanner et al., 2011). This dynamic process is coordinated by constantly changing protein-DNA and protein-protein interactions. In order to perform these different tasks (open the sliding clamp, threading DNA through the clamp, closing and releasing the clamp on DNA), intermolecular interactions are modulated by alterations in conformation within the γ clamp loader which interacts with both the sliding clamp and DNA (Bloom, 2006). These changes of conformation are thought to be modulated via ATP binding/hydrolysis at some or all of the three ATP binding sites of the γ complex (Bloom, 2006). The number of ATP molecules bound to the γ complex (clamp loader) at each step of the clamp loading and the order in which each ATP binding site is occupied is not firmly defined yet. It was shown that it is the loss of ssDNA at the end of an Okazaki fragment that triggers the release of Pol III core from the β-clamp and DNA (Figure 4; Georgescu et al., 2009). The OB-fold domain in the α subunit, which binds both ssDNA and the β-clamp, acts as

a sensor that modulates the affinity of Poll III for β-clamp in response to ssDNA (Georgescu et al., 2009, Wing et al., 2008). When the OB domain encounters a nick site at a finishing Okazaki fragment, it no longer binds to ssDNA and changes its conformation so that the α subunit loses its affinity for the β-clamp (Figure 4).
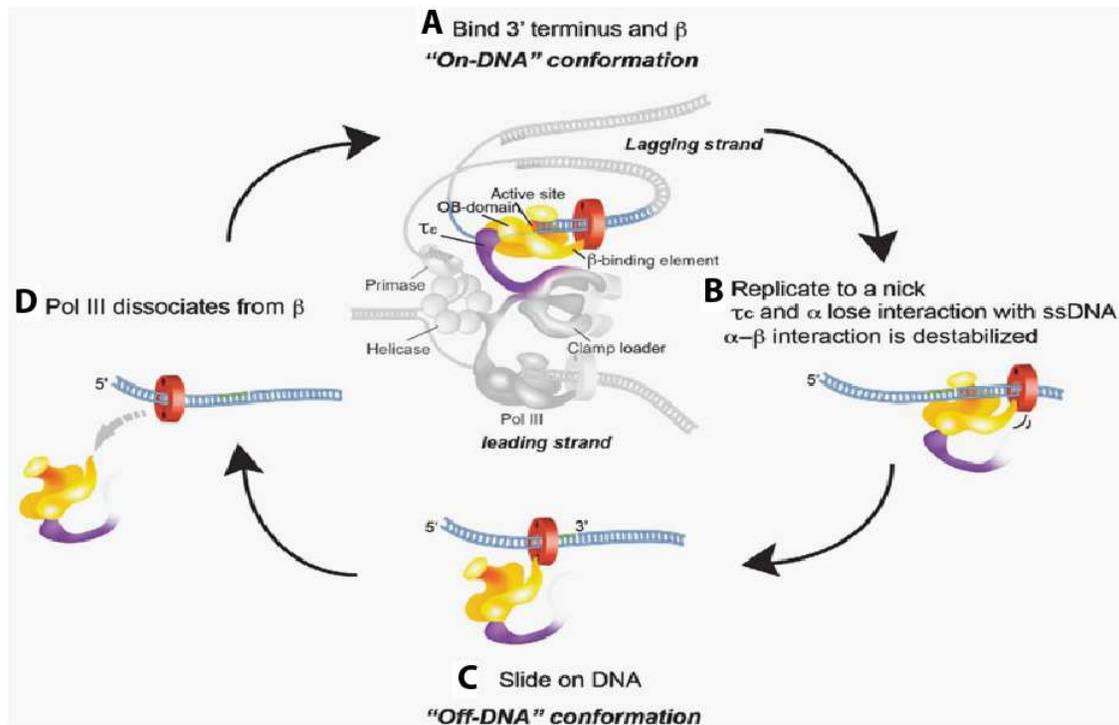
Figure 4: Collision release model of the polymerase during Okazaki fragment synthesis on the lagging strand. (A) At a replication fork, the Pol III α subunit is bound to the β-clamp through its C-terminal arm to which the OB domain is attached. The τ subunits bind Pol III through the C-terminal of α and DnaB helicase. Binding of the OB domain and τ to ssDNA leads to a tight interaction with DNA and β. (B) The OB domain and τ no longer binds ssDNA on finishing Okazaki fragment. (C) τ no longer holds the Pol III core on DNA, Pol III slides off dsDNA, and the α-β interaction is destabilized due to the loss of ssDNA binding activity of the OB domain. (D) Pol III is released from β (Adapted from Georgescu et al., 2009).

In addition, the τ subunit of the γ clamp loader interacts with the α subunit close to the OB domain and ssDNA. In the presence of a nick site, τ loses its affinity for ssDNA and no longer holds Pol III core to DNA (Georgescu et al., 2009). Therefore the τ subunit acts as a switch to

enhance Pol III binding at a primed DNA site, but not at a completed Okazaki fragment or nicked site, and the OB domain is a sensor for sliding clamp binding.

In order to complete lagging strand synthesis, RNA primers must be removed and replaced by DNA prior to ligation to the neighboring DNA fragments (previous and next Okazaki fragments). The DNA polymerase I (Pol I) is the major factor that converts RNA primers into DNA (Okazaki et al., 1971). Pol I has a small 5'-3' exonuclease domain and a second large domain containing both the the 5'-3' polymerase and the 3'-5' proofreading exonuclease (Joyce and Grindley, 1984). RNase H is another factor in removing the RNA primers but was shown to be dispensable (Fukushima et al., 2007). It is thought that upon the completion of each Okazaki fragment, Pol III core disengages from its β-clamp and is replaced by Pol I (Lopez de Saro and O'Donnell, 2001). After conversion of RNA primers into DNA, these DNA are ligated with neighboring DNA fragments by DNA ligase (Lehman, 1974) which also interact with the β-clamp (Lopez de Saro and O'Donnell, 2001).

1.2.5 Replication at DNA lesions

DNA lesions occur naturally from endogenous and/or environmental factors. Although the lesions are usually repaired before replication, some lesions escape the cell's repair mechanisms and remain present at the time of replication. The cell has two lesion tolerance mechanisms that allow the replisome to bypass damaged DNA, the error-prone translesion synthesis (TLS) pathway which uses specialized polymerase, and the error-free damage avoidance (DA) pathway which encompasses multiple processes related to homologous recombination.

Three alternative DNA polymerases can perform the translesion synthesis (TLS) pathway: Pol II, Pol IV and Pol V (Delmas and Matic, 2006). Pol V is part of the SOS response to DNA damage whereas Pol IV is present at constitutive levels (Furukohri et al., 2008,

Goodman, 2000, Schlacher et al., 2006). To bypass DNA lesions, the DNA Pol III core is replaced in the replisome by DNA polymerase IV (Pol IV) or Pol V on the β-clamp sub-assembly of the HE for translesion synthesis (TLS; Furukohri et al., 2008). These polymerases are able to synthesize DNA across the lesions, after which Pol III can resume its activity (Friedberg et al., 2002). It was shown that interactions between Pol III and Pol IV also occur and enable minimal interruption of replication at DNA lesions (Furukohri et al., 2008). All *E. coli* DNA polymerases bind and compete for the same binding domain of the β-clamp which acts as a polymerase switch (Burnouf et al., 2004, Lopez de Saro et al., 2003). Therefore the structure of DNA leads to a change in conformation that promotes the release and re-loading of the appropriate polymerase in a diffusion–dependent manner; the cellular translesion polymerase concentration being up-regulated during the SOS response initiated by DNA damage (Delmas and Matic, 2006 and references therein).

The DA pathway is less characterized than the TLS pathway. It is only recently that a method was developed to introduce site-specific lesions in the genome of a living cell and study the molecular basis of each pathway independently (Pages et al., 2012). DA uses the daughter strand gap-filling pathway (Rupp and Howard-Flanders, 1968) and a recA-independent pathway likely to involve the template switching mechanism (Pages et al., 2012).

It was recently shown that the replisome has the inherent ability to stay associated with the DNA at a single lesion and reinitialize leading-strand synthesis downstream of the damage without the need for replication restart proteins and SOS-response induction (Yeeles and Marians, 2011). Therefore the replisome has an inherent tolerance to punctual DNA lesions but can also dissociate under acute replication stress (i.e. multiple DNA lesions) where the SOS-inducible systems such as TLS and DA are required for survival.

1.2.6 Termination of DNA replication

1.2.6.1   Features of the replication termination components

In *E. coli*, forks moving in opposite directions on the circular chromosome meet in the termination region of the chromosome opposite to *oriC* at specific sites rather than by random collision (Bird et al., 1972, Germino and Bastia, 1981, Louarn et al., 1977). *In vivo* marker frequency experiments initially established that termination occurs at two specific loci in a direction dependent manner. Each site is oriented to block either the clockwise moving fork or the counter clockwise moving fork (de Massy et al., 1987, Hill et al., 1987, Hill et al., 1988a, Kuempel et al., 1977, Louarn et al., 1979). The insertion and deletion mutants used to map the termination region enabled the discovery that one of the two termination foci was associated with the trans-acting gene responsible for fork arrest, the terminator protein Tus (reviewed in Neylon et al., 2005, Hidaka et al., 1989, Hill et al., 1989, Kobayashi et al., 1989, Sista et al., 1989). The 5'-end of *tus* mRNA was shown to be within the *Ter* site of this foci which is now called *TerB* (Roecklein et al., 1991). As a result, the binding of Tus to *TerB* inhibits its own transcription by blocking RNA polymerase access to the promoter. Further mapping of these two loci identified the presence of two polar termination sites (*Ter*) with the same orientation in each locus (now called *TerA, TerB, TerC* and *TerD*). The four *Ter* sites share a highly conserved core sequence (Francois et al., 1989, Hidaka et al., 1988, Hill et al., 1988b). This core sequence was used as a probe to screen the Kohara λ bacteriophage library of *E. coli* and a fifth *Ter* site was detected (*TerE*) and shown to have fork arrest activity *in vivo* (Hidaka et al., 1991). This brought up the number of *Ter* sites to five on the chromosome.

The high sequence similarity between these *Ter* sites prompted the search for more *Ter* sites in the Genbank database using a 23 bp consensus sequence (Sharma and Hill, 1992). A sixth *Ter* site (*TerF*) was identified and assayed for fork arrest activity using the *Ter* assay (Horiuchi and Hidaka, 1988). However, the oligonucleotide used in these experiments did not

carry the mutation at position 18 and resulted in an overestimation of the binding affinity and fork arrest activity of *TerF* (Coskun-Ari and Hill, 1997, Sharma and Hill, 1992). A seventh *Ter* site (*TerG*) was also identified by a sequence similarity search (Neidhardt, 1987). A mutational analysis of *TerB* identified the critical residues involved in Tus binding and fork inhibition activities and enabled the derivation of a new 11 bp consensus sequence from which three more *Ter* sites (*TerH*, *TerI* and *TerJ)* were identified in the *E. coli* GenBank database but were not assayed for DNA replication arrest activity (Coskun-Ari and Hill, 1997). This study brought the number of *Ter* sites to ten (referred as primary *Ter* sites throughout this manuscript) which are arranged in two clusters of five sites flanking the centre of the terminus region (Figure 5A). All *Ter* sites within a cluster are oriented with the same polarity (impedes forks approaching from only one direction) and the two clusters have opposite orientation. This arrangement creates a fork trap allowing the two replisomes to enter but not to exit the termination region, avoiding forks progressing in the terminus-to-origin direction (Figure 5A; Coskun-Ari and Hill, 1997, Neylon et al., 2005). *TerA*, *D*, *E*, *H* and *I* are orientated to block anti-clockwise moving forks while *TerB*, *C*, *F*, *G*, and *J* are oriented to block clockwise moving forks. More recently, four new *Ter* sites (*TerK*, *L*, *Y* and *Z*) were discovered using a 14 bp consensus sequence covering positions 6 to 19 in the *Ter* core sequence (Figure 5A-B; Duggin and Bell, 2009). One of these newly identified *Ter* sites (*TerK*) is located within the previously defined termination region encompassing the ten primary *Ter* sites and the other three are proximal to the origin (Figure 5A). Interestingly, two of them (*TerZ* and *Y*) are oriented to block origin-to-terminus moving forks about 500 and 1000 kbp downstream from the origin but their function is unclear.
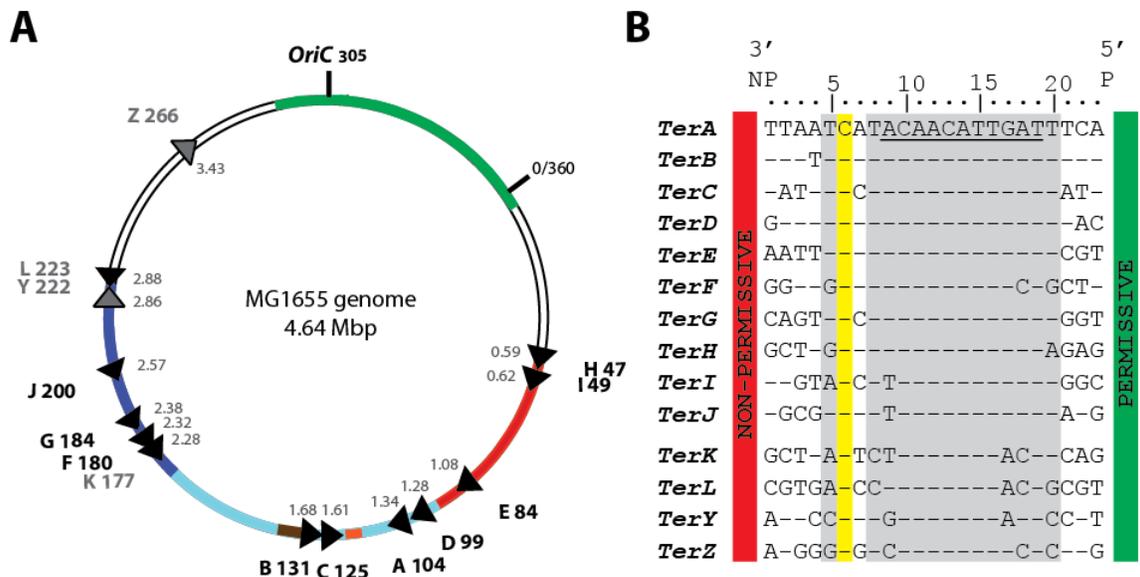
Figure 5: *Ter* sites position, orientation and sequence in *E. coli* MG1655. (A) Map of the *E. coli* genome indicating the position and orientation of the 14 *Ter* sites identified (Duggin and Bell, 2009). The point of the arrow indicates the non-permissive face of the Tus-*Ter* complex. *TerA*, *TerD*, *TerE*, *TerI* and *TerH* are oriented to block anti-clockwise moving forks whereas *TerC*, *TerB*, *TerK*, *TerF*, *TerG*, *TerJ* and *TerL* are oriented to block clockwise moving forks. *TerY* and *TerZ* (white arrows) are oriented to block the left replisome moving in the origin-to-terminus direction. The outer circle represents the chromosome domains (green: *oriC* domain, white: unstructured domains, red: right domain, dark blue: left domain and light blue: termination domain) established by Seitz et al. (2000) and Valens et al. (2004). The numbers inside the circle indicate *Ter* site positions in Mbp. (B) Sequence similarities of the 14 *Ter* sites. The conserved cytosine at position 6 is highlighted in yellow. Sequences are oriented with their non-permissive face (NP) on the left. The shaded nucleotide sequences correspond to the nucleotides interacting with Tus (Kamada et al., 1996) and the underlined sequence is the 11-bp core sequence determined by Coskun-Ari and Hill (1997).

The affinity of Tus to *TerB* has been well characterized (Gottlieb et al., 1992, Mulcair et al., 2006, Skokotas et al., 1995, Skokotas et al., 1994). The equilibrium dissociation constant (K_D) of the Tus-*TerB* complex was reported to be within the nanomalar to subpicomolar range depending on the buffer conditions used (Gottlieb et al., 1992, Mulcair et al., 2006, Neylon et al., 2000). Coskun-Ari and Hill mutated each nucleotide in the core sequence of *TerB* from position 6 to position 21, and tested the efficiency of these mutants to bind Tus and arrest replication forks (Coskun-Ari and Hill, 1997). Some of these mutations reflected the structure of other *Ter* sequences (i.e. the T to G mutation at position 18 in the *TerB* core sequence

matches the *TerF* core sequence), and provided insight into the affinity of other *Ter* sites. Based on these data, *TerA-D*, *TerE* and *TerG* were predicted to be strong *Ter* sites, *TerH* to be a moderately strong site and *TerF*, *I* and *J* to be weak sites relative to *TerB* (Coskun-Ari and Hill, 1997).

It is only fairly recently that the *in vivo* fork arrest activity of the 14 *Ter* sites has been investigated in both a plasmid context and the chromosome (Duggin and Bell, 2009). Out of these 14 *Ter* sites, paused forks were only observed at *TerA*, *B*, and *C* and to some extent at *TerD*, *G, H* and *I* under normal growth conditions. The remaining *Ter* sites (*TerE, F*, *J, K, L, Y* and *Z*) were classified as pseudo-*Ter* sites (*pTer*) sites since they did not encounter a replisome under natural Tus expression level or were not functional (Duggin and Bell, 2009). It is only upon Tus over-expression that fork pausing was detected at the *pTer* sites and to a small extent at *TerK* and *L*. This was surprising since an increase in cellular Tus concentration should have resulted in a tighter fork trap and a decrease in the proportion of forks reaching the outer *Ter* sites. These observations could be explained if the inner most *Ter* sites are already saturated at wild type Tus level. In this case, overexpressed Tus do not reduce the frequency of forks arriving at the outer and weaker *Ter* sites but rather increases their intrinsic efficiency in fork arrest as they become more heavily occupied by Tus.

In order to determine the intrinsic ability of *Ter* sites to halt replication, each *Ter* site was cloned into the pACYC184 plasmid, a vector supporting unidirectional replication, so that completion of replication relied on the replisome ability to pass through the Tus-*Ter* complex cloned in the non-permissive orientation (Duggin and Bell, 2009).
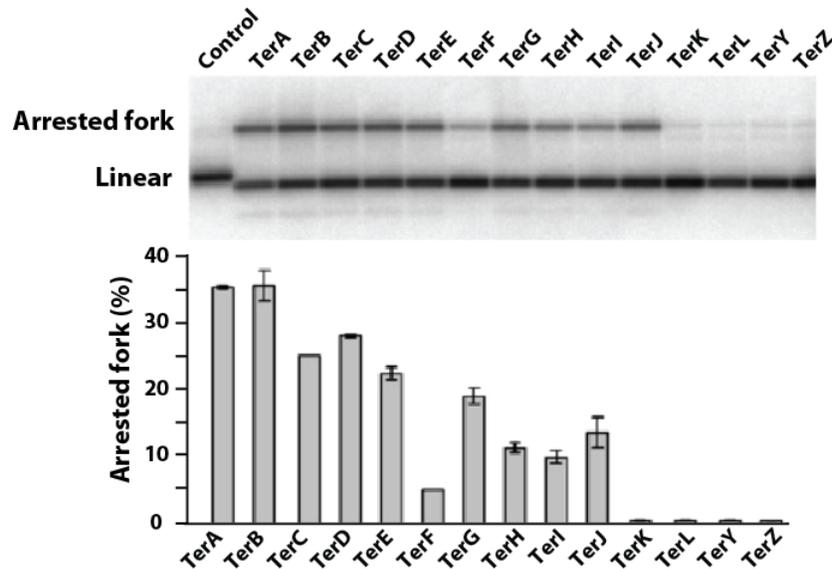
Figure 6: Fork pausing efficiency of *E. coli Ter* sites. Each *Ter* site was cloned in pACYC184 so that the unidirectional replication fork in the plasmids would approach the blocking orientation of the *Ter* site. The efficiency of pausing was quantified as the ratio between linear and forked DNA revealed by Southern blotting (Reproduced from Duggin and Bell, 2009).

All the ten primary *Ter* sites (*TerA-J*) gave rise to arrested fork intermediates but they significantly differed in their efficiency of pausing in accordance with the predicted binding affinity from the *TerB* mutation analysis (Coskun-Ari and Hill, 1997, Duggin and Bell, 2009). The strongest sites were *TerA*, and *TerB* which blocked about 35 % of forks, followed by two groups of *Ter* sites having similar efficiencies; one group comprised *TerC*, *TerD*, *TerE* and *TerG,* and the second group comprised the less efficient *TerH*, *TerI* and *TerJ* (data reproduced in Figure 6). *TerF* had the lowest efficiency of the ten primary *Ter* sites in fork pausing and the last four *Ter* sites *TerK*, *TerL*, *TerY* and *TerZ* were quasi-unable to arrest *in vivo* replication fork (1 % of arrested fork). It was suggested that the *pTer* sites might be able to inhibit fork progression in artificial conditions but not in their wild-type chromosomal context. Their exact role in the fork trap remains unclear (Duggin and Bell, 2009).

The polarity of replication termination was originally explained by the rather unusual binding of Tus to *Ter* DNA (Kamada et al., 1996). Tus has two (amino-and carboxy-) domains both classified as α + β structures (Figure 7A). The two central β-sheets from the two domains form a central large cleft through extensive contacts with the bases and backbone of the *Ter* DNA (Figure 7B-C; Kamada et al., 1996). Two interdomain β-strands intercalate deeply into the major groove deforming the B-form of *Ter* DNA. The Tus-*Ter* contacts are asymmetrically distributed along the *Ter* DNA; the two strands interact with Tus at the non-permissive (i.e. blocking) face and only one strand interacts with Tus at the permissive (non-blocking) face (Figure 7B). The α-helical regions from both domains are concentrated at the non-permissive face (Figure 7A) and are thought to protect the interdomain β-strands from direct contact with the unwinding DnaB helicase (Kamada et al., 1996). Nevertheless, these properties did not seem to be sufficient to explain the polarity of the Tus-*Ter* complex as a Tus mutant with reduced *Ter* binding affinity *in vitro* was found to still have some fork arrest activity *in vivo* (A173V; Skokotas et al., 1994). Conversely other Tus mutants with a maintained affinity for *Ter* (E47Q (Henderson et al., 2001), E49K (Skokotas et al., 1995)) had a reduced ability to halt DNA replication *in vivo*. These discrepancies suggested that another factor contributed to the polarity of fork arrest. The mechanism of DNA replication termination has been fueling an interesting debate in recent years.
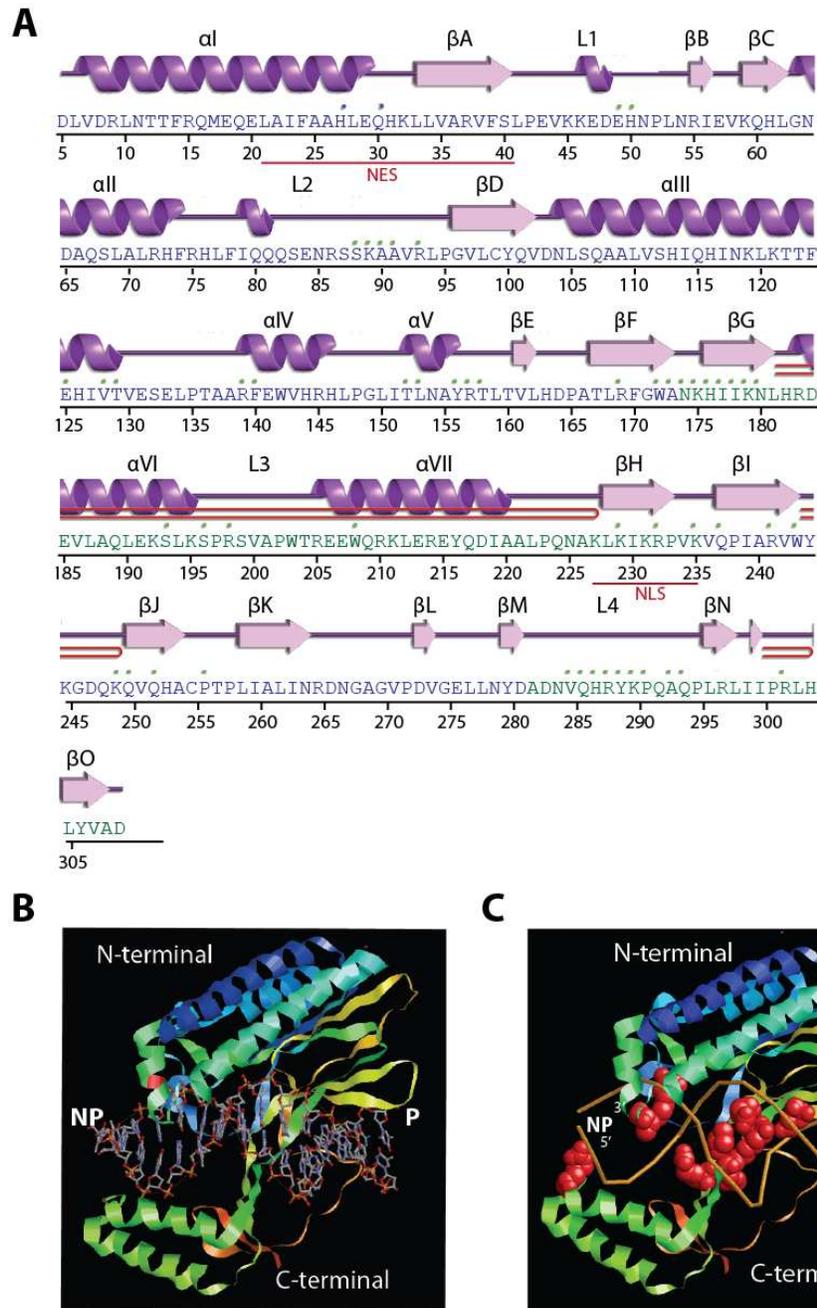
Figure 7: Structure of the Tus-*TerA* complex. (A) Secondary structure of Tus (PDB ID 2I05). Helices are named as α, β-sheets as β and loops as L. NES: nuclear export signal, NLS: nuclear localization signal (Kaczmarczyk et al., 2010). Blue dots and green dots are residues that contact with metals and DNA respectively as per the PDBsum database. Residues in blue form the N-terminal domain and residues in green form the C-terminal domain. The double red lines represent beta hairpins. (B) Tus-*TerA* structure (PDB ID 2I05) with F140 in red. (C) Sequence-dependent interactions in Tus-*TerA*. Residues in contact with bases are shown as red spacefill spheres of Van der Waals radius and include R198, K89, T136, R232, V234, Q237, K175, H176, Q252, A173, Q250. Only *TerA* backbone is represented for clarity purposes.

1.2.6.2   DnaB-Tus and Tus-*Ter*-lock mediated fork arrest

The Tus-*Ter* complex acts primarily as an inhibitor to the DNA unwinding activity of the DnaB helicase which is at the forefront of the replisome. A specific protein-protein interaction between Tus and DnaB was suggested early on based on the observations that Tus could arrest a variety of helicases but not all of them. However, discrepancies in the literature about which helicases were arrested by Tus-*Ter* or not prevented drawing a clear answer about a specific Tus-DnaB interaction. For example Bedrosian and Bastia (1991) found that Tus-*Ter* impeded SV40 T antigen helicase activity in an orientation-dependent manner whereas Hidaka et al. (1992) found the opposite polarity for the same system. The Tus-*Ter* complex was found to inhibit Rep and UvrD helicases in one study (Lee et al., 1989) but not in others (Khatri et al., 1989, Hiasa and Marians, 1992). Tus-*Ter* was also shown to arrest *E. coli* PriA, eukaryotic helicase B (Hidaka et al., 1992) and *E. coli* RNA polymerases (Mohanty et al., 1996) in a polar fashion. These discrepancies may have arisen from the difference in substrate used since arrest of helicases seemed to be affected by the length of the unwounded substrate, the shorter the substrate, the stronger the pausing (Hiasa and Marians, 1992). The most direct evidence of a physical interaction between Tus and DnaB was obtained with a yeast forward two-hybrid analysis where low levels of interaction *in vivo* were reproducibly detected (Mulugu et al., 2001). Tus mutants were selected for their reduced interaction with DnaB in the absence of *Ter* DNA. Two out of four mutants (P52L and E47Q) could still arrest DnaB. Only one E49Q, retained *Ter*-binding affinity and reduced ability to cause fork arrest in cell extract. This mutant is the strongest evidence of a Tus-DnaB interaction (Mulugu et al., 2001) but there is no affinity or kinetic data on this weak interaction yet.

Mulcair et al. (2006) investigated the possibility that a specific DNA structure created by the unwinding action of the DnaB helicase at the non-permissive face of the Tus-*Ter* complex could be responsible for the observed polarity. Forked oligonucleotides were designed

to mimic the progressive unwinding action of DnaB and the effect of these forked species on Tus affinity was measured by surface plasmon resonance (SPR). The presence of a fork at the permissive face of Tus resulted in the fast dissociation of the protein ($t_{1/2} < 5$ s) whereas a fork at the non-permissive face up to C(6) in the *Ter* core sequence resulted in an increase in affinity of Tus for *TerB* to nearly no dissociation ($t_{1/2}$ of 5300 s for forked *TerB* compared to 124 s for fully ds-*TerB;* Mulcair et al., 2006). The crystal structure of Tus with a forked *TerA* core sequence (identical to *TerB*) showed that the cytosine at position 6 in the core sequence (highlighted in yellow in Figure 5B) moves 14 Å from its normal position to bind tightly in a specific cytosine-binding pocket at the surface of the non-permissive face of Tus (Figure 8; Mulcair et al., 2006). This conformation is called the Tus-*Ter*-lock or TT-lock and could explain the polar feature of the complex. In the crystal structure of Tus in complex with forked *Ter* DNA (Mulcair et al., 2006), it was shown that E49 is a residue involved in the cytosine-binding pocket of Tus and hence in TT-lock formation. Therefore the complete loss of DnaB arrest activity of this mutant observed by Mulugu et al. (2001) could also be the result of impaired TT-lock formation. On the other hand, P52L and E47Q mutants may have retained some TT-lock formation ability that rescued some arrest activity *in vivo* despite lower Tus-DnaB interaction *in vitro*. The P42L mutant was the weakest binder to *TerB* and DnaB and was therefore fully defective in fork arrest. Other discrepancies between either *in vivo* and *in vitro* data, or different assays can be explained by the later discovery of the TT-lock as much as by the Tus-DnaB interaction. Another example is the Tus A173V mutant which had a 130-fold lower affinity for double-stranded *TerB* (gel mobility shift assay) but retained substantial (75 %) replication arrest activity (Saecker, 2001, Skokotas et al., 1995, Skokotas et al., 1994). This residue contacts the core sequence at position 16 (Kamada et al., 1996). Its mutation would lead to a loss in affinity but would not impede lock formation at the non-permissive face and/or the DnaB-Tus interaction.
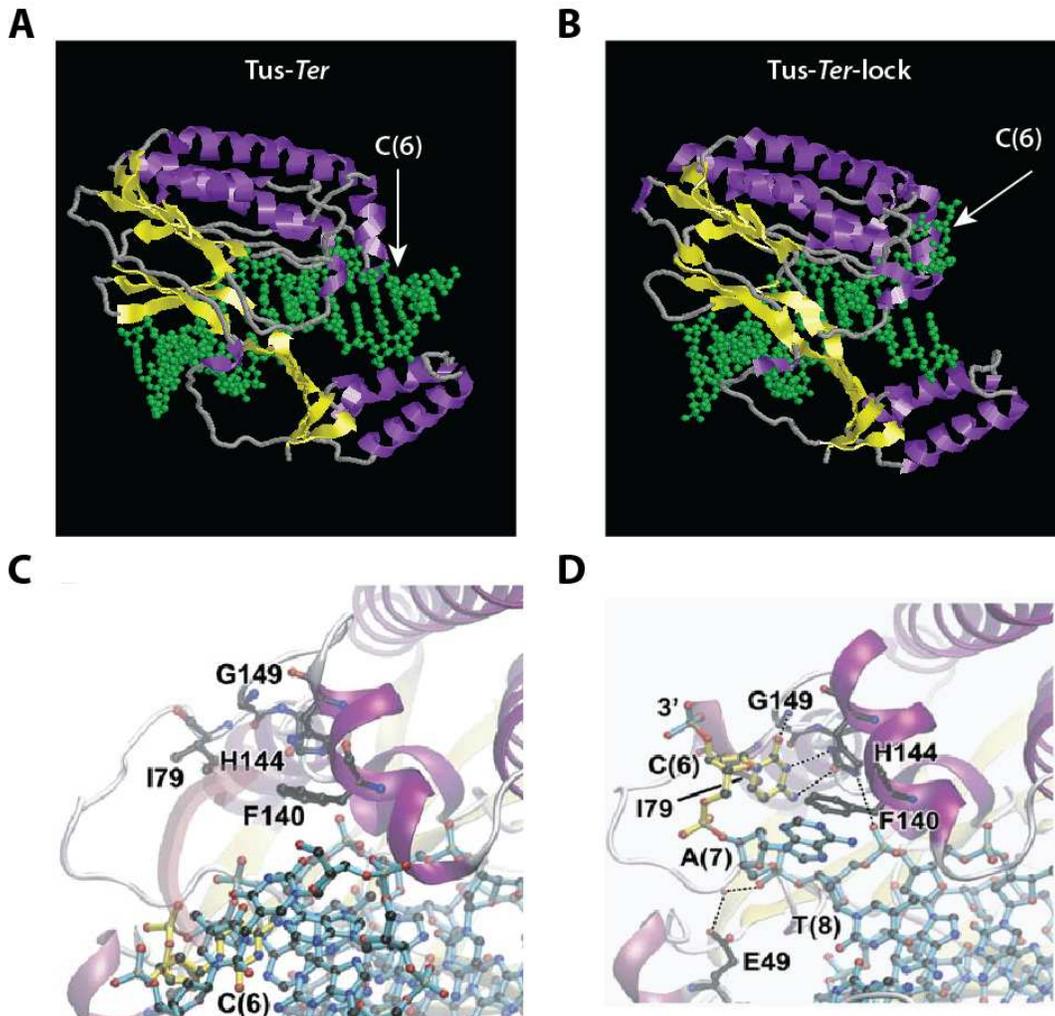
Figure 8: Tus-*Ter* and Tus-*Ter*-lock structures. Structure of Tus in complex with (A) *TerA* (PDB ID 1ECR, Kamada et al., 1996) and with (B) forked *TerA* at the non-permissive face (PDB ID 2EWJ, Mulcair et al., 2006). The DNA molecule is shown in green in a ball and stick representation, α-helices are represented in purple, β-sheets in yellow and loops in grey. (C) Structure of the cytosine binding pocket at the surface of Tus prior to and (D) after DNA unwinding. The locked C(6) is shown in yellow (Adapted from Mulcair et al., 2006).

An elegant strand displacement assay was developed to discriminate DnaB-Tus interaction from TT-lock mediated fork arrest (Bastia et al., 2008). The DnaB helicase sliding onto double-stranded *TerB* was still inhibited by the Tus-*Ter* complex in a polar manner. The polar fork arrest of Tus-*Ter* was maintained without melting of DNA and TT-lock formation. Nevertheless, when the assay was performed with a *TerB* containing a mispaired bubble (C(6) included), the blocking was increased (see Figure S3 in Bastia et al., 2008). This increase in

blocking activity was said to be non-significant due to a high standard error associated with the data obtained with bubbled *TerB*. This variation could be attributable to the suboptimal docking of Tus on the bubble substrate which could have masked and/or affected the base flipping locking mechanism. Indeed, a 10-fold decrease in association rate was observed between fully annealed *TerB* and a bubble containing *TerB* (Mulcair et al., 2006). This lower association rate is likely to be due to the absence of specific contacts between A5 and G6 (mismatched in the bubble *TerB*) with R198 in Tus (Coskun-Ari and Hill, 1997, Kamada et al., 1996). Also, the structure of the bubble *TerB* restricts the complete flipping of C(6) into the cytosine binding pocket and impairs the lock formation (half life of 2240 s for a 5 bp bubble-*Ter* compared to 5300 s for a fully unwounded *Ter* site over 5 bp; Mulcair et al., 2006). This bubble structure in the DnaB helicase assay would therefore result in the underestimation of the TT-lock contribution to polar fork arrest. Therefore the major mechanism responsible for polar fork arrest is still under debate.

1.2.6.3   The alternative site of termination, the *dif* site

Hendrickson and Lawrence (2007) proposed that the *dif* locus, which is located between *TerA* and *TerC* (Figure 5A), is the only site of termination and that the fork trap is instead used to stall repair-associated or other non-oriC initiated forks. The *dif* site is the site of action of the XerCD site-specific DNA recombinase which is required for chromosome unlinking and segregation (Blakely and Sherratt, 1994, Grainge et al., 2011). The *dif* hypothesis was raised based on the skewness of base compositional differences of the leading and lagging strands which is correlated with replication direction and can therefore predict termination sites (Grigoriev, 1998, Lobry, 1996, Rocha, 2004, Arakawa et al., 2007). It was shown that the skew switches from one strand to another at or very near the *dif* site, but not at *Ter* sites, and it was suggested that *Ter* sites are used to halt replication forks originating from DNA repair events

rather than from *oriC* (Hendrickson and Lawrence, 2007). The occurrence of termination at the *dif* site (in between the two innermost *Ter* sites, Figure 5A) was not excluded from the other studies looking at fork arrest at the inner *Ter* sites because they lacked the resolution to discriminate between the *TerC* and the *dif* site (Hendrickson and Lawrence, 2007). Neverthess, the analysis of replication intermediates in the *TerC-dif* region by 2D-gel analysis found no significant pausing near the *dif* site but significant fork arrest at *TerC*, supporting the fork trap model and not the *dif* model (Duggin and Bell, 2009). The fork trap model was also supported by the latest genomic compositional skew analysis of the termination region (Kono et al., 2012) that showed that a single finite termination site at *dif* is not sufficient to explain the genomic compositional bias observed in the published genome sequence, whereas the fork trap can.

1.2.6.4   Replication termination of other circular chromosome

Other prokaryotic chromosomes

While site specific polar replication termination sites are present in both prokaryotes and eukaryotes (see section 1.2.6.5 for site specific termination in eukaryotes), the components mediating fork arrest have little similarities between species, even amongst the prokaryotic taxa. In the archeabacteria, replication termination occurs by random collision of replication forks (Duggin et al., 2011). In the eubacteria, termination of replication has been extensively studied for the gram positive *Bacillus subtilis* and the components of the system have no sequence homology to the *E. coli* system (Bussiere et al., 1995, Duggin, 2006, Duggin et al., 2005, Hastings et al., 2005, Vivian et al., 2007, Wake and King, 1997, Weiss and Wake, 1984). The *B. subtilis* fork trap is composed of nine termination sites (Griffiths et al., 1998) with five *Ter* sites oriented to block clockwise moving fork and four *Ter* sites oriented to block the anti-clockwise moving fork. Each *Ter* site is 30 bp long comprising two imperfect inverted 16 bp repeats (called A and B sites), each binding a dimer of RTP. Fork arrest requires RTP binding

to the B site and the subsequent cooperative binding of a second RTP dimer to the A site (Duggin et al., 1999, Langley et al., 1993, Smith et al., 1994). Also, only forks approaching the B site will be hindered (Smith and Wake, 1992). The molecular mechanism of the RTP-*Ter* complex polarity is thought to reside in a specialized structure that contributes to the mechanism of replication fork arrest in a manner that is independent of its high affinity for DNA but that might require the cooperativity of several assymetrical features of the RTP:DNA complex (Duggin, 2006, Vivian et al., 2007). Polar *Ter* sites are also present in the R6K plasmid which carries only two copies located asymmetrically with respect to the two main origins of replication (Horiuchi and Hidaka, 1988). The two *Ter* sites of the R6K plasmid share 15 and 12 nucleotides with *TerA* and *TerB* respectively and are also organized to form a fork trap (Neylon et al., 2005).

Mitochondrial chromosome

Mitochondria have evolved from α-proteobacterium and have conserved their circular chromosome (16.6 kbp) following endosymbiosis in eukaryotic cells (Gray, 2012). The chromosome has two origins where replication fires unidirectionally. For some time, it was believed that leading strand synthesis (H strand) was initiated from one origin (OH) and that lagging strand synthesis (L strand) fired in the opposite direction once the leading strand exposed the second origin (OL). More recent work found Y-arc shaped DNA molecules in between the two origins by 2-D agarose gel electrophoresis, indicative of the classic synchronous leading- and lagging-strand replication forks (Holt et al., 2000). Subsequent work proposed a third model called RITOLS (ribonucleotide incorporation throughout the lagging strand) where the lagging strand is initially laid down as RNA before being converted to DNA (Yang et al., 2002). These three models are currently under debate (reviewed in Kasiviswanathan et al., 2012). The H strand frequently stalls 700 bp downstream of OH

producing a triple stranded structure called the D-loop. At the 3'-end of the D-loop, a short (15 bp) termination-associated sequence (TAS) was identified (Brown and Clayton, 2002, Clayton, 1991) and shown to bind a 48 kDa DNA-binding protein in bovine mitochondria (Madsen et al., 1993). In other species, it was speculated that one of the uncharacterized proteins of the mTERF family (for mitochondrial termination factor) responsible for transcription termination, replication-termination regulation and protein synthesis, may be the long-sought-after TAS-binding protein (Falkenberg et al., 2007, Pellegrini et al., 2009). This was supported by the finding that an mTERF orthologue in sea urchin (mtDBP) binds to a site at the 3' end of the D-loop and acts as a bidirectional contrahelicase (Polosa et al., 2005). It was later shown that mTERF induces DNA replication pausing in human mtDNA in the D-loop and to additional sites upon overexpression (Hyvarinen et al., 2007). Interestingly the binding of mTERF to its substrate is stabilized by a base flipping mechanism of three nucleotides (adenine, thymine and cytosine) forming π-stacking interactions with the mTERF R162 (strictly conserved), F234 and Y288, as well as hydrogen bonds (Yakubovskaya et al., 2010). Like Tus in *E. coli*, mTERF1 binding also induces a DNA bend of 25° (20° in *E. coli*) and DNA relaxing (underwinding). However this base flipping mechanism occurs in fully double-stranded binding sites without DNA melting like in *E. coli*.

1.2.6.5   The fork trap, DNA transcription and chromosome organization

Contrary to eukaryotes, transcription and replication are not uncoupled in bacteria but occur simultaneously. A characteristic of the fork trap is that the outermost *Ter* sites (with the exception of *TerY*) are positioned within ORFs whereas the inner *Ter* sites (with the exception of *TerB* which lies within the *tus* promoter; Roecklein and Kuempel, 1992) are positioned between ORFs (Table 1). Since the outermost *Ter* sites are less likely to encounter a fork

approaching the non-permissive face of the complex, they may have another biochemical activity related to transcription-replication coordination.

The Tus-*Ter* complex was shown to mediate the polar arrest of *E. coli* RNA polymerases *in vitro* and *in vivo*, protecting replication termini from RNA chain elongation at the non-permissive face of the complex (Guajardo and Sousa, 1999, Mohanty et al., 1996). Nevertheless, the orientation of most *Ter* sites within ORFs is such that RNA polymerases are unlikely to encounter the blocking face of the Tus-*Ter* (Table 1). Only two *Ter* sites, the predicted weak *TerL* (within *iap* protease gene) and *TerB* (within the *tus* promoter) are presenting the non-permissive face to the RNA polymerase. Tus blocks its own transcription and prevents the RNA polymerase from binding to the promoter.

In yeast, polar replication fork barriers (RFBs) have been identified within ribosomal DNA (rDNA) intergenic spacer and the mating type region in *S. pombe* (Dalgaard and Klar, 2001, Eydmann et al., 2008). rDNA *Ter* sites are recognized by Fob1p in *S. cerevisiae* and by the transcription terminator Reb1 or the genome-stabilizing protein Sap1 in *S. pombe* (Eydmann et al., 2008, Mohanty and Bastia, 2004, Krings and Bastia, 2005). The protein Rft1 recognises the RTS1 element in the mating type region (*mat1*). In *S. pombe*, rDNA intergenic regions contain four *Ter* sites which are bound to either Reb1 or Sap1 but only Sap1 bound *Ter* site is polar (Krings and Bastia, 2005). These sites are oriented to promote the selective pausing of the fork that moved in opposite direction to the direction of transcription. Interestingly Sap1 is also a chromatin–organizing protein, a mating type switch and a replication origin binder (Krings and Bastia, 2005, Bastia and Singh, 2011 and references therein). On the other hand Reb1 was shown to mediate interaction between *Ter* sites resulting in chromosome kissing control of DNA replication termination (Bastia and Singh, 2011, Singh et al., 2010). Reb1 also has transcription activation functions and appears to coordinate replication and transcription processes. It was suggested that chromosome kissing through *Ter* sites could help Reb1

binding to promoters of sporulation protein genes and modulate cell-cycle-directed gene expression (Bastia and Singh, 2011). RFBs were also found at tRNA genes (Deshpande and Newlon, 1996) and RNA polymerase II-dependent transcription units in *S. cerevisiae* (Azvolinsky et al., 2009). Fachinetti et al. (2010) identified 71 termination regions (*TER*, 5 kb each) in *S. cerevisiae* with the majority (62/71) being polar pause elements. Interestingly, *TER* sites were occupied by the topoisomerase 2 before fork arrival and participated to genome stability. The replication termination sites in eukaryotes have therefore developed additional functions in transcription-replication regulation and chromosomal organization.

In *E. coli*, *Ter* sites are generally located at the border of the major chromosomal domains with the exception of *TerB*, *TerC* and *TerJ* (Figure 5A; Scolari et al., 2011, Valens et al., 2004). In these interdomain regions, transcription is directly linked to genome spatial organization mediated by Fis and N-NS binding (Scolari et al., 2011). The edges of the *Ter* domain delimited by *TerA*, *D* and *E* on one side and *TerF*, *G* and *K* on the other side (Figure 5A) are flanked by the whole flagella regulon and key regulators of biofilm formation (Scolari et al., 2011). Their symmetrical distribution was suggested to help maintain the relative proportion of flagellar proteins during replication (Scolari et al., 2011). Although there is no evidence of Tus-*Ter* involvement in transcription, the occurrence of the outer *Ter* sites within ORFs along with the conservation of polar fork barriers in higher organisms and their role in transcription-replication regulation suggest that the *E. coli* Tus-*Ter* complex could also have a role in transcription (Table 1).

The position of a *Ter* site within an ORF varies for each *Ter* site (i.e. can be found at the beginning, middle or end of the ORF). It was shown in eukaryotes that a significant proportion of TF binding sites are found within ORFs and are correlated with non-coding RNAs (ncRNA; Cawley et al., 2004). In bacteria, ncRNA are at the cross roads between regulons of many biological pathways including transcription reprogramming, carbon

metabolism, iron homeostasis, envelope homeostasis and toxicity (reviewed in Repoila and Darfeuille, 2009) and are usually expressed as part of the stress response regulatory system (Gottesman, 2005). Although there is no evidence of Tus acting as an accessory transcription factor other than for its own regulation at *TerB*, *Ter* associated genes fall within the above cited pathways (Table 1).

Table 1: Position and orientation of *Ter* sites within transcribed regions.

Ter site	ORF and *Ter* polarity	Gene identification	Gene function	Comments	Reference
H		cusA (NP_415107)	Copper/silver efflux system	Inner membrane protein in conjunction with CusB, CsuC and CsuF	Franke et al., 2003
I		*entC* Isochorismate synthase (NP_415125)	Siderophore enterobactin (iron carrier) biosynthesis	Secreted protein, role in biofilm formation. Regulated by Fur (ferric uptake regulator)	Hancock et al., 2010
E	19 15	Between *efeU* and *efeO* (NP_415537)	Defective Iron transporter	Functional in *E. coli* O157:H7 but not in K12 strains. Also regulated by Fur.	Baichoo and Helmann, 2002, McHugh et al., 2003
D	180 300	Between *narK* (NP_415741) and *narG* (NP_415742)	Nitrate reductase (Anaerobic respiration)	*narG* is part of the operon encoding respiratory nitrate reductase and *narK* encodes a putative nitrite transporter	Hartig et al., 1999
A	17 148	Between *yciM* (NP_415796) and *pyrF* (NP_415797	Nucleotide metabolism-Translation initiation factor	*pyrF* encodes an outer membrane protein and is part of an operon with *yciH* (unknown function, similarities with a translation initiation factor). *yciM* has a role in biofilm formation	Jensen et al., 1984, Lomakin et al., 2006, Niba et al., 2007
C	134 49	Between *yneE* (NP_416037) and *uxaB* (YP_025302)	Swarming - Carbohydrate metabolism	*yneE* encodes a membrane protein and *uxaB* is part of the exu regulon coding for a dehydrogenase	Blanco and Mata-Gilsinger, 1986, Inoue et al., 2007

Ter site	ORF and *Ter* polarity	Gene identification	Gene function	Comments	Reference
B	43	Between *rstB* (NP_416126) and *tus* (NP_416127)	Sensory kinase - Replication termination	Within *tus* promoter. *rstB* encodes a transmembraneous sensory histidine kinase that phosphorylate the transcription regulator RstA	Hidaka et al., 1989, Hill et al., 1989, Roecklein et al., 1991, Yamamoto et al., 2005
K		*yeiG* (NP_416686)	Unknown function	*TerK* is directly at the begin of the *yeiG* (unlnown function, may possess enzymatic activity against antibiotics)	Novikova et al., 2007, Soo et al., 2011
F		*rcsC* (NP_416722)	Sensory system for capsule synthesis and cell division	Control expression of the capsule operon *cps* and cell division control gene *ftsZ*	Carballes et al., 1999, Stout and Gottesman, 1990
G		*menH* (or *yfbB*) (NP_416766)	Manaquinone synthesis pathway (essential electron carrier)	SHCCHC synthase (respiration)	Jiang et al., 2008
J		*maeB* (or *ypfF*) (NP_416958)	Malic enzyme (glycolysis/glucogenesis and TCA cycle)	*TerJ* is located towards the end of *maeB*	Bologna et al., 2007
Y	10 55	Between *ygbM* (NP_417219) and *ybnN* (NP_417220)	Putative isomerase - permease	*ygbM* sequence has homology to xylose isomerase genes and *ygbN* encodes an inner membrane protein from the gluconate transporter familly	Daley et al., 2005
L		*iap* (NP_417233)	Protease	Isozyme conversion of alkaline phosphatase	Ishino et al., 1987
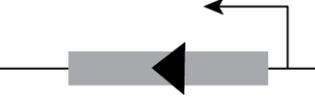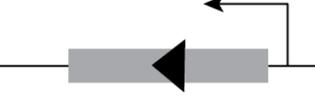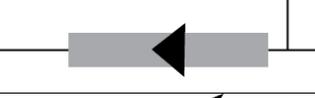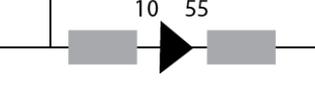
Ter site	ORF and *Ter* polarity	Gene identification	Gene function	Comments	Reference
Z		*fmt* (NP_417746)	Methionyl-tRNA formyltransferase	Essential role in translation initiation	Guillon et al., 1993
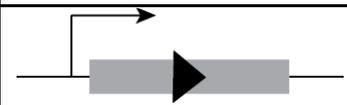
Ter sites are represented by a black triangle with the tip of the triangle representing the non permissive face of the complex. The grey box represents the ORF. The thin black arrow represents the direction of ORF transcription. The numbers indicate the distances in bp between *Ter* and ORFs when *Ter* was found in between two ORFs. Gene numbers correspond to the RefSeq database (NCBI). The first and second genes cited are upstream and downstream of the *Ter* site respectively when reading the top strand of the chromosome in the clockwise direction. The red line highlights the limit between the two oppositely orientated *Ter* clusters.

Nine *Ter* sites are located within ORFs and five in between two ORFs. The genomic context of all *Ter* sites span a variety of metabolic pathways including metal ion transport (*TerH*, *I*, *E*, and *Y*), sensory system for capsule synthesis and swarming (biofilm; *TerC* and *F*), respiration (*TerD*, *G* and *J*), carbohydrate metabolism (*TerC*), transcription-translation (*TerA*, *B*, *L* and *Z*) and cell division (*TerF*). The identified genes are generally associated with cell response to environmental stimuli (either membrane associated proteins, signaling proteins or secreted proteins). Within the identified gene, two have a putative function (*nark* and *ygbM*) and two have an unknown function (*yciM* and *yeiG*). It is not known whether the genes associated with *Ter* sites are differentially expressed in a *tus*-null mutant and in response to different stimuli.

1.3 About this thesis

In *E. coli,* the molecular mechanism responsible for the polarity of the Tus-*Ter* complex in arresting the DnaB helicase is still being debated. Since the discovery in 2006 of the Tus-*Ter*-lock mechanism, no effort has been made to investigate further its formation *in vivo* and/or with the remaining *Ter* sites. Indeed, most of the work intended to characterize the mechanism of polar fork arrest has been focused on the innermost *Ter* sites (*TerA-C*) leaving the essential question of whether or not the remaining *Ter* sites maintained the same biological properties unanswered. The affinity of Tus for the 14 *Ter* sites has only been inferred from mutational analysis of *TerB* and a direct comparative study of the kinetics of all complexes is lacking for most *Ter* sites. The variation in pausing efficiency observed by Duggin and Bell further justifies the biochemical characterization of all Tus-*Ter* complexes.

The main aim of this work was therefore to characterize the structure of the fork trap through comparative analysis of the properties of the ten primary Tus-*Ter* complexes *in vitro* and *in vivo*. The specific questions addressed are:

- What are the affinity and kinetic parameters of Tus-*Ter* and Tus-*Ter*-lock complexes?

- Are all *Ter* sites able to form a TT-lock?

- Is the TT-lock responsible for fork arrest?

- What is the distribution of Tus on *Ter* sites *in vivo*?

- Are outermost *Ter* sites able to stop a replisome?

Several new methods were developed to answer these questions and monitor protein-DNA interactions. Parts of this thesis have been published and the original maunscipts were adapted to incorporate the development and use of these new methods for the characterization of the Tus-*Ter* complexes.

The general materials and methods are described in Chapter 2.

Chapter 3 describes the development and validation of a medium throughput thermal stability assay for the rapid screening and ranking of ligands. This technique called GFP-Basta was validated with three different proteins and published in Molecular BioSystems. Only the results obtained for Tus were included in this chapter.

Chapter 4 describes the kinetic and affinity characterization of the ten primary Tus-*Ter* complexes and their ability to form a TT-lock. This study provides detailed mechanistic information on the Tus-*Ter* complex. This chapter has been published in Molecular BioSytems.

Chapter 5 presents the development of a new high-throughput method derived from GFP-Basta for the rapid determination of protein stability and ligand binding. This technique called DSF-GTP was applied to the characterization of 12 proteins (published in RSC Advances) but only the validation part of the method using Tus as a model protein was included in this chapter. DSF-GTP was applied to analyse the effect of ionic strength on the stability of the ten Tus-*Ter* and Tus-*Ter*-lock complexes (unpublished data).

Chapter 6 describes an alternative new qPCR-based protein-DNA binding assay that was used to determine the affinity of Tus-*Ter* complexes. This chapter has been published in Analyst.

In Chapter 7, a ChIP-qPCR method was developed to determine the *in vivo* distribution of Tus on *Ter* sites using exogeneous expression of Tus-GFP. Finally, ectopic *Ter* sites were introduced into the *E. coli* genome to determine the effect of the TT-lock on replication dynamics and whether outer *Ter* sites can stop replication forks.

The results from this study are summarized and discussed in Chapter 8, and some perspectives are introduced.

Chapter 2: General material and methods

2.1 Materials

2.1.1 Reagents and enzymes

The specific chemicals, buffers and solutions used are described in the appropriate methods. All antibodies were purchased from Abcam except for the HRP-conjugated goat anti-chicken IgY (103-035-155) which was purchased from Jackson Immuno Research. All enzymes but EcoRI (Promega) were purchased from New England Biolabs, as were T4 DNA ligase, Taq DNA polymerase with Thermopol buffer. Quick load 1 kb DNA ladder and 100 bp DNA ladder, iProof High-Fidelity DNA polymerase and iTaq polymerase were obtained from Bio-Rad.

2.1.2 Growth media and agar

Luria and Burrows (LB) medium

Broth: 10 g peptone, 5 g yeast extract, 5 g NaCl dissolved in 1 litre of $_dH_2O$. For agar plates, 15 g agar (1.5 %) were added.

Overnight Expression Medium

Broth: 60 g of Overnight Expression Media granules (Novagen) dissolved in 1 litre of $_dH_2O$ with 1 % glycerol.

Terrific Broth (TB)

Broth: 12 g peptone, 24 g yeast extract, 4 ml of glycerol dissolved in 900 ml of $_dH_2O$ and autoclaved. Once the media cooled to 60°C, 100 ml of 0.89 M filter sterilized potassium phosphate was added.

2.1.3 Bacterial strains and plasmids

Table 2: Bacterial strain genotypes.

Strain	Description	Genotype	Source
DH12S	Used for cloning and preparation of genomic or plasmid DNA	Φ80d *lacZ* ΔM15 *mcrA* Δ(*mrr-hsd*RMS-mcrBC) *ara*D139 Δ(*ara, leu*)7697 Δ*lac*X74 *gal*U *gal*K *rps*L(Str^R) *nup*G *rec*A1 / F' {*proAB+ lac*I^q *lac*ZDM15}	Invitrogen
BL21(*DE3*) RIPL	Contains λ prophage carrying the T7 RNA polymerase gene under *lac* promoter and contains extra copies of the rare codons argU, ileY, and leuW as well as the proL tRNA genes	B F– *omp*T *hsd*S(r$_B^-$ m$_B^-$) *dcm*$^+$ Tet^r *gal* λ(*DE3*) *end*A Hte [*argU proL* Cam^r] [*argU ileY leuW* Strep/Spec^r]	Stratagene
KRX	Contains the T7 RNA polymerase gene under the control of a rhamnose promoter	[F$^-$ *tra*D36 Δ*omp*P *pro*A$^+$B$^+$ *laq*I^q, Δ(*lac*Z)M15] Δ*omp*T *end*A1 *rec*A1 *gyr*A96(Nal^r) *thi*-1 *hsd*R17 (r$_K^-$, m$_K^+$) e14$^-$ (McrA$^-$) *rel*A1 *sup*E44 Δ(*lac-pro*AB) Δ(*rha*BAD)::T7 RNA polymerase	Promega

Table 3: Parent plasmids used in this study.

Name	Description	Promoter	Selection marker	Source	Reference
pIM013	GFP-cloning cassette	T7 RNA polymerase	Ampicillin	Dr. I. Morin	Moreau et al., 2010
pMM001	His$_6$-Tus	T7 RNA polymerase	Ampicillin	M. Moreau	Dahdah et. al., 2009
pIM033	Used to insert gene under T7 promoter region, without tag	T7 RNA polymerase	Ampicillin	Dr. I. Morin	Unpublished
pPMS1259	His$_6$-Tus-GFP	T7 RNA polymerase	Ampicillin	Dr. P. Schaeffer	Dahdah et al., 2009

2.1.4 Oligonucleotides

All oligonucleotides were purchased from Sigma-Aldrich.

2.1.4.1 Primer sequences used to amplify genomic *Ter-* and *oriC-* containing regions.

The sequences of the primers used to amplify *Ter*-containing genomic DNA regions from *E. coli* strain DH12S by PCR or for qPCR analysis are shown in Table 4. These oligonucleotides were designed to amplify a DNA fragment no longer than 150 bp (except for *TerE* amplicon). The genome of BL21(*DE3*)RIPL and DH12S were not yet fully sequenced at the time of primer design, therefore, the primer sequences were based on the genomic DNA sequence of *E. coli* K12 MG1655, in DNA regions where no mutation was observed in the B7A and DH10B *E. coli* strains. Oligonucleotides were resuspended at 100 µM in TE$_{50}$ (10 mM Tris (pH 7.5), 0.1 mM EDTA and 50 mM KCl).

Table 4: Primer sequences for amplification of genomic *Ter-* or *oriC-* containing regions.

	Forward Primer (5-3')	JCU #*	Reverse Primer (5-3')	JCU #*	Amplicon size (bp)
TerA	CAACCATTAAACCGATTCGCGGTC	72	AGTTGCGATTTCTCCCCTGG	73	145
TerB	TTACCTCTGCCTGACACTACGC	74	TGTTGAGTCGGTCTACGAGATCG	75	123
TerC	CTGCATGTGGCACCTGTTAATGA	76	GCTGTACGTCCGTTGTGCTA	77	123
TerD	GGCATGATGTCGCGCTTTTTTTATG	78	GGGTATTAAGGAGTATTCCCCATGG	79	125
TerE	GAAGTCGCCGTCTGGTTTAT	180	TACGGCGGAAGTTAATGGTC	181	172
TerF	CACATCTTCGGGAGTCGGTTC	82	GGTTGAGTGGTAAACGCTGCTG	83	131
TerG	CCAAGCGAGTACCCCACCAG	84	CACGGTTGTATGTTGATCTCCCA	85	142
TerH	TGAAGGACAAACTGGAAACGCTGA	86	CAGACTACCGCCACCACAAT	87	148
TerI	ATTGCTGGAACGGTTGATTGCG	88	CTCGCCGTCTTTACGTAGCA	89	118
TerJ	GACGATACGACGCACCGATG	90	CTGGTGATGCCGAACATGGAAG	91	150
oriC	CGCACTGCCCTGTGGATAACAA	92	CCCTCATTCTGATCCCAGCTTA	93	115

* JCU numbers included in this table and all the following tables are the reference number of each oligonucleotide in the Schaeffer's laboratory database.

2.1.4.2 Oligonucleotides used for GFP-Basta

Oligonucleotides used for GFP-basta were designed to include the 23 bp *Ter* or *Ter-lock* sequence followed by a stabilizing 10-mer GC rich region in order to elevate their T_m values above 70°C (Table 5). A 34 bp sequence from the *oriC* of *E. coli* MG1255 was used as a non-specific DNA control. All oligonucleotides were resuspended at 200 µM in TE$_{50}$ (10 mM Tris (pH 7.5), 0.1 mM EDTA and 50 mM KCl). Complementary oligonucleotides were annealed by heating 2 minutes at 80°C and progressively cooling down to room temperature.

Table 5: *Ter* and *Ter-lock* oligonucleotides used for GFP-Basta.

	Ter	JCU #	*Ter-lock*	JCU #
TerA	AATTAGTATGTTGTAACTAAAGT*GGGGGCGGGG* TTAATCATACAACATTGATTTCA*CCCCCGCCCC*	97 98	TATGTTGTAACTAAAGT*GGGGGCGGGG* TTAATCATACAACATTGATTTCA*CCCCCGCCCC*	140 98
TerB	AATAAGTATGTTGTAACTAAAGT*GGGGGCGGGG* TTATTCATACAACATTGATTTCA*CCCCCGCCCC*	99 100	TATGTTGTAACTAAAGT*GGGGGCGGGG* TTATTCATACAACATTGATTTCA*CCCCCGCCCC*	140 100
TerC	ATATAGGATGTTGTAACTAATAT*GGGGGCGGGG* TATATCCTACAACATTGATTATA*CCCCCGCCCC*	101 102	GATGTTGTAACTAATAT*GGGGGCGGGG* TATATCCTACAACATTGATTATA*CCCCCGCCCC*	141 102
TerD	CATTAGTATGTTGTAACTAAAT*GGGGGGCGGGG* GTAATCATACAACATTGATTTAC*CCCCCGCCCC*	103 104	TATGTTGTAACTAAAT*GGGGGGCGGGG* GTAATCATACAACATTGATTTAC*CCCCCGCCCC*	142 104
TerE	TTAAAGTATGTTGTAACTAAGCA*GGGGGCGGGG* AATTTCATACAACATTGATTCGT*CCCCCGCCCC*	105 106	TATGTTGTAACTAAGCA*GGGGGCGGGG* AATTTCATACAACATTGATTCGT*CCCCCGCCCC*	143 106
TerF	CCTTCGTATGTTGTAACGACGAT*GGGGGCGGGG* GGAAGCATACAACATTGCTGCTA*CCCCCGCCCC*	107 108	TATGTTGTAACGACGAT*GGGGGCGGGG* GGAAGCATACAACATTGCTGCTA*CCCCCGCCCC*	144 108
TerG	GTCAAGGATGTTGTAACTAACCA*GGGGGCGGGG* CAGTTCCTACAACATTGATTGGT*CCCCCGCCCC*	109 110	GATGTTGTAACTAACCA*GGGGGCGGGG* CAGTTCCTACAACATTGATTGGT*CCCCCGCCCC*	145 110
TerH	CGATCGTATGTTGTAACTATCTC*GGGGGCGGGG* GCTAGCATACAACATTGATAGAG*CCCCCGCCCC*	111 112	TATGTTGTAACTATCTC*GGGGGCGGGG* GCTAGCATACAACATTGATAGAG*CCCCCGCCCC*	146 112
TerI	AACATGGAAGTTGTAACTAACCG*GGGGGGCGGGG* TTGTACCTTCAACATTGATTGGC*CCCCCGCCCC*	113 114	GAAGTTGTAACTAACCG*GGGGGGCGGGG* TTGTACCTTCAACATTGATTGGC*CCCCCGCCCC*	147 114
TerJ	ACGCAGTAAGTTGTAACTAATGC*GGGGGCGGGG* TGCGTCATTCAACATTGATTACG*CCCCCGCCCC*	115 116	TAAGTTGTAACTAATGC*GGGGGCGGGG* TGCGTCATTCAACATTGATTACG*CCCCCGCCCC*	148 116
oriC	CCGGCTTTTAAGATCAACAACCTGGAAAGGATCA GGCCGAAAATTCTAGTTGTTGGACCTTTCCTAGT	117 118	TTTAAGATCAACAACCTGGAAAGGATCA GGCCGAAAATTCTAGTTGTTGGACCTTTCCTAGT	149 118

Italicized sequence corresponds to the 10-mer CG rich region used to elevate oligonucleotides melting temperature.

2.1.4.3   Oligonucleotides used for SPR analysis

For SPR experiments, all *Ter* and *Ter-lock* DNA were designed to include a single-stranded decamer overhang after base 23 to allow their hybridization to a 10 bp-long biotinylated complementary oligonucleotide ("velcro"; Table 6). Individual oligonucleotides were resuspended at 100 µM in TE_{50} (10 mM Tris (pH 7.5), 0.1 mM EDTA and 50 mM KCl) and 25 µl of the oligonucleotide containing the 10-mer overhang was mixed with 50 µl of the complementary oligonucleotides, and 175 µl of SPR_{250} buffer (50 mM Tris (pH 7.7), 0.1 mM EDTA, 0.2 mM β-mercaptoethanol and 250 mM KCl). Hybridization was achieved by heating at 80°C for 2 minutes followed by slow cooling to room temperature.

Table 6: *Ter* and *Ter*-lock oligonucleotides used for SPR.

	Ter	JCU #	*Ter-lock*	JCU #
TerA	AATTAGTATGTTGTAACTAAAGT*GGGGGCGGGG* TTAATCATACAACATTGATTTCA	97 163	TATGTTGTAACTAAAGT*GGGGGCGGGG* TTAATCATACAACATTGATTTCA	140 163
TerB	AATAAGTATGTTGTAACTAAAGT*GGGGGCGGGG* TTATTCATACAACATTGATTTCA	99 164	TATGTTGTAACTAAAGT*GGGGGCGGGG* TTATTCATACAACATTGATTTCA	140 164
TerC	ATATAGGATGTTGTAACTAATAT*GGGGGCGGGG* TATATCCTACAACATTGATTATA	101 165	GATGTTGTAACTAATATG*GGGGCGGGG* TATATCCTACAACATTGATTATA	141 165
TerD	CATTAGTATGTTGTAACTAAATG*GGGGGCGGGG* GTAATCATACAACATTGATTTAC	103 166	TATGTTGTAACTAAATG*GGGGGCGGGG* GTAATCATACAACATTGATTTAC	142 166
TerE	TTAAAGTATGTTGTAACTAAGCA*GGGGGCGGGG* AATTTCATACAACATTGATTCGT	105 167	TATGTTGTAACTAAGCAG*GGGGCGGGG* AATTTCATACAACATTGATTCGT	143 167
TerF	CCTTCGTATGTTGTAACGACGAT*GGGGGCGGGG* GGAAGCATACAACATTGCTGCTA	107 168	TATGTTGTAACGACGATG*GGGGCGGGG* GGAAGCATACAACATTGCTGCTA	144 164
TerG	GTCAAGGATGTTGTAACTAACCA*GGGGGCGGGG* CAGTTCCTACAACATTGATTGGT	109 169	GATGTTGTAACTAACCAG*GGGGCGGGG* CAGTTCCTACAACATTGATTGGT	110 145
TerH	CGATCGTATGTTGTAACTATCTC*GGGGGCGGGG* GCTAGCATACAACATTGATAGAG	111 170	TATGTTGTAACTATCTC*GGGGGCGGGG* GCTAGCATACAACATTGATAGAG	112 146
TerI	AACATGGAAGTTGTAACTAACCG*GGGGGCGGGG* TTGTACCTTCAACATTGATTGGC	113 171	GAAGTTGTAACTAACCG*GGGGGCGGGG* TTGTACCTTCAACATTGATTGGC	114 147
TerJ	ACGCAGTAAGTTGTAACTAATGC*GGGGGCGGGG* TGCGTCATTCAACATTGATTACG	115 172	TAAGTTGTAACTAATGC*GGGGGCGGGG* TGCGTCATTCAACATTGATTACG	116 148
Velcro	Biotin-CCCCGCCCCC			173

The italicized sequence is the complementary strand of the velcro used to immobilize *Ter* and *Ter*-lock sites on NLC chips.

2.2 Methods

2.2.1 Construction of plasmids

2.2.1.1   Preparation of plasmid DNA

E. coli DH12S was used routinely as host strain during construction of plasmids. Fresh overnight cultures of DH12S cells carrying the plasmid of interest were harvested by centrifugation. Plasmid DNA was extracted and purified from DH12S cells using the Nucleospin plasmid kit (Macherey-Nagel) and processed according to manufacturer's instructions. Plasmid DNA was eluted from the spin column in 30 µl of elution buffer.

2.2.1.2   Preparation of genomic DNA

Genomic DNA was obtained from *E. coli* DH12S cells grown in 2 ml LB broth overnight at 37°C and harvested by centrifugation. Genomic DNA was directly extracted using the Wizard genomic DNA purification kit (Promega) according to manufacturer's instructions. The DNA was recovered from isopropanol and ethanol precipitation by adding 100 µl of TE (10 mM Tris (pH 7.5), 0.1 mM EDTA) or TBS (20 mM Tris pH 7.5, 150 mM NaCl) and incubating the mixture at 65°C for 1 hour. The DNA quality was verified by agarose gel electrophoresis and quantified using Nanodrop.

2.2.1.3   Polymerase chain reaction

The polymerase chain reaction (PCR) was used to amplify *E.coli* genes from strain DH12S genomic DNA for subsequent cloning or to amplify *Ter-* and *oriC*-containing regions for the qPCR-DNA binding assay (*cf* Table 4 and Table 7 for primer sequences). A typical PCR reaction (50 µl) contained 1 ng of genomic DNA diluted in autoclaved $_{dd}H_2O$, 1X iProof High-Fidelity DNA polymerase buffer, 2 units of iProof High Fidelity DNA polymerase (Bio-Rad), 200 µM dNTP and specific primers at 300 µM each. The protocol consisted of a denaturation step of 30 s at 98°C, followed by 30 cycles at 98°C, 10 s, 60°C, 10 s, and 72°C, 10 s (according to amplicon size), and a final extension step at 72 °C for 5 minutes. Product size was verified by agarose gel electrophoresis.

2.2.1.4   Agarose gel electrophoresis

DNA fragments were routinely separated using a 1 % agarose gel. Agarose I gel/TBE blend 2 % (Amresco) powder was dissolved in 0.5X TBE. Gels contained Gel Red (1X) to enable visualisation of DNA under UV light. DNA samples were in 1X DNA loading dye (50 mM EDTA, 10 % glycerol, bromophenole blue). One microgram of either 1 kbp DNA ladder

(NEB) or 100 bp DNA ladder (NEB) was loaded on the gels according to the size of the DNA fragment to be analyzed. Gels were run in TBE buffer (90 mM Tris, 90 mM boric acid, 3 mM EDTA) at a constant voltage of 120 V using the PowerPac (Bio-Rad). DNA was visualised using a 302 nm transiluminator or a molecular Imager Bio-Rad Gel Doc^{TM} XR System.

2.2.1.5   Cloning

The *dnaA* and *ssb* genes were obtained by PCR amplification of DH12S genomic DNA (*cf* section 2.2.1.2) using primers designed to incorporate enzyme restriction sites for subsequent cloning into either pET-GFP (pIM013) or PelB-Tus (pIM033; Table 7). The pET-GFP cloning cassette encoding an N-terminal His$_6$-tag and a C-terminal GFP-tag was used to clone *dnaA-gfp* and *ssb-gfp* (Moreau et al., 2010, Morin et al., 2012). The alternative vector pIM033 was used to clone *dnaA* and *ssb* under the control of the T7 promoter (Table 7 and Table 8).

Table 7: Primer sequences for PCR amplification of genes of interest and restriction sites.

Protein of interest	Forward primer (5'-3')	Restriction site	Reverse primer (5'-3')	Restriction site	JCU #
SSB	AAAAAACATATGGCC AGCAGAGGCGTAAAC AAGG	NdeI	AAAAAAGAATTCTTAGAAC GGAATGTCATCATCAAAGT CCATC	EcoRI	235/236
SSB-GFP	AAAAAAGATATCGCC AGCAGAGGCGTAAAC AAGG	EcoRV	AAAAAAGCTAGCGAACGGA ATGTCATCATCAAAGTCCA TC	NheI	218/219
DnaA	AAAAAACATATGTCA CTTTCGCTTTGGCAG CAGTGTC	NdeI	AAAAAAGTCGACTTACGAT GACAATGTTCTGATTAAAT TTG	SalI	237/238
DnaA-GFP	AAAAAACTTAAGTCA CTTTCGCTTTGGCAG CAGTGTC	AflII	AAAAAAGCTAGCCGATGAC AATGTTCTGATTAAATTTG	NheI	275/276

Underlined sequences correspond to enzyme restriction sites.

PCR products and vectors were separately digested with the restriction enzymes specified in Table 7 in the reaction buffers recommended by the manufacturer (New England Biolabs).

Double digests were carried out in compatible buffers for both restriction enzymes or sequentially allowing for buffer compensation. Generally, 50 µl digest reaction mix contained 1 µg of purified DNA (insert or vector), 1X enzyme specific buffer, 2 units of restriction enzyme and BSA when required. Digestion controls were carried out in 20 µl reaction mix and all samples were incubated for 1 hour at the recommended temperature. Digestion reactions were loaded into a 1 % agarose gel in TBE at 120 V. Digestion of pIM33 with NdeI/SalI or with NdeI/EcoRI excised the pelB-Tus sequence. DNA fragments of interest were excised from agarose gel using a sterile scalpel blade under a 302 nm transilluminator and extracted from the agarose using the AxyPrep DNA gel extraction kit (Axygen Biosciences) following manufacturer's instructions. The DNA fragments were eluted twice in 30 µl elution buffer.

A fraction of purified insert and vector was analyzed on a 1 % agarose gel to determine the amounts to use for ligation. Two ligation mixtures (10 µl) were prepared containing either 1:1 or 1:3 insert to vector ratio and 400 units of T4 DNA ligase (NEB). A control ligation was set up containing only the digested vector and T4 DNA ligase. All ligation mixtures were incubated at 16°C overnight. The following day, 50 µl of chemically competent *E. coli* DH12S cells were transformed by heat shock with the 10 µl of ligation mixture as described in section 2.2.2. Positive transformants were selected on agar plates supplemented with ampicillin (100 µg/ml). Several transformants were transferred to a small starter culture grown at 37°C to late log phase to purify plasmid DNA as described in section 2.2.1.1. Correct DNA insertion was then verified by enzyme digestion on the freshly isolated plasmid using the appropriate restriction enzymes (i.e. cutting in the insert). Alternatively, transformants were screened by colony PCR as follow: a single colony was transferred with a sterile plastic loop on a new LB agar plate containing antibiotics (master plate) and to a PCR tube. A PCR master mix was prepared for the number of colonies screened containing 0.5

µM of each specific primer, 1X Taq DNA polymerase buffer, 200 µM dNTP mix (total) and

2.5 units of Taq polymerase (10 µl per reaction). The PCR protocol consisted of a

denaturation step of 30 s at 95°C, followed by 30 cycles of 10 s at 95°C, 10 s at 60°C, 10 s

(depending on amplicon size) at 68°C and a final extension step of 5 minutes at 68°C. Band

size was verified by agarose gel electrophoresis. Verified plasmids (Table 8) were sequenced

at the Australian Genome Research Facility (AGRF). Sequencing reactions contained 10 µl

of purified plasmid and vector primer(s) at 0.7 µM (final concentration).

Table 8: Recombinant plasmids.

Plasmid Name	Protein Encoded	Parent plasmid	Reference
pMM084	His_6-SSB-GFP	pIM013	Moreau et al., 2012
pMM085	SSB	pIM033	Unpublished
pMM200	DnaA	pIM033	Unpublished
pMM220	His_6-DnaA-GFP	pIM013	Moreau et al., 2012

2.2.2 Bacteria transformation

2.2.2.1   Chemical competence

E. coli DH12S, BL21(*DE3*)RIPL or KRX cells were grown in 10 ml LB broth supplemented

with the appropriate antibiotics at 37°C until OD_{600} reached between 0.4 and 0.8. Cell

cultures were transferred to ice and centrifuged in 1 ml aliquots at 4°C for 30 s at 6,000 *g*

(Microfuge 22R centrifuge, Beckman Coulter). Cells were washed once in 1 ml ice cold

CCMB (80 mM $CaCl_2$, 20 mM $MnCl_2$ 10 mM $MgCl_2$, 10 mM K-acetate, 10 % glycerol

(v/v)). Following a centrifugation step, cells were resuspended in 1 ml ice cold CCMB and 50

µl aliquots were snap-frozen in liquid nitrogen. Competent cells were stored at -80°C.

2.2.2.2　　Heat shock transformation

Competent cells (50 µl) were thawed on ice and 1 µl of purified plasmid DNA or 10 µl of ligation mixture was added and incubated on ice for 30 minutes. Cells were heat-shocked at 42°C for 30 s for KRX cells and 1 minute for DH12S or BL21(*DE3*)RIPL cells. Cells were placed back on ice for two minutes before adding 450 µl of LB broth. Cells were grown for 1 hour at 37°C and 50 µl were streaked on LB agar plates containing the selective antibiotic(s). The plates were incubated overnight at 37°C.

2.2.3 Large scale protein expression and purification

E. coli BL21(*DE3*)RIPL was used for overexpression of His_6-tagged Tus, Tus-GFP and GFP and strain KRX was used to express DnaA, SSB and their GFP-tagged homologue (*cf* Table 2 for strain genotypes).

2.2.3.1　　Protein overexpression

A fresh overnight culture of BL21(*DE3*)RIPL carrying either His_6-Tus, His_6-Tus-GFP or His_6-GFP encoding plasmids was used to inoculate 100 ml of Overnight Express Instant TB Medium (Novagen) supplemented with ampicillin (100 µg/ml) and chloramphenicol (50 µg/ml). Cells were grown at 37°C until OD_{600} 0.5 and then shaken at 16°C for 60 hours.

KRX cells transformed with pMM200 (DnaA) or pMM220 (DnaA-GFP) were streaked onto LB agar plates supplemented with 100 µg/ml ampicillin and 0.4 % glucose and incubated overnight at 37°C. Single colonies were then streaked on a master plate and grown for a further 6 hours at 37°C. A loop of bacteria from the master plate was used to inoculate 100 ml of LB broth supplemented with ampicillin (100 µg/ml) and incubated overnight at 25°C (230 rpm). For DnaA overexpression, the temperature was slowly increased to 37°C and protein expression was induced at OD_{600} 1.8 with 0.1 % rhamnose (final concentration)

for 3 hours. DnaA-GFP expression was induced at 16°C for a further 24 hours to increase the proportion of folded and soluble proteins.

Starter cultures of KRX cells carrying the plasmid encoding SSB or SSB-GFP were used to innoclutate 100 ml of terrific broth supplemented with 100 µg/ml ampicillin. The culture was incubated at 37°C (230 rpm) until OD_{600} 0.8-1 after which the temperature was decreased to 16°C or 25°C for SSB-GFP and SSB respectively and protein expression was induced with 0.1 % rhamnose (final concentration) for 24 hours. All cell cultures were harvested by centrifugation at 4°C for 30 minutes at 800 g (Eppendorf centrifuge 5810R). Cell pellets were snap-frozen in liquid nitrogen and stored at -80°C until purification.

2.2.3.2   Purification of replisomal proteins

Cells from 100 ml of culture (~3 g) were thawed on ice and resuspended in their respective lysis buffer at 7 ml/g. DnaA expressing cells were resuspended in DnaA buffer (50 mM HEPES-KOH (pH 7.6), 1 mM EDTA, 1 mM β-mercaptoethanol, 20 % sucrose (w/v)) supplemented with 625 mM KCl and 20 mM spermidine. DnaA-GFP expressing cells were resuspended in DnaA-GFP buffer (DnaA buffer supplemented with 200 mM KCl and 10 mM $MgCl_2$) containing 20 mM spermidine. The presence of $MgCl_2$ in DnaA-GFP buffer helped solubilizing the recombinant proteins. SSB lysis buffer (50 mM Tris (pH 7.8), 1 mM Na_3EDTA, 0.2 M NaCl, 10 % sucrose (w/v) and 15 mM spermidine) was used to resuspend SSB and SSB-GFP expressing cells. All cell suspensions were lysed by two passages in a French press at 12,000 psi and the soluble lysate was collected by centrifugation at 39,000 g for 30 minutes at 4°C (Beckman Coulter Avanti J-20 XP centrifuge). All following steps were carried out at 4 °C.

Tus, Tus-GFP and GFP proteins were affinity purified using Profinity IMAC nickel-charged resin (Bio-Rad) as described in Dahdah et al. (2009). Proteins were eluted with 200

mM imidazole and precipitated by adding ammonium sulfate (0.5 g/ml of lysate) followed by a centrifugation step at 30,000 g for 30 minutes (Microfuge 22R, Beckman-Coulter). Protein pellets were resuspended in buffer A (45 mM Na_2HPO_4, 5 mM NaH_2PO_4 (pH 7.8), 2 mM β-mercaptoethanol and 10 % glycerol).

The DnaA proteins were precipitated from the lysate by addition of ammonium sulphate (0.20 g/ml) and collected by centrigutation at 30,000 g for 30 minutes (Microfuge 22R, Beckman-Coulter). The protein pellet was dissolved in 10 times the initial lysate volume in DnaA-MN buffer (DnaA buffer supplemented with 10 mM $MgCl_2$ and 50 mM NaCl) to dilute residual ammonium sulfate and KCl and allow binding of proteins to Macro-Prep High S support resin (Bio-Rad). Equilibrated resin (3 ml for 270 ml of 10 times diluted lysate) was incubated with the lysate with gentle agitation for 1 hour before transfer into a glass column (Econo-column, 2.5 x 20 cm Bio-Rad). The column was washed 5 times with 3 ml of DnaA-MN buffer and the protein was eluted with 3 times 3 ml of DnaA elution buffer (DnaA buffer supplemented with 10 mM $MgCl_2$ and 1 M NaCl). Protein fractions were concentrated by ammonium sulfate precipitation (0.2 g/ml) followed by centrifugation at 30,000 g for 30 minutes. Protein pellets were re-solubilized in DnaA buffer.

DnaA-GFP lysates were passed twice through a glass column (econo-column, 2.5 x 20 cm Bio-Rad) packed with 3 ml of Profinity IMAC nickel-charged resin (Bio-Rad). Contaminant proteins were washed with 3 times 3 ml of DnaA-GFP wash buffer (DnaA buffer supplemented with 10 mM imidazole) and the proteins were eluted from the column with 3 ml DnaA-GFP elution buffer (DnaA buffer supplemented with 200 mM imidazole). Proteins were precipitated with ammonium sulphate (0.5 g/ml). After 30 minutes centrifugation at 30,000 g, protein pellets were resuspended in DnaA buffer.

To purify SSB, polymin P was added drop by drop to the lysate (0.4 %) to precipitate all DNA bound proteins. After centrifugation at 30,000 g for 30 minutes, the pellet containing

SSB proteins was resuspended in TGE buffer (50 mM Tris (pH 8.3), 1 mM Na$_3$EDTA, 20 %

glycerol (v/v), 0.4 M NaCl). Insoluble material was centrifuged for 20 minutes at 6,000 *g* and

soluble SSB was precipitated by adding ammonium sulfate (0.15 g/ml). Protein pellets

obtained after centrifugation at 30,000 *g* for 30 minutes were dissolved in SSB resuspension

buffer (20 mM Tris (pH 8), 0.5 M NaCl, 1 mM Na$_3$EDTA, 10 % glycerol (v/v), 1 mM β-

mercaptoethanol). The SSB-GFP proteins were directly extracted from the lysate by

ammonium sulphate precipitation at 0.2 g/ml and were resuspended in 50 mM phosphate

after centrifugation (pH 8). All proteins samples were snap-frozen in liquid nitrogen and

stored at -80°C.

2.2.4 Protein separation and analysis

2.2.4.1 SDS-PAGE analysis of proteins

Protein samples were separated and analyzed on 0.75 mm SDS-PAGE gels using the Mini-

Protean 3 cell (Bio-Rad) gel apparatus. Gels were prepared by mixing 4.2 ml of 12.5 % Next

gel solution (Amresco) with 30 µl of 10 % ammonium persulfate (APS) and 3 µl of TEMED.

The solution was poured into the gel cast plates and left to polymerize at room temperature

for 30 minutes. Protein separation was carried out with Next gel running buffer (Amresco).

Alternatively, a 10 % SDS-PAGE was prepared with acryl/bis stacking and resolving gel

(Amresco). For a 0.75 mm gel, 3.5 ml of resolving gel solution (10 % Acryl/bis 29:1, 40%

(w/v) solution (Amresco), 250 mM Tris (pH 8.8), 0.1 % SDS, 0.1 % APS and 0.1 %

TEMED) was poured and sealed with $_d$H$_2$O to allow polymerization. The water layer was

removed and 1.5 ml of stacking gel (5.3 % Acryl/bis 29:1, 40% (w/v) solution (Amresco),

125 mM Tris (pH 6.8), 0.1 % SDS, 0.1 % APS and 0.1 % TEMED) was poured over the

resolving gel. The gel was set for 30 minutes at room temperature and run with Tris-Glycine

buffer (25 mM Tris, 200 mM glycine, 0.1 % SDS). Proteins samples were mixed with 2X

sample loading buffer (125 mM Tris (pH 6.8), 20 % glycerol. 4 % SDS, 0.005 % bromophenol blue) and heat denatured for 5 minutes at 95°C. For protein induction gels, aliquots of cell cultures taken before and after induction were centrifuged and the cell pellets were resuspended in a given volume of 2X sample loading buffer to obtain a cell suspension concentration equivalent to an OD_{600} of 10. After heat denaturation (5 minutes at 95°C) and vortexing, 10 µl of cell suspension was loaded onto the gel and separated at 120 V (NEXT GEL) or 150 V (resolving and stacking gel). Proteins bands were stained with a Coomassie staining solution (50 % Methanol, 10 % acetic acid, 0.05 % Brilliant blue R) and the background was destained with a solution of 10 % acetic acid and 40 % propan-2-ol until clear.

2.2.4.2 Bradford assay

The concentration of proteins was measured by Bradford assay using the Bradford reagent (Sigma). The standard was established using 0.5, 1, 1.5, 2 and 3 µl of BSA stock solution at 10 mg/ml into 750 µl of Bradford reagent. After 30 minutes at room temperature, OD_{595} was measured with a SmartSpec 3000 (Bio-Rad) and the concentration of the unknown sample was determined according to the linear regression obtained with the BSA standard in Microsoft Excel.

2.2.4.3 Fluorescence assay

The concentration of Tus-GFP was determined using a Tus-GFP sample of known concentration as standard. Tus-GFP standard was serially diluted in a black 96-well plate (Nunc) and the fluorescence was measured using the fluorescence plate reader Victor V (Wallace Perkin-Elmer) with 355/535 nm excitation and emission filters (+/- 40 nm). The

concentration of the unknown sample was determined according to the linear regression performed on the Tus-GFP standard in Microsoft Excel.

2.2.4.4   Western blot analysis

SDS-PAGE gels (*cf* section 2.2.4.1) were directly transferred into transfer buffer (12 mM Tris, 9.7 mM Glycine, 0.037 % SDS, 25 % methanol) after electophoresis. Immuno-blot PVDF membrane (Bio-Rad) was activated for 10 s in 100 % methanol prior to immersion in transfer buffer with Xtra thick blot papers (Bio-Rad). Pads, membrane and SDS-PAGE were assembled in a Trans-blot SD semi-dry transfer cell (Bio-Rad) and subjected to electrophoresis for 20 minutes at 15 V. The PVDF membrane was then blocked with PBST (10 mM sodium phosphate (pH 7.4), 150 mM NaCl, 0.05 % Tween) containing 5 % skim milk powder (50 ml per gel) for 1 hour at room temperature. Antibodies (primary and secondary) were diluted in PBST with 1 % skim milk (10 ml per gel) and incubated with the PVDF membrane in a sealed plastic bag for 1 hour at room temperature under gentle agitation. The PVDF membrane was washed 3 times 5 minutes in PBST between primary and secondary antibody incubations and between secondary antibody and revelation. Fast 3,3'-diaminobenzidine (DAB) tablets set (Sigma) were dissolved in 5 ml $_{dd}H_2O$ in the dark and applied to the washed PVDF membrane to reveal bound HRP-conjugated secondary antibody. The membrane was covered with aluminum foil for 5 minutes and DAB was washed off the membrane with PBST. Membranes were let to dry overnight in the dark and were scanned the following day.

2.2.5 Production and purification of polyclonal antibodies

2.2.5.1   Immunization of chickens (*Gallus gallus*) and collection of anti-sera

Purified Tus, SSB and DnaA proteins (*cf* section 2.2.3) were diluted to 2 µg/µl in their respective buffer (*cf* section 2.2.3.2) and mixed with an equal volume of Emulsigen adjuvant

(MVP technologies). Each chicken (*Gallus gallus*) received 100 µg of purified proteins (100 µl of proteins in adjuvant) per injection in breast muscles using a 26G needle. The immunization and bleeding schedule was as follows:

Table 9: Chicken immunization schedule.

Immunization	Day	Bleeding
	0	Pre-immune serum collection
Primary	0	
1st boost	14	
2nd boost	28	
	35	Bleeding 1
	42	Bleeding 2
3rd boost	81	
	85	Bleeding 3

The blood was collected from the wing vein using a 23G needle. A small volume (≤ 100 µl) of tri-sodium citrate was present in the needle to slow down the coagulation of chicken blood and enable blood collection. Collected blood was left at room temperature for an hour prior to centrifugation at 370 *g* for 5 minutes at room temperature. The sera (supernatant) were mixed with an equal volume of 100 % glycerol and stored at -20°C until purification.

2.2.5.2   IgY purification

The sera were ammonium sulphate precipitated at 0.2 g/ml for 1 hour on ice and centrifuged at 30,000 *g* for 30 minutes at 4°C. The pellet was resuspended in filtered PBS (10 mM sodium phosphate (pH 7.4), 150 mM NaCl) supplemented with 10 % (bleeding 1 and 2) or 50 % (bleeding 3) glycerol.

2.2.6 GFP-Basta

The development and validation of GFP-Basta is described in Chapter 3. The detailed method and buffers used are described in section 3.2.4, p65. Protein samples were denatured in low-profile 0.2 ml PCR tubes with caps (Bio-Rad) in a thermocycler (MyCycler, Bio-Rad). The protein florescence was measured in black 96-well plates (Nunc) with a fluorescence plate reader (Victor V Wallace Perkin-Elmer).

2.2.7 DSF-GTP assay

The differential scanning fluorescence of GFP-tagged proteins (DSF-GTP) was developed and validated in Chapter 5. The detailed methods are described in section 5.2.2, p110. All the experiments were performed with the IQ5 iCycler (Bio-Rad). The automatic peak recognition program was run with the RStudio interface.

2.2.8 qPCR DNA-binding assay

The qPCR DNA-binding assay was developed in Chapter 6 and the detailed procedures can be found in section 6.2.2, p126. The streptavidin plates (Thermoscientific, Reacti-BindTM Streptavidin coatded HBC black 96-well plates), the biotinylated goat anti-GFP antibody (Ab 6658; Abcam) and the IQ5 iCycler (Bio-Rad) were used for the qPCR DNA-binding assay.

2.2.9 SPR

The ProteON XPR36 and NLC chips (Bio-Rad) was used to determine the kinetics of Tus binding to *Ter* sites at 20°C. The detailed procedure is included in Chapter 4 section 4.2.3, p83. The max RU value was checked for each measurement, outliers were removed and the average of all k_a and k_d measurements were used to determine the K_D (k_d/k_a).

2.2.10 Chromosome immunoprecitation (ChIP)

ChIP was performed in a microplate (96-well Maxisorb) coated with goat anti-GFP IgG (Abcam; Ab6673) to capture GFP-tagged proteins of interest from exponentially growing *E. coli* cells. The method is detailed in Chapter 7 section 7.2.3, p139.

Chapter 3: Quantitative determination of protein stability and ligand binding using a green fluorescent protein reporter system

This chapter has been published in Molecular BioSystems[1] (Moreau, M. J., Morin, I. &

Schaeffer, P. M) and describes the validation of a new method developed to characterise Tus-

Ter complexes in high-throughput. In this article, the technique was validated using Tus and

two other proteins, namely the glycerol kinase (GK) and the chloramphenicol

acetyltransferase (CAT) from E. coli. Only the data obtained for the Tus protein are included

in this chapter as the GK and CAT cloning and purifications were generated dy Dr I. Morin.

3.1 Introduction

The stable conformation of proteins or native state is obtained by interactions between amino

acids forming secondary structure motifs (α-helices β-sheets or random coils) held together

by hydrogen bonds, as well as hydrophobic and ionic interactions to form the tertiary folded

structure. These intra-molecular contacts drive the folding pathway of proteins to a final

native structure corresponding to the lowest Gibbs free energy conformation (Anfinsen, 1973,

---

[1] This chapter contains data published in Moreau, M. J., Morin, I. & Schaeffer, P. M. 2010. Quantitative determination of protein stability and ligand binding using a green fluorescent protein reporter system. Mol Biosyst, 6, 1285-92. [2] Moreau, M. J. and Schaeffer, P. M. 2012. Differential Tus-*Ter* binding and lock formation: implications for DNA replication termination in *Escherichia coli*. *Mol Biosyst,* 8**,** 2783-91.

Clark, 2004, Dobson et al., 1998). The measure of protein stability is therefore a measure of the energy engaged in all these interactions. Upon ligand binding, the newly formed inter-molecular interactions add to the native energy state and contribute to the total stability of the protein (Brandts and Lin, 1990, Schellman, 1975, Straume and Freire, 1992).

Currently, methods such as differential scanning calorimetry (DSC; Jelesarov and Bosshard, 1999) isothermal denaturation (ITD; Senisterra et al., 2008), differential scanning fluorimetry (DSF; Ericsson et al., 2006, Niesen et al., 2007, Pantoliano et al., 2001, Senisterra and Finerty, 2009, Vedadi et al., 2006) and light scattering (Leung et al., 1996, Senisterra et al., 2006, Vedadi et al., 2006) can be used to measure the increase in thermal stability of a protein upon binding to a ligand, drug, or inhibitor. The ligand-dependent change in T_m (ΔT_m) is directly proportional to the concentration and to the binding affinity of the ligand (Jelesarov and Bosshard, 1999, Lo et al., 2004, Matulis et al., 2005). These methods determine the stability of a protein by measuring the fraction of folded or unfolded/aggregated protein as a function of time or temperature. DSF is one of the most promising technologies used for high-throughput (HT) characterization of protein stability and ligand binding, as it is adaptable to any soluble protein even without known function. Unfortunately, neither DSF nor the remainder of thermal denaturation based methods can be used with partially purified or mixtures of proteins, as they cannot identify which of the proteins in the mixture is unfolding.

In order to compare the affinity of Tus for all ten *Ter* sites and circumvent the above mentioned issues, I developed a fast and simple *in vitro* method using the green fluorescent protein (GFP). GFP was used as a reporter system to quantify the stability of a partially purified protein and its ligand-associated stabilization. GFP is a very stable protein that has previously been fused to proteins of interest (POIs) to monitor *in vivo* protein folding and solubility (Waldo, 2003, Waldo et al., 1999). A similar *in vivo* assay was also designed for

the screening of Aβ peptide aggregation inhibitors (Kim et al., 2006). Here, the GFP-based protein stability assay (GFP-Basta) takes advantage of the fact that most proteins, when subjected to thermal denaturation, follow an unfolding pathway leading to irreversible aggregation as illustrated by the reaction coordinate diagram (Figure 9A; Chi et al., 2003). The hypothesis was that if GFP was to be used as a reporter of protein unfolding and aggregation, then the unfolding of the POI and GFP domains in the fusion protein should be totally uncoupled (independent unfolding) to avoid influencing each-other's unfolding kinetics (Figure **9**B). Also, if the aggregation process has been completed, then the measurement of the residual population of folded proteins (non-aggregated) could simply be determined by measuring the fraction of protein that remains soluble (F_{fold}) after heat treatment. Consequently, the thermal stability of the POI could directly be obtained through the measurement of the fluorescent F_{fold} of the GFP fusion protein after heat denaturation. In this simple case, the apparent aggregation rate constant k_{agg} reflects the unfolding kinetics of the POI as the rate-limiting step is the unfolding process. As a result, the full range of physical and chemical conditions where GFP is stable and fluorescent can be used to monitor the aggregation properties of a less stable POI (Figure 9B).

This chapter describes the validation of a new thermal stability assay capable of quantitatively determining the thermal stability of a POI in the presence of other proteins and rank protein ligands according to their K_D. It requires neither special equipment nor extensive purification steps.
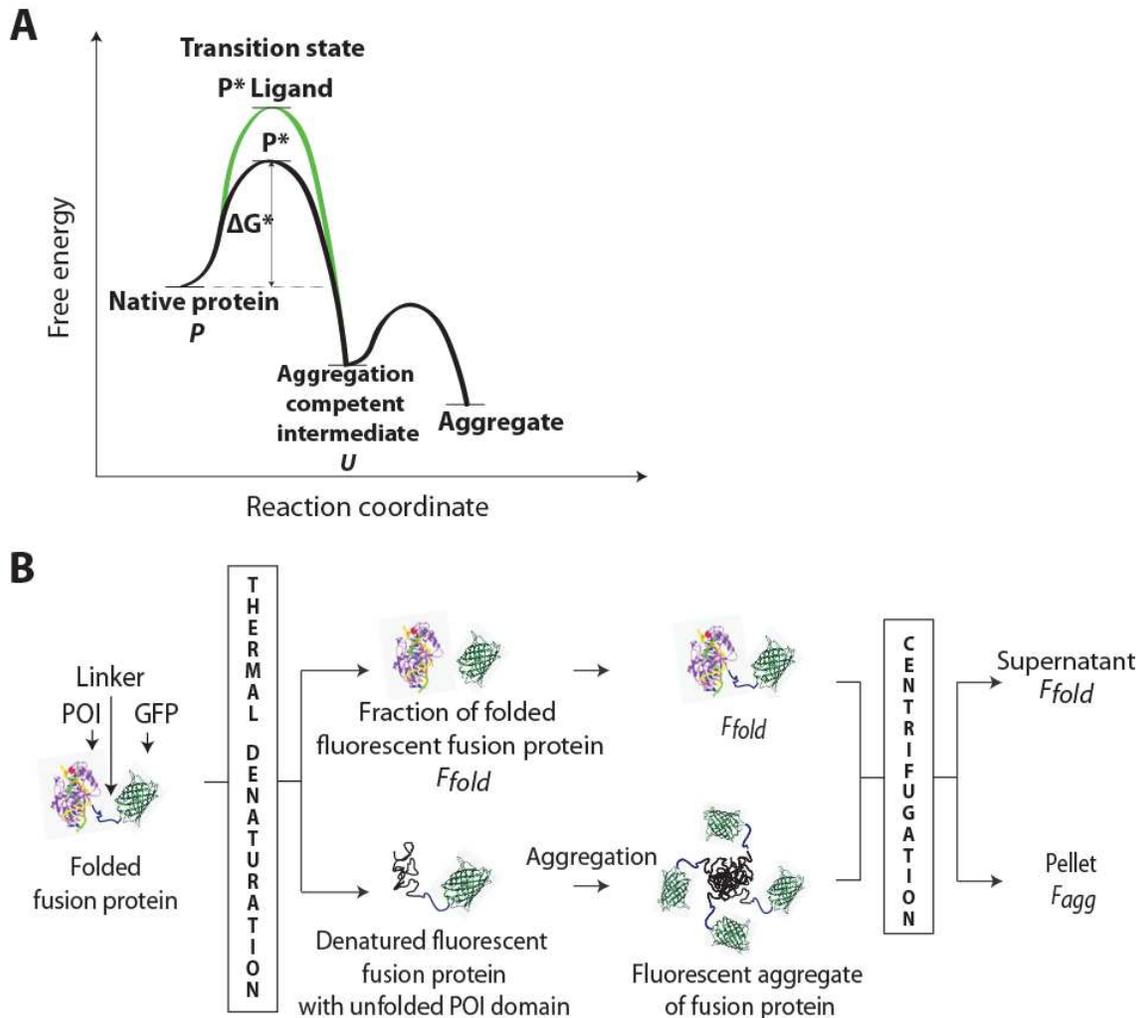
Figure 9: Principle of GFP-Basta. (A) Reaction coordinates of irreversible protein aggregation. ΔG* is the change in free energy of activation (Chi et al., 2003). (B) Principle of GFP-Basta. The thermal denaturation of POI-GFP fusion proteins produces a heterogeneous population of folded and denatured fluorescent proteins. The fraction of denatured proteins forms aggregates, which are further discarded from the solution to allow the measurement of the soluble fraction F_{fold}.

Three well characterized proteins, the monomeric DNA-binding protein Tus (Kamada et al., 1996), the trimeric chloramphenicol acetyl transferase (CAT; Panchenko et al., 2006) and the tetrameric glycerol kinase (GK; Koga et al., 2008, Thorner and Paulus, 1973) were used to validate GFP-Basta, but only the results obtained for Tus and its ligands are reported in this chapter (*cf* Moreau et al., 2010 for additional validation data with GK and CAT).

3.2 Material and methods

3.2.1 Protein expression and purification

The plasmid encoding His_6-Tus, His_6-Tus-GFP, and His_6-GFP were previously described (Dahdah et al., 2009, Ozawa et al., 2005). The three proteins were expressed and purified as described in Dahdah et al. (2009). After ammonium sulfate precipitation the proteins were resuspended in buffer A (45 mM Na_2HPO_4, 5 mM NaH_2PO_4 (pH 7.8), 10 % glycerol (v/v), 2 mM β-mercaptoethanol). The purity of proteins was assessed by SDS-PAGE (NEXT-GEL Amresco, *cf* section 2.2.4.1 p54) and band quantification using the image analysis software ImageJ (http://rsbwed.nih.gov/ij/).

3.2.2 DNA ligands

Ter oligonucleotides were obtained from Sigma and diluted in TE_{50} (10 mM Tris (pH 7.5), 0.1 mM EDTA, 50 mM KCl). DNA ligands were prepared by heating at 80°C for 2 minutes followed by slow cooling of complementary pairs of oligonucleotides. These DNA ligands correspond to previously described sequences (underlined) with the exception that they have each been extended with a GC rich dsDNA region in order to obtain T_m values above 70°C for each of them. Sequence details along with their respective K_D values (Mulcair et al., 2006) are listed in Table 10.

3.2.3 Determination of thermal aggregation profiles and T_{agg}

Samples (6 or 10 µl) of Tus, Tus-GFP and GFP (10 µM each final concentration) were mixed and incubated in a thermocycler (MyCycler, Biorad) set on algorithm measurement for 15 µl sample volume for 5 minutes along a temperature gradient from 38 to 53.5°C. Protein concentrations were typically between 10-13 µM in buffer A (*cf* section 3.2.1). After

incubation, reactions were stopped by transferring the samples to ice for 10 minutes prior to centrifugation at 18,000 g. for 20 minutes at 4°C in a Beckman Coulter centrifuge (rotor: F12x8.2). The supernatants (3 or 5 µl) were then analyzed by SDS-PAGE (10 % NEXT-GEL Amresco). The gels were illuminated on a transilluminator at 365 nm followed by Coomassie blue staining. Coomassie-stained protein bands corresponding to F_{fold} were integrated using ImageJ (http://rsbwed.nih.gov/ij/), normalized against the fluorescence of an untreated but centrifuged sample and plotted against the temperature. To determine the T_{agg} at which 50 % of proteins were aggregated, the thermal aggregation profile data were fit to the following sigmoid function:

$$F_{fold} = 1 - \left(\frac{1}{1 + e^{\frac{T_{agg}-T}{c}}} \right)$$

where F_{fold} is the normalized fluorescence intensity at temperature T, and c is the Hill slope factor. In the presence of *TerB*, the change in aggregation transition temperature ΔT_{agg} could be calculated as follows:

$$\Delta T_{agg} = T_{agg(Tus-GFP-Ter)} - T_{agg(Tus)}$$

3.2.4 S Method

The S method refers to the use of a spectrofluorometer for fluorescence measurement. Protein samples (6-10 µl) were incubated along a temperature gradient in a thermocycler for 5 minutes for the determination of T_{agg} or in isothermal conditions and increasing times to determine k_{agg}. After heat treatment and centrifugation as described above, half the volume of the supernatant was transferred to a black 96-well plate (Nunclon), diluted with buffer A and the fluorescent F_{fold} was determined with a fluorescence plate reader (Victor V Wallace Perkin-Elmer). The excitation and emission filters were set at 355 nm and 535 nm respectively, with 40 nm band-width. Data were normalized against the fluorescence of an

untreated but centrifuged sample. To evaluate the effect of additives, Tus-GFP (13 µM) or

GFP (control, 12µM) in buffer A were mixed with equal volumes of different additives in

water. To determine the effect of DNA ligands, reaction samples containing 5.4 µl of Tus-

GFP (11 µM in buffer A supplemented with 272.2 mM KCl) and 0.6 µl of DNA ligand (100

µM in TE$_{50}$, pH 8) or TE$_{50}$ (pH 8) for the control were incubated 10 minutes at 25°C to allow

complex formation prior to the heat denaturation step. Reaction volume was 10 µl and 5 µl of

the soluble fraction was analyzed by plate reader after centrifugation. The k_{agg} (s^{-1}) measured

the loss of fluorescence of the soluble fraction of proteins over time. The k_{agg} values were

determined by the exponential fit of normalized F_{fold} as follows:

$$F_{fold} = e^{(-k_{agg})t}$$

where t is the time in seconds.

3.2.5 EMSA

A modified version of an electrophoretic mobility shift assay (EMSA) was used to determine

the T_{agg} of Tus-GFP-*TerB* complex. Briefly, equal volumes of Tus-GFP (~70 µM in buffer A)

and *TerB* (100 µM in TE$_{50}$) were mixed, diluted 5 times in buffer A and incubated at 25°C for

10 minutes. The samples (8 µl) were heated in a thermocycler at the specified temperatures

for 5 minutes followed by 10 minutes on ice. Treated samples were loaded (4 µl) onto a 1 %

TBE-agarose gel where the F_{fold} of complexes were separated from the fraction of aggregated

proteins (F_{agg}) at 80 V for 20 minutes. Here, the binding of *TerB* to the Tus-GFP induces a

shift in electrophoretic mobility of the complex towards the anode, whereas, unbound Tus-

GFP (i.e. aggregated) stays in the wells. GFP fluorescence was detected at 365 nm and

integrated with ImageJ to determine the F_{fold} of Tus-GFP-*TerB* which showed increased

mobility towards the anode. The T_{agg} of Tus-GFP was obtained in the same buffer with the S

method to determine the ΔT_{agg}.

3.3 Results

3.3.1 Design of the model fusion proteins

The fusion construct consists of an N-terminal His_6-POI domain followed by a minimal LGSGGH linker sequence and a C-terminal GFP. The linker was first used for the construction of a fully functional Tus-GFP fusion protein to develop a TT-lock-based immunoPCR system (Dahdah et al., 2009). Tus binds to 21 bp *TerA-J* sequences (Kamada et al., 1996, Mulcair et al., 2006, Neylon et al., 2005) and the association and dissociation rate constants of complex formation can be altered by mutating the *Ter* sequence, providing an invaluable tool to evaluate the effect of ligand affinity on Tus stability using GFP-Basta. The GFP was chosen due to its high excitability in the UV and its extreme stability in various conditions. The limits of GFP-Basta are therefore dependent on the stability of GFP in the various tested conditions.

3.3.2 Principle and validation of the GFP reporter system

To show that the POI and GFP domains unfold independently, the respective stabilities of Tus, Tus-GFP, and GFP were compared by incubating the proteins for 5 minutes at temperatures ranging from 25 to 53.3°C followed by a cooling and centrifugation step to remove protein aggregates. F_{fold} was then determined by SDS-PAGE in order to measure the residual fraction of wtTus. For this experiment, equal amounts of the three proteins were mixed to avoid variations in buffer composition and protein concentrations. The T_{agg} values (temperature at which 50 % of proteins are aggregated) from the thermal aggregation profiles for Tus and Tus-GFP were 45.4 and 44.2°C respectively (Figure 10B). The same was observed for CAT, GK and their GFP fusions (data not shown; *cf* Moreau et al., 2010).
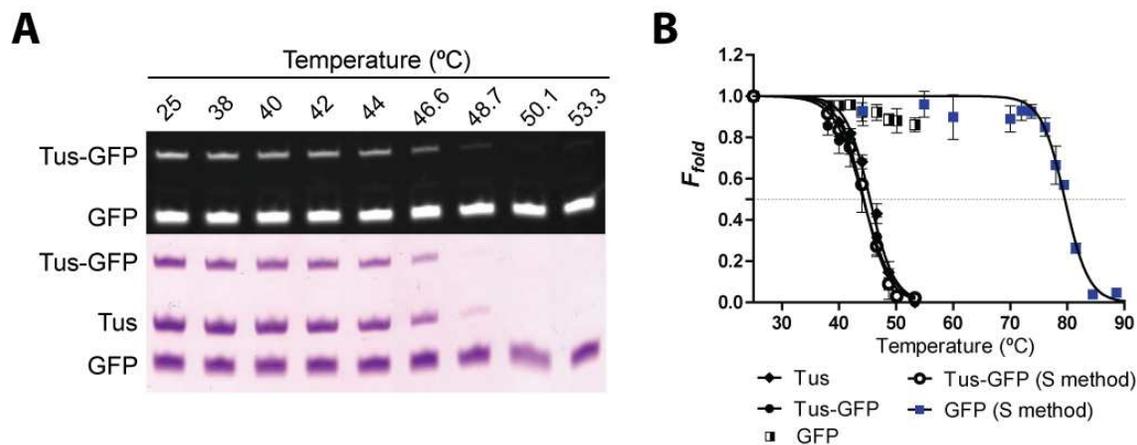
Figure 10: Thermal aggregation profiles of Tus, Tus-GFP and GFP. (A) SDS-PAGE of F_{fold} for equal amounts of Tus, Tus-GFP and GFP (10 µM each) heat treated during 5 minutes at temperatures ranging from 25°C to 53.3°C. Fluorescence (top gel) was recorded with illumination at 365 nm before Coomassie blue staining (bottom gel). **(B)** Thermal aggregation profiles of Tus, Tus-GFP and GFP obtained from triplicate SDS-PAGE, and by the S method (*cf* section 3.2.4) performed in the same conditions. The T_{agg} is the temperature at which 50 % of the proteins are aggregated.

This demonstrated that the unfolding of the POI leading to its aggregation was unaltered in the GFP-fusion protein and that no substantial effect was induced by the GFP domain. The data indicate that the aggregation rate constants (k_{agg}), and therefore all preceding unfolding processes, must be essentially identical for the POI-GFP and POI. As expected, GFP was not affected in this temperature range (Ishii et al., 2007a, Ishii et al., 2007b). The T_{agg} of GFP was determined to be 79.6°C by measuring its residual fluorescence after heat denaturation and centrifugation at a higher temperature range using a fluorescence plate reader as readout (S method). The T_{agg} of Tus-GFP was also reproduced using the S-method (44.3°C, Figure 10B). Furthermore, the total fluorescence including F_{fold} and F_{agg} (fraction of aggregated proteins) remained unaltered if heat denaturation occurred at temperatures below ~75°C for 5 min, meaning that GFP is still folded in the fusion protein aggregates and that aggregation of the POI portion did not trigger the unfolding of GFP. The aggregation of POI-GFP is therefore the result of the unfolding of its most unstable POI domain. These results validate

68

GFP-basta and demonstrate that the minimal LGSGGH linker is long enough to uncouple the unfolding of POI and GFP in the fusion protein.

3.3.3 Isothermal aggregation and evaluation of additives

To evaluate the kinetic parameters of this system (i.e. aggregation rate k_{agg}), a 96-well plate format was designed that enabled the measurement of the residual F_{fold} of Tus-GFP over time. Isothermal aggregation reactions were monitored at a temperature close to the previously determined T_{agg} of Tus-GFP (46°C) to quantify the effect of stabilizing or destabilizing salts and additives on the k_{agg} of Tus-GFP (Figure 11A). Here, an increase in protein stability due to an additive is reflected by a decrease in k_{agg} compared to the Tus-GFP control sample (without additive; Figure 11C).
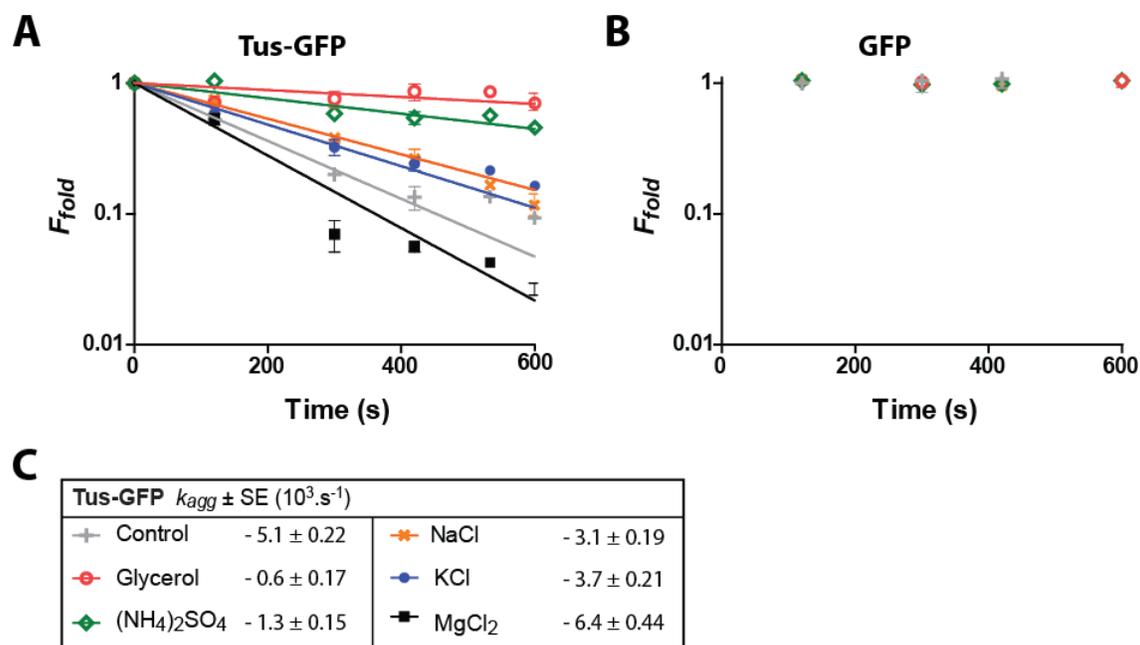
Figure 11: Effects of additives on Tus-GFP. (A) Aggregation rates (k_{agg}) at 46°C of Tus-GFP (12 μM) in the presence of additives. Additives were present at final concentrations of either 25 % glycerol, 0.3 M $(NH_4)_2SO_4$, 0.3 M NaCl, 0.4 M KCl or 0.3 M $MgCl_2$ in 0.5X buffer A. (B) Aggregation rates of GFP (13 μM) in the same conditions as Tus-GFP. (C) k_{agg} values obtained at 46°C by the S-method in triplicate.

As each additive or salt could influence the stability of the GFP portion of the fusion protein, it was essential to use a GFP control to show that the additive did not affect the fluorescence or stability of GFP (Figure 10B). The effect of additives commonly used in storage buffers to improve protein stability was tested and as expected, glycerol was found to have the largest stabilization effect, but $MgCl_2$ had a detrimental effect on Tus-GFP. These additives did not affect the stability or fluorescence of GFP. It is therefore possible to quickly screen for optimal protein storage conditions using GFP-Basta.

3.3.4 *TerB* induced stabilization of Tus

Tus is a DNA binding protein that binds to 21 bp DNA sequences called *Ter* (Kamada et al., 1996, Mulcair et al., 2006, Neylon et al., 2005). The tight binding ($K_D \sim$ nM) (Mulcair et al., 2006) of *TerB* to Tus-GFP should therefore induce a strong ligand-induced stabilization effect resulting in a large shift in T_{agg} (ΔT_{agg}). The T_{agg} of the Tus-GFP-*TerB* complex was first measured by a modified electrophoretic mobility shift assay (EMSA; Figure 12A). Tus-GFP and *TerB* were mixed in equimolecular quantities in low salt conditions ($K_D <$ pM) and treated at room temperature for 10 minutes to allow complex formation prior to being heat-treated at temperatures ranging from 37 to 67°C. Here, no centrifugation step is required as the Tus-GFP aggregates are retained in the wells of the agarose gel due to their low mobility and *Ter*-bound Tus-GFP proteins corresponding to F_{fold}, migrate more rapidly due to their increased net negative charge (*cf* section 3.2.5 for detailed procedure). The fluorescent bands corresponding to F_{fold} were integrated and revealed a T_{agg} of 63.5°C for Tus-GFP-*TerB* complex which compared to the T_{agg} of 42.8°C for Tus-GFP obtained with the S method in the same conditions (Figure 10B) corresponds to an increase in thermostability of 20.7°C (Figure 12A).
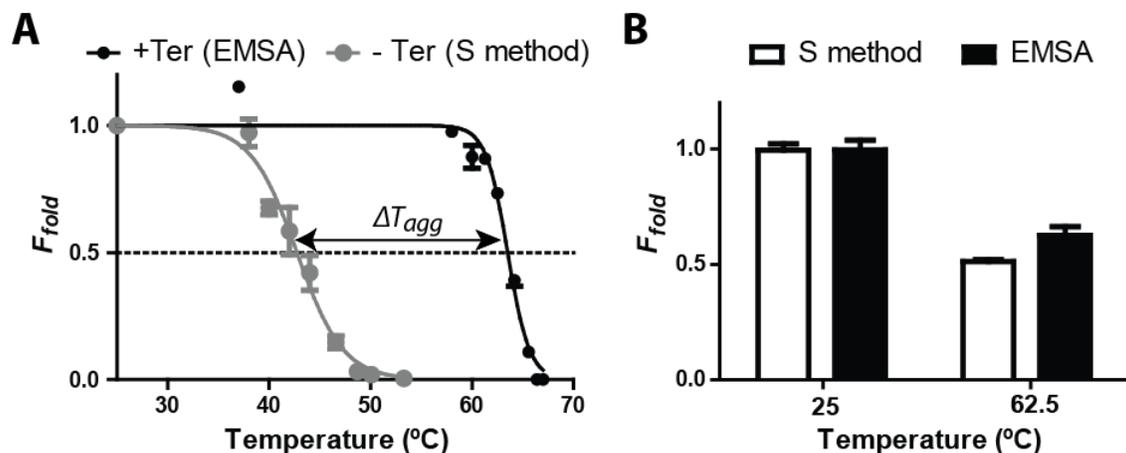
Figure 12: Thermal aggregation profiles of Tus-GFP and Tus-GFP-*Ter* complex. (A) The T_{agg} of Tus-GFP-*TerB* complex was obtained by EMSA (63.5 ± 0.2°C). The T_{agg} of Tus-GFP was obtained by S method (T_{agg}= 42.8 ± 0.3°C). The double-headed arrow indicates a ligand-induced ΔT_{agg} of 20.7°C. (B) Isothermal denaturation at 62.5°C for 5 minutes (triplicates) of Tus-GFP-*TerB* complex obtained by EMSA (F_{fold}= 0.51) and S method (F_{fold}= 0.63) in identical buffer conditions.

Heat induced aggregation of Tus-GFP-*TerB* complex was also determined with the S method and compared with the EMSA method in isothermal conditions (Figure 12B). The F_{fold} obtained with the two different methods were in good agreement confirming that F_{fold} obtained with the S method consists mainly of folded and active proteins.

3.3.5 Relationship between ligand affinity and aggregation rates of Tus-*Ter* complex

The ligand induced stabilization on Tus-GFP of various well-characterized *Ter* variants was investigated by isothermal aggregation reactions using the S method. The dissociation constants (K_D), for various Tus-*Ter* complexes (*TerB*, *Ter*-AG, *Ter*-AAG and TT-lock) were previously determined by surface plasmon resonance (SPR; Mulcair et al., 2006). The same oligonucleotides were used to determine their effect on Tus-GFP k_{agg} but they have each been extended with a 10-mer GC rich dsDNA region in order to obtain T_m values above 70°C for each of them (Table 10). Here, the relationship between K_D and k_{agg} was investigated in

conditions were the Tus-GFP-*Ter* complexes were at concentrations at least ~100-fold above their respective K_D to ensure that at least 99 % of proteins were in their bound form.

Table 10: Sequences and K_D values of *TerB* variants in 250 mM KCl. The *Ter* sequences used by Mulcair et al. (2006) to determine the K_D values using SPR are highlighted in grey.

	Ter	K_D (nM)
TerB	5'-CTTTAGTTACAACATACTTATCCCCGCCCCC GAAATCAATGTTGTATGAATAGGGGCGGGGG	1.4
Ter-AG	5'-CTTTAGTTACAACATACTTATCACCCCGCCCCC AATCAATGTTGTATGAATAGTGGGGCGGGGG	16.5
Ter-AAG	5'-CTTTAGTTACAACATACTTATCACCCCGCCCCC ATCAATGTTGTATGAATAGTGGGGCGGGGG	113
TT-lock	5'-CCCCCGCCCCCAATACTTTAGTTACAACATACTTAT GGGGGCGGGGGTTATGAAATCAATGTTGTAT	0.4

The k_{agg} values of the complexes were determined at 50°C in 250 mM KCl, where unbound Tus-GFP proteins aggregate very quickly. As expected, the k_{agg} values of the complexes increased with increasing K_D values (Figure 13A). The ligand-induced stabilization effects due to the gain of inter- and intramolecular interactions could simply be extracted by dividing the k_{agg} of Tus by the k_{agg} of the complex and this value should correlate with the K_D of the interaction.

Indeed, a linear correlation was obtained between $\ln(K_D)$ from published SPR data (Mulcair et al., 2006) and the $\ln(k_{agg(\text{Tus})}/k_{agg(\text{Tus-}Ter)})$ obtained with GFP-Basta (Figure 13B). Meng and coworkers recently studied the stabilization effect of glycerol on the irreversible thermal denaturation of creatine kinase using the activated-complex theory (Meng et al., 2004).

A

$k_{agg} \pm$ SE $(10^3.s^{-1})$	
K_D (nM)	
◆	0.45 ± 0.03
+ TT-lock	0.4
☐	1.39 ± 0.11
+ TerB	1.4
▲	6.24 ± 0.32
+ Ter-AG	16.5
■	9.55 ± 0.49
+ Ter-AAG	113
■	13.3 ± 0.76
Tus	-

B
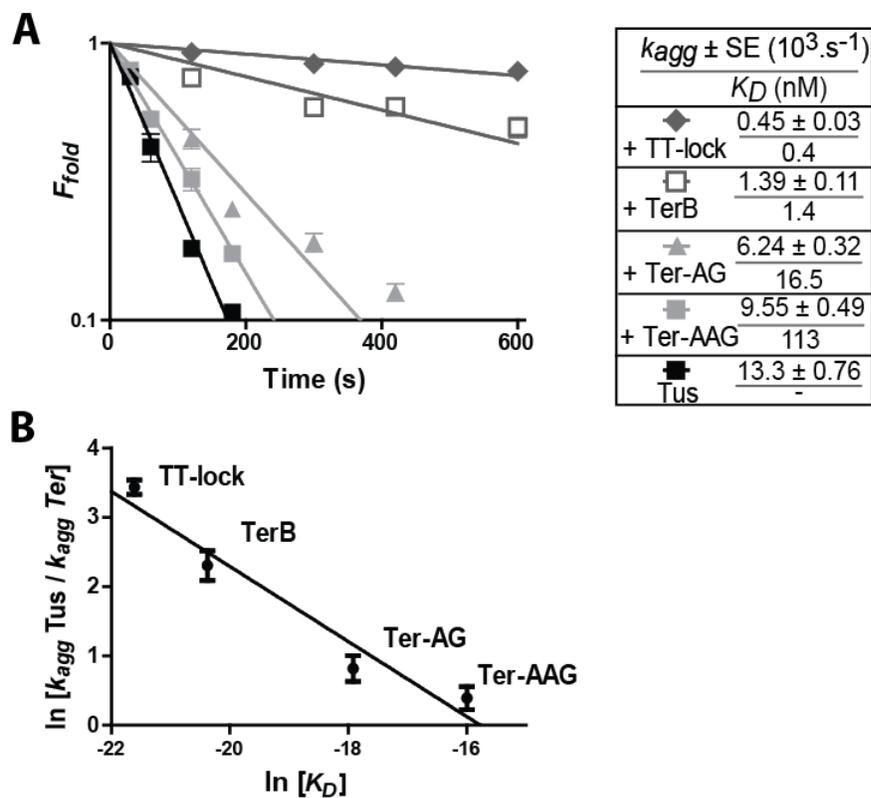
Figure 13: Correlation between k_{agg} and K_D. (A) k_{agg} **of Tus-GFP in complex with *Ter* variants (*TerB*, *Ter*-AG, *Ter*-AAG or TT-lock) were determined at 50°C by the S method. (B) Correlation between ln(K_D) from published SPR data (Mulcair et al., 2006) and the ln($k_{agg(\text{Tus})}/k_{agg(\text{Tus-}Ter)}$).**

This theory was used to demonstrate the relationship between ln(K_D) and the ln($k_{agg(\text{Tus})}/k_{agg(\text{Tus-}Ter)}$) seen in Figure 13B. The activated complex is an intermediate transition state between reactants and products. The activated-complex theory postulates the existence of an equilibrium between reactants (P) and the activated complex (P*). In this case, the kinetic scheme of irreversible denaturation and aggregation of a protein is expressed as:

$$P \overset{K^*}{\Leftrightarrow} P^* \xrightarrow{k_{agg}} U \rightarrow Aggregate$$
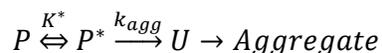
Where U is the unfolded aggregation competent intermediate (Figure 9A). If the fraction of unfolded proteins after heat denaturation is driven into an irreversible aggregation pathway then the extent of aggregation should therefore reflect the proportion of unfolded proteins. In this case, the apparent aggregation rate constant k_{agg} is related to the change in free energy of

activation ΔG* and can be expressed in accordance with the activated-complex theory (Meng et al., 2004) as:

$$k_{agg} = (\frac{k_B T}{h})e^{\frac{-\Delta G^*}{RT}}$$

which can be transformed to:

$$\Delta G^* = -RT ln k_{agg}(h/k_B \text{ T})$$

The difference in change of free energy of activation (ΔΔG*) between Tus-GFP (ΔG*$_{(Tus)}$) and Tus-GFP-ligand (ΔG*$_{(Tus\text{-}Ter)}$) can be obtained with the following expression:

$$\Delta\Delta G^* = \Delta G^*_{(Tus)} - \Delta G^*_{(tus-Ter)} = -RT \ln(\frac{k_{agg(Tus)}}{k_{agg(Tus-Ter)}})$$

ΔG* is connected with the equilibrium constant by the relationship ΔG* = –RTln K*. This term can be replaced in the previous equation giving:

$$ln K^*_{(Tus)} - ln K^*_{(Tus-Ter)} = \ln(k_{agg(Tus)}/k_{agg(Tus-Ter)})$$

In the situation where most of Tus is in complex with its ligand, the term lnK* Tus-*Ter* can be represented as the sum of lnK*(Tus) and the ligand–induced stabilization of Tus given by lnK*(ligand effect). The previous equation can therefore be simplified as:

$$-ln K^*_{(ligand\ effect)} = \ln(\frac{k_{agg(Tus)}}{k_{agg(Tus-Ter)}})$$

lnK*(ligand effect) is proportional to the ΔΔG* induced only by ligand binding and should therefore be proportional to lnK_D of the Tus-ligand complex. To test this, the term *lnK**$_{(ligand\ effect)}$ was replaced with the term lnK_D in the last equation and a linear correlation was obtained between ln($k_{agg(Tus)}$/$k_{agg(Tus\text{-}Ter)}$) and lnK_D. GFP-Basta can therefore be used to accurately estimate the K_D values of *Ter* variants using this reference curve (Figure 13B).

3.4 Discussion

The T_{agg} of Tus and Tus-GFP were very similar with no more than ~1°C difference and the same was observed for CAT and glycerol kinase (Moreau et al., 2010). Given that the T_{agg} of GFP was higher than the T_{agg} of POI-GFP, the main driving force in POI-GFP aggregation must be the unfolding of POI that subsequently causes aggregation of the whole fusion protein. This result demonstrates that the thermal denaturation of a fusion protein can essentially be driven by the less stable domain as long as these domains are uncoupled. A small loss of GFP fluorescence of ~10 % occurred between 35 and 53°C and has been described in the literature (Vessoni Penna et al., 2004). The T_{agg} of GFP was determined to be 79.6°C using the S method. The thermodynamic properties of GFP have been thoroughly studied (Crameri et al., 1996, Ishii et al., 2007a, Ishii et al., 2007b, Penna et al., 2005, Tsien, 1998, Ward and Bokman, 1982). Ward and Bokman reported a 50 % loss of GFP fluorescence at 78°C (Ward and Bokman, 1982). More recent studies measured the stability of GFP by isothermal denaturation at 80, 85 and 95 °C (Ishii et al., 2007a). These experiments attest that the loss of GFP fluorescence due to the unfolding of the β-barrel around the fluorophore (Tsien, 1998) is accompanied by aggregation and that the S method is an accurate tool for measuring protein stability. The GFP variant used in this study (uvGFP) is resistant to pH between 5.5 and 12 with optimum pH between 7 and 8 (Penna et al., 2005). The assay must therefore be carried out under experimental conditions that do not significantly affect GFP stability. The loss of fluorescence signal occurring around 75°C under physiological conditions and the intrinsic instability of GFP at pH < 5.5 are limitations for GFP-Basta compared to other methods such as DSF (Ericsson et al., 2006, Niesen et al., 2007, Pantoliano et al., 2001, Vedadi et al., 2006) or ITD (Epps and Taylor, 2001, Foster et al., 1999, Leung et al., 1996) but they are nevertheless negligible as a large majority of proteins are much less stable than GFP.

Thermal aggregation profiles were obtained to define the T_{agg} of the proteins and therefore the temperature at which isothermal aggregation reaction should be performed. Thermal scans are not very sensitive to detect small changes in thermostability. The isothermal aggregation method was therefore favored, since it is more sensitive and requires fewer data points to determine accurate k_{agg}. Under conditions of high irreversibility, a kinetic analysis is both more appropriate and more informative (Lepock et al., 1992). Indeed, GFP-Basta could readily detect modest stabilizing/destabilizing effects of additives on overall protein stability (Moreau et al., 2010). GFP-Basta was also able to identify and rank *Ter* ligands according to their reported K_D (Mulcair et al., 2006) for Tus using the S method. Indeed, a linear correlation was obtained between lnK_D from published SPR data (Mulcair et al., 2006) and the ln($k_{agg(Tus)}/k_{agg(Tus-Ter)}$) from GFP-Basta data (Figure 13B) and was demonstrated mathematically using the activated-complex theory (Meng et al., 2004). Although more data need to be acquired to confirm this relationship with other proteins and ligands, it nevertheless shows that GFP-Basta can accurately identify and classify similar ligands according to their affinity.

Many additional applications can be predicted. GFP-Basta could be used to screen libraries of mutants for improved thermostability or to identify unstable domains in multidomain proteins. It could also be used as a rapid screening assay to identify protein-protein interactions as the aggregation of the POI can be monitored in a mixture of proteins. It is also expected that the effect of inhibitors disrupting protein-protein interactions could be identified using GFP-Basta. Most of these applications are not possible with other thermal denaturation based methods as their main limitation is to be performed with pure proteins. GFP-Basta also eliminates the requirement for fluorescent dyes, some of which have been shown to provoke unwanted interferences upon non-specific binding to proteins (Lavinder et

al., 2009). Finally, GFP-Basta is also amenable to the screening of inhibitors for protein aggregation diseases as GFP is very stable (Kim et al., 2006).

GFP-Basta is adaptable to different formats (e.g. 96-well plates, SDS-PAGE, EMSA) and the protocol is fast. Each experiment (i.e. condition) was studied in 6-10 µl reaction volume, allowing for the screening of about 300 conditions per mg of protein. It is expected that, by using only one time point measurement and appropriate robotics, one set of reactions should take ~40 minutes from denaturation to data acquisition. GFP-Basta could therefore be used for the HT screening of stabilizing compounds and ligands.

3.5 Conclusion

In conclusion, GFP-Basta can be used to provide quantitative information on the stability and ligand affinity of proteins, regardless of their quaternary structure. GFP-basta can identify stabilizing compounds and ligands and, most importantly, affords a mean to correlate the affinity of various DNA ligands with respect to their effect on the k_{agg} of Tus using the activated-complex theory (Meng et al., 2004). GFP-basta can therefore be used to study the affinity of Tus for the ten *Ter* sites which is presented in the following chapter.

Chapter 4: Differential Tus–*Ter* binding and lock formation: implications for DNA replication termination in *Escherichia coli*

The majority of this chapter has been published in Molecular BioSystems[2] (Moreau, M.J. and Schaeffer, P. M.). The crosslinking study is part of a paper in preparation[3] (Oakley, A.J., Moreau, M. J., Schaeffer, P.M., and Dixon N.E.) and has been added to this chapter as it provides structural insights on the TT-lock.

4.1 Introduction

In *E. coli*, two replisomes are assembled at the unique origin of replication *oriC* and proceed bidirectionally to replicate the circular chromosome until they meet in the replication termination region (Neylon et al., 2005). This termination region is defined as the section of the chromosome containing a series of termination (*Ter*) sites (Figure 14A). These sequences were originally identified as 21 bp in length (Hidaka et al., 1988, Hill et al., 1988a) with a highly conserved 11 bp core sequence (Figure 14B). The ten primary *TerA-J* sites are spread

---

[2] Moreau, M. J. and Schaeffer, P. M. 2012. Differential Tus-*Ter* binding and lock formation: implications for DNA replication termination in *Escherichia coli. Mol Biosyst,* 8**,** 2783-91.
[3] Oakley, A.J., Moreau, M. J., Schaeffer, P.M., and Dixon N.E. Flexibility in the Tus-*Ter* Complex Modulates its Function. In preparation.

over 2800 kb and are arranged in two clusters of five sites, with one cluster on each side of

the region directly opposite to *oriC*. This polar organisation creates a "fork trap" to constrain

forks meeting in the terminus region (Duggin et al., 2008, Hill et al., 1987). Replisomes can

proceed through the first cluster of *Ter* sites on their way to the terminus, but will be stopped

by the second cluster containing *Ter* sites in opposite orientation (Figure 14A).
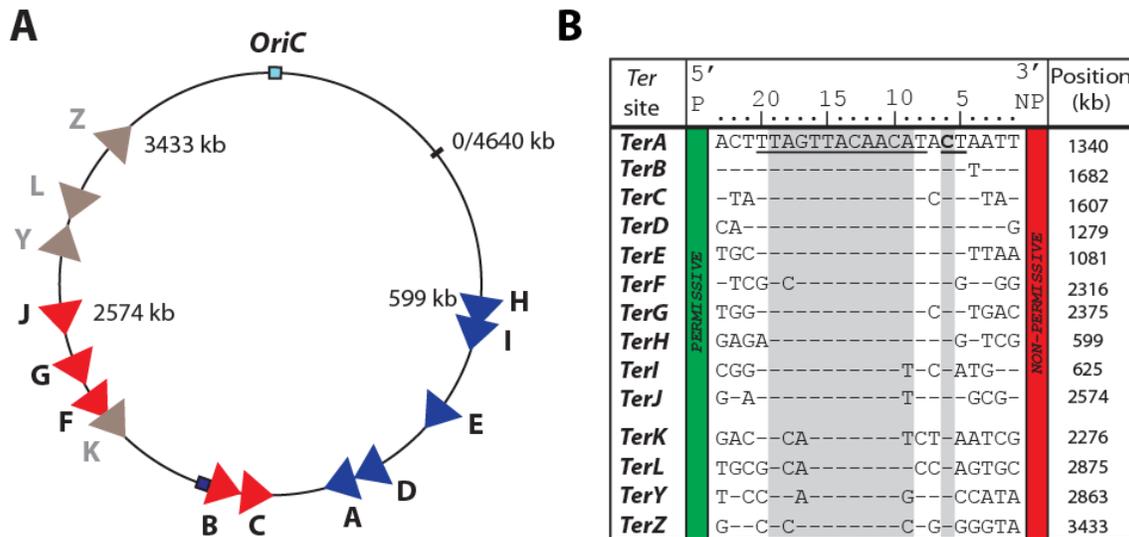
Figure 14: Genomic location and sequence identity of *Ter* sites. (A) Relative position of *Ter* sites on the *E. coli* DH12S chromosome. The arrowheads represent the orientation of *Ter* sites with the base of the arrowhead representing the permissive face of the Tus–*Ter* complex and the tip representing the non-permissive face. The cluster of red arrowheads arrests fork progressing in the clockwise direction and the blue cluster stops the anti-clockwise progressing fork. The brown arrowheads represent four recently identified *Ter* sites. The blue square represents the *tus* gene. (B) Sequence of all *Ter* sites and their position on the chromosome. The C(6) responsible for the TT-lock formation is in bold. Bases forming direct contacts with Tus in the crystal structure are underlined.

The cluster including *TerB*, *C*, *F*, *G* and *J* is oriented to block a clockwise moving fork

whereas *TerA*, *D*, *E*, *I* and *H* are oriented to block anti-clockwise moving forks (de Massy et

al., 1987, Hill et al., 1987, Neylon et al., 2005). This polar fork arrest is mediated by the

asymmetrical binding of Tus to *Ter* sites that creates a complex with a permissive face

allowing fork progression and a non-permissive face that stalls the fork (Neylon et al., 2005).

It is postulated that when DnaB helicase at the forefront of the replisome encounters the permissive face of the Tus-*Ter* complex, Tus is displaced upon duplex DNA separation and the fork progresses unimpeded. However, when DnaB unwinds the duplex DNA ahead of the non-permissive face of the Tus-*Ter* complex, a G·C base pair located at position 6 in the *Ter* core sequence is broken and the C(6) moves 14 Å from its normal position to bind tightly in a cytosine-specific binding pocket at the surface of the non-permissive face of Tus. This base flipping results in a tighter interaction between Tus and *Ter*, called the Tus-*Ter*-lock (TT-lock), which stalls the replication fork until the second replisome arrives (Mulcair et al., 2006).

Surprisingly, it has been shown that the protein Tus is not essential to *E. coli* survival (Hill, 1992, Hill et al., 1987, Hill et al., 1989, Roecklein et al., 1991) and that it is conserved only in closely related bacteria (Neylon et al., 2005). Although an analogous system exists for replication termination in *B. subtilis*, it involves a different replication terminator protein (RTP) and termination sites using a different mechanism (Bussiere et al., 1995, Duggin, 2006, Vivian et al., 2007). More recently and after a long search, 71 chromosomal termination regions *TER* containing fork pausing elements were identified in budding yeast; these involve binding of Top2 DNA topoisomerase highlighting the biological importance of replication fork barriers (Fachinetti et al., 2010).

In *E. coli*, the *dif* site was recently proposed as an alternative termination site (Hendrickson and Lawrence, 2007). It is the site of action of the XerCD site-specific DNA recombinase and is located 18 kbp from *TerC* (Blakely and Sherratt, 1994, Hendrickson and Lawrence, 2007). According to the replication fork trap model, the position of *Ter* sites restricts replication fork fusion to the terminus region. Duggin and Bell examined DNA replication intermediates at *Ter* sites and *dif* and identified two definitive signatures of site-specific termination at *Ter* sites thus supporting the fork trap model (Duggin and Bell, 2009).

To date, 14 *Ter* sites have been identified in *E. coli* (*TerA-L*, *TerY* and *TerZ*), of which nine have been derived by consensus sequence search using the *E. coli* GenBank database (Coskun-Ari and Hill, 1997, Duggin and Bell, 2009, Sharma and Hill, 1992). However, the recently identified *TerK*, *L*, *Y* and *Z* were found to be very weak DNA replication fork pausing sites (Duggin and Bell, 2009). The remaining *Ter* sites (*TerA-J*) were able to arrest forks but they significantly differed in their efficiency (*cf* Figure 6 p22; Duggin and Bell, 2009). The TT-lock formation was recently proposed to only act as a fail-safe mechanism after it was demonstrated that its formation was not essential to block the activity of DnaB helicase *in vitro* (Bastia et al., 2008).

The consensus based-identification of *Ter* sites, the variation in their efficiency in pausing replisomes, and the controversy about the TT-lock raises two essential questions: (a) what is the affinity of Tus for the different *Ter* sites and does it correlate with the sites that are most often used in DNA replication pausing; and (b) are all ten *Ter* sites capable of forming the TT-lock to block fork progression and does it correlate with their efficiency in pausing forks *in vivo*. A combination of SPR and GFP-Basta (Chapter 3; Moreau et al., 2010) was used to determine the thermodynamic and kinetic parameters of Tus binding to the ten primary *TerA-J* sites and their respective lock-forming sequence variants (*Ter-lockA-J*) in order to better understand the role of the ten *Ter* sites in termination and to evaluate their ability to form TT-locks. This study provides detailed mechanistic information on the Tus-*Ter* interactions and explains their differences in fork arrest efficiency *in vivo*.

4.2 Material and methods

4.2.1 Protein expression and purification

The His_6-Tus-GFP (referred as Tus-GFP) and His_6-Tus (referred as Tus) proteins were expressed in *E. coli* BL21(*DE3*)RIPL and affinity purified with Profinity IMAC Ni-charged

resin as previously described in Dahdah et al. (2009). After ammonium sulphate precipitation, the Tus-GFP protein pellets were resuspended in buffer A (50 mM Tris, pH 7.6, 250 mM KCl, 0.1 mM EDTA and 0.2 mM β-mercaptoethanol) and dialysed (SnakeSkin pleated dialysis tubing 10,000 MWCO; Pierce) twice against 200 ml of buffer A at 4°C. Tus was expressed and purified as for Tus-GFP. After ammonium sulphate precipitation, Tus was resuspended in buffer A. The purity of proteins was assessed by SDS-PAGE (NEXT-GEL Amresco) and concentration was determined by standard Bradford assay.

4.2.2 GFP-Basta

Oligonucleotides used for GFP-basta are described in Table 5, p45. *Ter* and *Ter*-lock DNA were designed to include the 23 bp *Ter* or *Ter-lock* sequences followed by a stabilizing 10-mer GC rich region in order to elevate their T_m values above 70°C. An example of the design is shown in Figure 15A-B.

The aggregation rate constants of Tus-GFP alone or in complex with each *Ter* or *Ter-lock* sequence were determined by the isothermal method of GFP-Basta (*cf* section 3.2.4 p65; Moreau et al., 2010). For these reactions, an equal volume of Tus-GFP (1.6 µM) in buffer A or buffer B (buffer A with 150 mM final KCl concentration) was mixed with an equal volume of *Ter* or *Ter-lock* DNA (2 µM) in the corresponding buffer A or B. The reactions were left 10 minutes at room temperature to allow complex formation. Each reaction (70 µl) was heated at a constant temperature in a MyCycler (BioRad), i.e. at either 52°C in 250 mM KCl or at 58°C in 150 mM KCl reaction. After heating, samples were transferred to ice for 10 minutes to stop the reaction. Aggregates were then centrifuged at 18,000 rpm for 20 minutes at 4°C in a Beckman Coulter Microfuge 22R centrifuge using the rotor F12x8.2. The residual fluorescence in the supernatant after thermal denaturation was quantified by transferring 60 µl of the supernatant into a black 96-well plate (Nunclon) and the residual fluorescence was

measured with a fluorescence plate reader (Victor V Wallace Perkin-Elmer). The excitation and emission filters were set at 460 nm and 535 nm respectively, with 40 nm bandwidth. The values obtained were normalized against the fluorescence of an untreated sample. Aggregation curves were fitted as described previously to obtain aggregation rate constants k_{agg} (*cf* section 3.2.4 p65). Aggregation half-lives ($t_{1/2\text{-}agg}$) were obtained as $\ln2/k_{agg}$.

4.2.3 SPR

For SPR experiments, all *Ter* and *Ter-lock* were designed to include a single-stranded decamer overhang (velcro) after base 23 to allow their hybridization to a biotinylated complementary oligonucleotide immobilized on the NLC Chip via a neutravidin coated surface (Table 6 p46). An example of the design is shown in Figure 16A. Individual oligonucleotides were resuspended in TE_{50} (10 mM Tris-HCl (pH 7.6), 1 mM EDTA, 50 mM KCl) to a final concentration of 100 µM. They were combined by mixing 25 µl of the oligonucleotide containing the 10-mer overhang with 50 µl of the complementary oligonucleotides, and 175 µl of buffer A. Hybridization was achieved by heating at 80°C for 2 minutes followed by slow cooling to room temperature.

Measurements were carried out at 20°C using a ProteON XPR36 (Bio-Rad) with freshly diluted Tus in buffer A or B. The biotinylated pCBio (5'-Biotin-CCCCGCCCCC-3') was used as a molecular "velcro" to capture the *Ter* oligonucleotides on the neutravidin NLC chip (Bio-Rad). The pCBio was immobilized onto the surface at 50 nM for 300 s at 25 µl/min. *Ter* or *Ter-lock* DNA were hybridized through their complementary single stranded G_5CG_4 overhang to the pCBio at a concentration of 25 nM and flow rate of 25 µl /min during ~100 s. The kinetics of complex formation between Tus and *Ter* were measured in buffer A and B. Six Tus concentrations ranging from 100 nM to 3.125 nM in buffer A and from 30 nM to 0.91 nM in buffer B were injected at a flow rate of 25 µl /min for 120 s, and dissociations

were analysed over 900 s. When required, Tus was dissociated from *Ter* sequences with 1 M NaCl injections (25 µl/min for 120 s). The surface was regenerated with 50 mM NaOH and 1 M NaCl (30 µl/min for 60 s), leaving the pCBio on the surface. Experiments were carried out at least in triplicate and fit to the Langmuir binding model with all the variables fitted locally. For graphical representation and to facilitate visual comparison of *Ter versus Ter-lock* sequences, only one representative concentration was shown for each sequence. All curves were normalised by the RU value obtained at the end of Tus injection (t = 120 s).

4.2.4 Photo-crosslinking

The sequences used for the photo-crosslinking are double stranded *TerB* modified at position 7 with an adenine instead of a thymine (P1), the TT-lock variant of wild type *TerB* (P2) and the TT-lock variant of the T7A mutated *TerB* (P3). They harbor a bromodeoxyuridine (BrdU, underlined) at position 7 in the core sequence as shown below:

```
P1:   5'-ATAAGAATGTTGTAACTAAAG
         TATTCBTACAACATTGATTTC


P2:   5'-     TATGTTGTAACTAAAG
         TATTCBTACAACATTGATTTC


P3:   5'-     AATGTTGTAACTAAAG
         TATTCBTACAACATTGATTTC
```

The photo-crosslinking experiment was performed as described in Dahdah et al. (2009) with the following modifications: wtTus and F140A Tus were dialyzed in crosslinking buffer (50 mM phosphate, pH 7.8, 10 % glycerol and 2 mM β-mercaptoethanol). Oligonucleotide pairs were prepared by mixing equal volumes of each oligonucleotide to a final concentration of 50 µM in TE (10 mM Tris (pH 7.5), 0.1 mM EDTA) supplemented with 125 mM KCl and annealed by heating 2 minutes at 73°C and slowly cooling down to room temperature. Wild

type Tus and F140A Tus were diluted in crosslinking buffer to 38 µM. Equal volumes (6 µl) of proteins at 38 µM and oligonucleotides at 50 µM were mixed and incubated at room temperature for 10 minutes to allow protein-DNA complex formation. For each complex, two drops of 5 µl were spotted under the cover of a Nunclon 96-well plate with a pre-cooled block (-20°C) placed on top to avoid overheating of proteins. The drops were then irradiated for nine minutes using a UV-transiluminator set at 312 nm (Vilber Lourmat). The two drops for each complex were then pooled and 10 µl of 2X SDS-PAGE loading dye were added. Samples were then heated for 2 minutes at 65°C and 15 µl were loaded in a 10% SDS-PAGE gel (Next gel 10 %, Amresco). Gels were run for 1 hour at 120 V, stained with Coomassie Brilliant Blue for 30 minutes and destained in a solution of 40 % isopropanol/10% acetic acid. Gels were scanned and bands were quantified using the ImageJ software.

4.3 Results

4.3.1 Comparison of the effect of *Ter* and *Ter-lock*-induced thermal stabilization on Tus-GFP by GFP-Basta

A new method for quantifying the strength of protein-ligand interactions was previously developed using a GFP reporter system called GFP-Basta, and showed it was a reliable method to study protein-DNA interactions (Chapter 3; Moreau et al., 2010, Morin et al., 2012). Here GFP-Basta was used to compare the binding of Tus to the ten different *TerA-J* sites (Figure 15A) and their lock-forming *Ter-lockA-J* variants (Figure 15B) to determine if the stability of these complexes correlates with their efficiency in fork pausing seen *in vivo*. *Ter-lock* sequences were partially single-stranded at the non-permissive end until the C(6) that is critical for TT-lock formation (Figure 15B).

A   *Ter*

CCCCGCCCCCACTTTAGTTACAACATA**C**TAATT
GGGGCGGGGGTGAAATCAATGTTGTATGATTAA

B   *Ter-lock*

CCCCGCCCCCACTTTAGTTACAACATA**C**TAATT
GGGGCGGGGGTGAAATCAATGTTGTAT
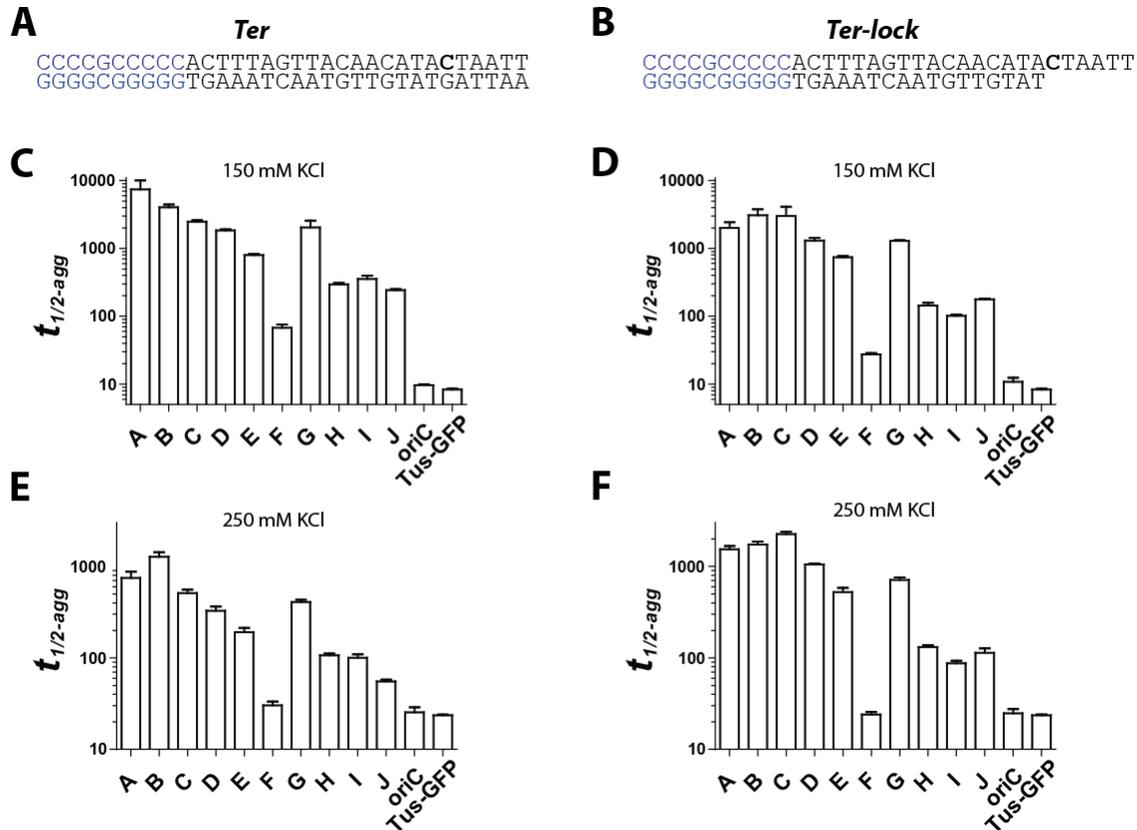
Figure 15: DNA-induced thermal stabilization of Tus-GFP. Sequence and structure information of the *TerA* (A) and *Ter-lockA* (B) sequences with their stabilizing decamer depicted in blue. First-order aggregation kinetics were measured at 58°C in 150 mM KCl for *Ter* (C) and *Ter-lock* (D) sites, and at 52°C in 250 mM KCl for *Ter* (E) and *Ter-lock* (F) sites. The error bars represent the upper and lower limit of the 95% CI of the mean obtained from $t_{1/2-agg}$ in duplicates. Values of $t_{1/2-agg}$ for each *Ter* and *Ter-lock* site in 250 mM KCl are given in Table 11. See also Appendix A for aggregation rates of reactions.

Aggregation rate constants (expressed in half-life values $t_{1/2-agg}$) of Tus-GFP in complex with either *Ter* or *Ter-lock* sequences were obtained in buffers containing 150 mM KCl at 58°C (Figure 15C-D; lower-salt and high-affinity conditions) and 250 mM KCl at 52°C (Figure 15E-F; high-salt and moderate-affinity conditions) to evaluate the effect of ionic strength on complex stability (Mulcair et al., 2006, Neylon et al., 2000). Additionally, a sequence derived from *oriC* as well as its partially single-stranded variant (*oriC-lock*, *cf* Table 5 p45) were used to evaluate the stabilization effects of non-specific DNA-binding on Tus.

Aggregation profiles obtained for the Tus-*Ter* and Tus-*Ter-lock* complexes in lower- or high salt conditions were similar within each group (Figure 15C-F and Table 11). In both salt conditions, the strongest binders were *TerA-E* and *G*, and their respective *Ter-lock* sequences, whereas moderate binders consisted of *TerH*, *I* and *J* and their respective *Ter-lock* sequences. Surprisingly, the stabilizing effects of *TerF* and the *Ter-lockF* on Tus were comparable with those of the non-specific *oriC* in 250 mM KCl, immediately suggesting that *TerF* is not a functional *Ter* site (*cf* values in Table 11). This further suggests that *TerF* forms mainly non-specific electrostatic interactions with Tus and cannot bind Tus-GFP at 250 mM KCl at a concentration of 1 μM. Nevertheless, in lower-salt conditions, a weak stabilizing effect was observed from binding of *TerF* or *Ter-lockF* to Tus that were significant compared to the nonspecific *oriC* suggesting that some specific interactions still occur. In high-salt, *Ter-lockA-E,G,J* were more stabilizing than their respective *Ter* sites suggesting that formation of the TT-locks is impaired for *TerF,H-I* (Figure 15F and Table 11). On the contrary, in low-salt, except for *TerC,* all remaining *Ter* sites were systematically more stabilizing than the *Ter-locks* (Figure 15C-D and Table 11). This is most likely due to the loss of non-specific electrostatic interactions resulting from six nucleotides that are missing in the partially single stranded *Ter-lock* oligonucleotides when compared to their double-stranded *Ter*. Here, the difference in net electrostatic interactions between the two species could very well explain these data.

Table 11: DNA-induced thermal stabilization of Tus-GFP.

Ter site	150 mM KCl $t_{1/2\text{-}agg}$ ± SEM (s)		250 mM KCl $t_{1/2\text{-}agg}$ ± SEM (s)	
	Ter	*Ter-lock*	*Ter*	*Ter-lock*
A	7445 ± 2653	2009 ± 436	746 ± 129	1532 ± 138
B	4064 ± 362	3085 ± 694	1271 ± 157	1721 ± 151
C	2484 ± 130	2997 ± 1107	511 ± 46	2250 ± 132
D	1848 ± 69	1305 ± 120	329 ± 36	1045 ± 27
E	799 ± 31	744 ± 35	191 ± 22	523 ± 59
F	68 ± 8	28 ± 1	30 ± 3	24 ± 1
G	2044 ± 519	1299 ± 32	407 ± 23	708 ± 50
H	294 ± 17	144 ± 14	107 ± 5	131 ± 7
I	355 ± 45	103 ± 3	100 ± 9	88 ± 5
J	244 ± 8	178 ± 4	55 ± 2	114 ± 13
oriC	10 ± 1	11 ± 2	25 ± 3	25 ± 3
No DNA	8 ± 0.2		24 ± 0.4	

Mean values and SEM of $t_{1/2\text{-}agg}$ for each *Ter* and *Ter-lock* site in low-salt (150 mM KCl) and high-salt (250 mM KCl) conditions. Tus-GFP aggregation reactions were measured at 52°C in 250 mM KCl and at 58°C in 150 mM KCl for all *Ter* and *Ter-lock* sequences.

The stabilization effects observed for *TerH-J* and their respective *Ter-locks* were not uniform. Only for *TerJ*, which is the weakest of the *Ter* sites in this group, could be observed a significant increase in stabilization effect when the *Ter-lockJ* was bound to Tus-GFP in high salt conditions (Figure 15D).

If we consider the combined stabilizing effects of either *Ter* or *Ter-lock* sites, *TerA-D* are clearly the strongest binders. These sites are equivalently placed in both *Ter* clusters at their most proximal regions to the terminus region. *TerG*, and *E*, are the next strongest binding sites, followed by the moderate binder *TerJ*, *H* and *I* and the weak binder *TerF*. As a result, there are three strong and two moderate to weak binding sites on each side of the termination region (Figure 14A).

4.3.2 Kinetics of binding of Tus to *Ter* and *Ter-lock* sequences by surface plasmon resonance

Although the GFP-Basta data already suggested that all *Ter* sites might not be able to form a TT-lock equally well, the difference in stabilization observed could still be due to additional electrostatic interactions in the double stranded *Ter* sites compared to the *Ter-lock*. To obtain a clearer answer to the question of whether all ten *Ter* sites are capable of forming a TT-lock to block fork progression, the kinetic parameters for the binding of Tus to the ten *TerA-J* and their respective *Ter-lock* sequences were determined by surface plasmon resonance using a ProteON XPR36 (BioRad) instrument.

A universal biotinylated-polyG "velcro" was designed to reversibly immobilize the different *Ter* and *Ter-lock* sequences on a neutravidin-coated surface so as to massively reduce the cost of this study (Figure 16A). Where possible, the kinetic parameters of Tus binding to each *Ter* and *Ter-lock* sites were determined in 250 mM KCl (Figure 16B) and 150 mM KCl (Figure 16C). It was not possible to fit all data sets obtained at a single KCl concentration because at 150 mM KCl, strong binders dissociated immeasurably slowly, while at 250 mM KCl weak and moderate binders could not reach their maximal binding values (R_{max}) within the tested concentrations. For instance, in 250 mM KCl, the binding of Tus to the moderate binders, *TerH*, *TerI* and *TerJ*, reached only 46%, 24% and 10% of the R_{max}, and 13%, 12%, and 11% to their *Ter-lock* analogues respectively at a concentration of 100 nM. The kinetic parameters – i.e. association rate constant (k_a), dissociation rate constant (k_d), half-life of dissociation ($t_{1/2}$) and equilibrium dissociation constant ($K_D=k_d/k_a$) were determined for each complex. For clarity, only K_D and $t_{1/2}$ values are shown in Figure 16. The k_a obtained were similar for most *Ter* and *Ter-lock* sequences (Table 12). Accurate k_a values could not be determined for *TerF,H-J, Ter-lockF,H-J* in high-salt for the reasons described above.

A

Velcro

```
N
L    B — CCCCGCCCCC
C
```

TerA

```
            ACTTTAGTTACAACATACTAATT
GGGGCGGGGGTGAAATCAATGTTGTATGATTAA
```

or

TerA-lock

```
            ACTTTAGTTACAACATACTAATT
GGGGCGGGGGTGAAATCAATGTTGTAT
```

+ →

```
N
L   B — CCCCGCCCCCACTTTAGTTACAACATACTAATT
C        GGGGCGGGGGTGAAATCAATGTTGTATGATTAA
```

```
N
L   B — CCCCGCCCCCACTTTAGTTACAACATACTAATT
C        GGGGCGGGGGTGAAATCAATGTTGTAT
```

B

250 mM KCl	A		B		C		D		E		F		G		H		I *		J *	
	ds	tt	ds	tt	ds	tt	ds	tt	ds	tt	ds	tt	ds	tt	ds	tt	ds	tt	ds	tt
$T_{1/2}$ (s)	315.8	3310	380	4367	243	3409	217	2589	137	2979	NA	NA	157	2277	31.6	59.1*	28.7	196	27.3	332
K_D (nM)	10.7	1.34	11.6	1.05	18	1.24	17	1.99	25.6	2.05	NA	NA	16.7	3.25	NA	NA	NA	NA	NA	NA

C

150 mM KCl	A		B		C		D		E		F		G		H		I		J	
	ds	tt	ds	tt	ds	tt	ds	tt	ds	tt	ds	tt	ds	tt	ds	tt	ds	tt	ds	tt
$T_{1/2}$ (s)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	31.8	19.4	NA	NA	1244	412	949	1069	357	1322
K_D (nM)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	88.4	NA	NA	NA	1.36	3.34	1.32	2.05	6.07	4.18

Figure 16: Surface plasmon resonance binding kinetics of Tus to the different *Ter* and *Ter-lock* sequences. (A) Principle of the reversible ''velcro'' surface. Red and blue sequences represent the ''velcro'' complementary base pairing sequences. (B) Binding kinetics obtained in 250 mM KCl (see also Table 12 for k_a, k_d and ± SEM values). (C) Binding kinetics in 150 mM KCl (see also Table 12 for k_a, k_d and ± SEM values). All sensorgrams were normalized to a R_{max} value of 1 to allow for direct visual comparison of their $t_{1/2}$ values of dissociation. All curves were normalised by the RU value obtained at the end of the Tus injection (t = 120 s). The blue and red curves represent the sensorgrams of Tus binding to *Ter* and *Ter-lock* sequences respectively. The $t_{1/2}$ and K_D are given as mean values. *: $t_{1/2}$ value was determined by direct visual analysis of the dissociation phase. For direct comparison of *Ter versus Ter-lock* sequences, only one representative curve is shown for each sequence.

In our SPR experiment, a ~20-fold lower k_a and a 3-fold longer $t_{1/2}$ were observed for *TerB* compared to the most recent study run in similar conditions (Mulcair et al., 2006). This variation in k_a is probably due to the radically different surface chemistry and oligonucleotide binding method used in the other study – i.e. alginate surface and "velcro" in our study *versus*

Table 12: Kinetic parameters of Tus affinity for *Ter* sites and their *Ter*-lock variants at 250 and 150 mM KCl.

	$10^{-5}\,k_a \pm$ SEM $(M^{-1}\,s^{-1})$	$10^{3}\,k_d \pm$ SEM (s^{-1})	K_D (nM)	$t_{1/2}$ (s)
TerA	2.05 ± 0.17	2.2 ± 0.07	10.7	315.8
Ter-lockA	1.56 ± 0.29	0.21 ± 0.0.5	1.34	3309.9
TerB	1.57 ± 0.12	1.83 ± 0.08	11.6	379.8
Ter-lockB	1.52 ± 0.20	0.16 ± 0.01	1.05	4366.9
TerC	1.58 ± 0.12	2.86 ± 0.17	18	242.6
Ter-lockC	1.64 ± 0.16	0.20 ± 0.01	1.24	3408.9
TerD	1.88 ± 0.16	3.2 ± 0.15	17	217.4
Ter-lockD	1.34 ± 0.17	0.27 ± 0.016	1.99	2588.8
TerE	1.97 ± 0.1	5.1 ± 0.15	25.6	137.0
Ter-lockE	1.13 ± 0.07	0.23 ± 0.005	2.05	2979.3
TerF	2.47 ± 0.38	21.8 ± 2.20	88.4	31.8
Ter-lockF	NA	35.7 ± 4.25	NA	19.4
TerG	2.63 ± 0.31	4.4 ± 0.1	16.7	157.4
Ter-lockG	0.94 ± 0.06	0.3 ± 0.02	3.25	2276.5
TerH	NA	22 ± 1.2	NA	31.6
	4.10 ± 0.44	0.56 ± 0.04	1.36	1244.2
Ter-lockH	NA	NA	NA	59.1*
	5.04 ± 0.35	1.68 ± 0.09	3.34	411.7
TerI	NA	24.1 ± 0.4	31	28.7
	5.54 ± 1.01	0.73 ± 0.08	1.32	949.1
Ter-lockI	NA	3.53 ± 0.4	18.9	196.4
	3.17 ± 0.44	0.65 ± 0.06	2.05	1069
TerJ	NA	25.4 ± 0.4	NA	27.3
	3.20 ± 0.36	1.94 ± 0.10	6.07	356.8
Ter-lockJ	NA	2.1 ± 0.18	NA	332.1
	1.25 ± 0.2	0.52 ± 0.05	4.18	1322

Shaded rows represent 150 mM KCl data while clear rows correspond to 250 mM KCl data. NA: Data not available.
* $t_{1/2}$ value was determined by direct visual analysis.

dextran and biotinylated abasic linker in the other. Indeed, for all *Ter* sites, the k_a values were similar for the double stranded and *Ter-lock* oligonucleotides suggesting that access to all surface immobilized ligands by the analyte was affected by the surface chemistry in a similar fashion. The data are therefore directly comparable and further supported by the GFP-Basta data.

Confirming the results obtained with GFP-Basta, Tus did not bind to *TerF* nor *Ter-lockF* in 250 mM KCl in the range of concentrations tested (Figure 16B). The SPR data for *TerA-J* correlated well with the stability profile obtained with GFP-Basta. All strong binders in their *Ter-lock* configuration were able to "lock" the Tus protein as demonstrated by a dramatic increase in their $t_{1/2}$ compared with their double-stranded analogues (Figure 16B). The sensorgrams for Tus binding to the moderate binders *TerH-J* were in agreement with the GFP-Basta data demonstrating that the weaker binding of these species is mainly due to a shorter $t_{1/2}$ compared to those obtained for the strong binders. The sensorgrams obtained with *Ter-lockH-J* revealed the inability of *TerH* to form a strong TT-lock in 250 mM KCl (Figure 16B). It also revealed that the $t_{1/2}$ obtained with *Ter-lockI* and *J* were similar to those obtained for Tus in complex with *TerA-E* and *G*. The *Ter-lockI-J* were able to induce a 10-15-fold longer $t_{1/2}$ than their respective *TerI-J*, demonstrating that they could form a locked complex with Tus. In lower-salt conditions, sensorgrams obtained for strong *Ter* and *Ter-lock* sites could not be fitted accurately, but were still of great value as they showed that the strong *Ter-lock* sites induced the slowest dissociation of Tus (Figure 16C). Interestingly, the $t_{1/2}$ of Tus for *TerH* was longer than for *Ter-lockH*, *TerI* exhibited the same $t_{1/2}$ as *Ter-lockI* for Tus, and *Ter-lockJ* had a longer $t_{1/2}$ than *TerJ* for Tus. Finally, *TerF* and *Ter-lockF* bound weakly to Tus in lower-salt conditions, with both exhibiting low affinity and very short $t_{1/2}$. This is in agreement with the data obtained for these species by GFP-Basta (Figure 15C-D) suggesting that binding of Tus to these species is only marginally more specific than to the non-specific

oriC and *oriC-lock*. The SPR data also revealed that the higher stability of the Tus-*Ter* complexes over the Tus-*Ter-lock* complexes previously observed with GFP-Basta in 150 mM KCl (Figure 15B-C) was due to a greater effect of the ionic strength on the $t_{1/2}$ of Tus-*Ter* than on their respective Tus-*Ter-lock* complexes (Figure 16B-C). This conclusion is in contrast with a previous study showing that the k_a was mainly affected by the ionic strength of the buffer for *TerB* (Neylon et al., 2000). Although not directly comparable, the data presented here suggest that for *Ter* species both k_a and $t_{1/2}$ are significantly affected by an increase in ionic strength. This of course does not mean that the lock is lost in low-salt conditions but rather reflects the large contribution of cooperative electrostatic interactions to the Tus-*Ter* and Tus-*Ter-lock* complex and reflects also the importance of the specific interaction of R198 with A(5) and G(6) which cannot occur with our synthetic *Ter-locks*.

4.3.3 F140 interaction with T(7)

In the crystal structure of Tus-*Ter-lock* structure (PDB ID 2OI6; Mulcair et al., 2006) it was observed that the T(7) base is stacked against the F140 phenyl ring without forming base specific contact. Crosslinking studies aiming at developing covalently linked protein-DNA conjuguates (Dahdah et al., 2009, Schaeffer and Dixon, 2009) showed that a *Ter*-lock oligonucleotide containing a bromodeoxyuridine (BrdU) at position 7 in the core sequence could be photo-crosslinked with high yield with Tus upon photochemical activation. Crosslinking was also obtained with fully double stranded *Ter* site though in much lower yields and slower kinetics. It was proposed that photoactivation produced a uridyl radical that could form a covalent carbon-carbon (C-C) bond with the neighbouring electron rich phenyl ring of F140 (Schaeffer and Dixon, 2009). Since F140 does not contact fully double stranded *Ter* directly in the crystal structure (PDB ID 1ECR), the low crosslinking yield obtained with this species was expected to involve a different mechanism and residue. In order to verify that

crosslinking was indeed occurring between base 7 and the F140 phenyl ring, the photo-crosslinking reactivity of Tus mutant F140A was examined for both *Ter* (P1) and *Ter*-lock species (P2, P3). It was expected that the low yielding photo-crosslinking reaction with double stranded *Ter* would still occur in the F140A mutant for the above mentioned reason.
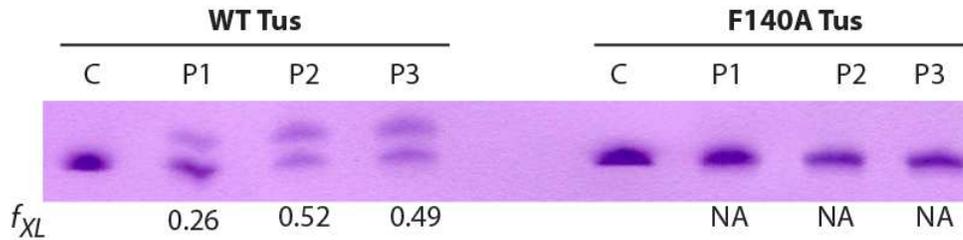
Figure 17: Crosslinking of wild-type and F140A Tus with *TerB* or *Ter-loc*kB. P1: fully ds*Ter* with BrdU substitution at position 7. P2: *Ter-lockB* variant of P1 with a mismatched BrdU-T(7). P3: *Ter-lockB* variant of P1 with a BrdU-A(7) base pair. Fractions of crosslinked protein-DNA complexes (f_{xl}) are indicated below the bands. f_{xl} for Tus with P1, P2 and P3 were (mean ± SEM) 0.26 ± 0.01, 0.52 ± 0.023 and 0.49 ± 0.023 respectively (data obtained in triplicates). No crosslinking was observed with the F140A Tus mutant.

As expected, F140A did not cross-link with the TT-Lock variants P2 and P3, which confirms that F140 is essential for the reaction to occur (Figure 17). However, Tus F140A crosslinking was also abolished with *TerB* (P1) and *Ter-lockB* with paired T(7) (P2), suggesting that it is the same phenylalanine that is involved in the C-C bond formation. In the Tus-*TerA* structure, F140 and A(7) are about 9 Å apart implying that breathing motions of protein or DNA occur with sufficient frequency (i.e. transient interaction) to allow these groups to interact. These results suggest that the helix α4 containing F140 has significant flexibility. This is in agreement with a molecular dynamic simulation of Tus binding to *TerB* which revealed potential transient protein-DNA interactions and significant flexibility of Tus-*Ter* and TT-lock complexes. (Oakley A. et al., in preparation).

4.4 Discussion

4.4.1 *TerF* and *TerH* are unable to form a significant TT-lock

The SPR analysis revealed that *TerF* and *TerH* were unable to form an obvious TT-lock and that binding of Tus to the *TerF* could only be achieved in the low-salt conditions that better reflect the physiological conditions found in the bacteria (Figure 16C). *TerF* has been identified by searching the *E. coli* genome for consensus *Ter* sequences (Sharma and Hill, 1992). Initially, the affinity of *TerF* for Tus was overestimated ~ 50-fold because the GC base pair at position 18 was replaced by TA (Coskun-Ari and Hill, 1997). *TerF* was found to have only ~5% fork pausing efficiency compared to 35% for *TerB* in a plasmid context (Duggin and Bell, 2009). Using SPR, a K_D of 8.8 x 10^{-8} M was obtained in low-salt, and no binding was detected at 250 mM KCl (Figure 16B-C). In addition to its low binding affinity, *TerF* is also unable to form a TT-lock. This result could explain why *TerF* cannot efficiently pause replication forks *in vivo* although the more distal *TerG* can (Duggin and Bell, 2009). In the same study *TerF* was only able to induce fork pausing when Tus was overexpressed to ~5% of total cellular protein content reflecting its weak affinity for Tus (Duggin and Bell, 2009). Conversely *TerH* ($t_{1/2} = 32$ s in Figure 16B), which is unable to form a significant TT-lock ($t_{1/2} = 59$ s in Figure 16B), was categorised as a moderate binder – i.e. binding to Tus can be observed in 150 mM and 250 mM KCl – with a K_D of 1.4 x 10^{-9} M in lower-salt. Interestingly, *TerH* was still found to pause forks with 12.5 % efficiency (Duggin and Bell, 2009). Thus, formation of the TT-lock, although clearly important, is not the only factor controlling polarity of fork arrest by Tus. The fork pausing activity of *TerH* is therefore believed to be the result of the remaining (and substantial) resistance of Tus to dissociation when forks approach the non-permissive face in the absence of a TT-lock, as shown by Mulcair et al. (2006).

4.4.2 Importance of non-conserved bases for Tus-*Ter* binding and TT-lock formation

To clarify the understanding of the binding of Tus to *Ter* or *Ter-lock* sites, the data obtained here and all other available base substitution data and their effect on Tus-binding were summarised in Table 13. The strong binders *TerA-E* and *G* have very similar K_D and kinetic values in high salt – i.e. K_D of 10-25 nM and $t_{1/2}$ of 140-380 s (Figure 16B). The *Ter-lockA-E* and *G* were also very similar with again only little more than 2-fold difference between the strongest and the weakest complex of this group – i.e. K_D of 1-3 nM and $t_{1/2}$ of 2280-4370 s (Figure 16).

Table 13: Effect of base substitutions on Tus-*Ter* binding and TT-lock formation.

Ter Site	Base							K_D (nM)	*Ter-lock* $t_{1/2}$ (s) [d]
A	(A)4	(T)5	(A)7	(A)9	(A)18	(T)20	(T)21	10.7	3310
B	(T) =							11.6	4367
C			(C) =[a]				(A) =[a]	18	3408
D								17	2589
E	(T) =						(C) 4X[a]	26	2979
F		(G)2X[b] 2X[c]			(C) 60X[a]	(G) 7.5X[a]	(C) 4X[a]	NB	NB
G	(T) =		(C) =[a]				(G) 10X[a]	17	2277
H		(G)2X[b] 2X[c]				(A) 3X[a]	(G) 10X[a]	37	~59
I	(T) =	(A)[e]	(C) =[a]	(T) 25X[a]			(G) 10X[a]	31	196
J	(G)2X[b]			(T) 25X[a]			(A) =[a]	240	332
K	(A)[e]	(A)[e]	(T)5X[a]	(T) 25X[a]	(C) 60X[a]		(C) 4X[a]	–	–
L	(G)2X[b]	(A)[e]	(C) =[a]		(C) 60X[a]	(G) 7.5X[a]	(C) 4X[a]	–	–
Y	(C)[e]	(C)[e]		(G)23X[a]		(C) 18X[a]	(C) 4X[a]	–	–
Z	(G)2X[b]	(G)2X[b] 2X[c]	(G) [e]	(C)47X[a]	(C) 60X[a]	(C) 18X[a]		–	–

[a]Increase in K_{obs} compared to *TerB* in potassium glutamate (Coskun-Ari and Hill, 1997). [b]Reduction in k_a in 250 mM KCl (Mulcair et al., 2006). [c]Reduction in $t_{1/2}$ in 250 mM KCl (Mulcair et al., 2006). [d]Dissociation $t_{1/2}$ obtained in 250 mM KCl from Figure 16B. [e]Base substitution data is not available. = binding is unchanged. *TerK,L* and *TerY* have a further substitution at position 17 that affects their binding by 17-fold and *TerK and TerL* have both a substitution at position 8 affecting their binding by 15-fold[a].

These data suggest that bases 1-4, 7 and 21-23 contribute little to the affinity and kinetics of the Tus-*Ter* or to the formation of the TT-lock complex in 250 mM KCl but their effects increase in lower salt. Substitution of base T(21) in *TerA* by a C or a G has been shown to increase the K_D by 4- and 10-fold respectively, but no change was observed if substituted to an A (Coskun-Ari and Hill, 1997) in high-affinity conditions. The affinity of Tus for *TerF,H-J* was predicted to be much weaker than for *TerA* due to the presence of single or multiple substitutions in the bases 9 (*TerI-J*), 18 (*TerF*), 20 (*TerF,H*) and 21 (*TerF,H*) that have been shown to significantly weaken the stability of the Tus-*Ter* complex (Table 13). The SPR data for *TerF* and *H* correlate well with these earlier findings. As expected *TerF* was found to be the weakest of all *Ter* sites and *TerH* was the strongest of the moderate binders (Figure 16B-C and Table 13). Interestingly, although *TerJ* was expected to be the strongest of the moderate binders because it has fewer base substitutions than *TerH-I*, i.e. only A(9) to T (Table 13), it was systematically found to be the weakest of this group. This substitution is most likely responsible for the overall lower affinity of Tus to *TerI* and *TerJ*, affecting both on- and off-rates. The data obtained with the weak and moderate binders were most valuable to refine the understanding of the TT-lock as a base substitution affecting its formation was identified. Indeed, within these *Ter* sites, the non-TT-lock forming *Ter-lockH* and *Ter-lockF* sites, are the only *Ter-lock* sites with a G at position 5 instead of the T present in all other strong TT-lock forming sites (Table 13). It has previously been shown that substitution of T(5) by G affected the binding of a *Ter-lock* (*cf* F5-*TerB*(G5), Figure 3 in Mulcair et al., 2006) to Tus by more than two-fold reduction in $t_{1/2}$ and four-fold increase in K_D, suggesting the presence of some base-dependent interactions or hindrances occurring at this position, albeit not being obviously important in either Tus crystal structures (Kamada et al., 1996, Mulcair et al., 2006). Although these changes are quite modest, the cumulative effects due to the presence of additional non-optimal substitutions in preceding bases (1-4) might further

affect TT-lock formation as a result of cumulative steric and/or electrostatic hindrances. *Ter-lockJ* produces the strongest TT-lock of the moderate binders probably because it has the conserved T(5) found in all strong binders.

Interestingly, *TerI* is the only *Ter* site to have an A at position 5 instead of the canonic T(5), but this does not seem to significantly affect TT-lock formation with *Ter-lockI*. The highly conserved T(5) within the strong binders is obviously important for TT-lock formation following strand separation at the non-permissive face of the complex and might help C(6) to better dock into its position, i.e. in the specific cytosine binding pocket at the surface of Tus. Furthermore, in the Tus-*TerA* structure the N3 of A(5) is in contact with R198 in the Tus protein which also contacts the N3 of G(6) (Coskun-Ari and Hill, 1997, Kamada et al., 1996). The mutation R198A resulted in a 130-fold increase in K_D mostly due to a ~50-fold decrease in k_a in 250 mM KCl, as well as a 5-8-fold increase in K_D in low-salt conditions indicating the importance of R198 for *Ter* binding (Neylon et al., 2000). Indeed, R198 is largely responsible for holding the C-domain against the *Ter* site at the non-permissive end of the complex (Kamada et al., 1996, Neylon et al., 2000). The R198 residue could be one of the key residues implicated for scanning of DNA by Tus in the search for a *Ter* site when pushed by the replisome. These are the first specific and non-specific interactions between a base and an amino acid residue that a progressing replication fork will disturb at the non-permissive face of the Tus-*Ter* complex. The A(5) of the T·A(5) base pair interacting with R198 is proposed to play a critical role in Tus-*Ter* complex formation and following strand separation, T(5) strengthens the TT-lock through the formation of additional electrostatic interactions (Figure 18). This is further supported by a similar reduction of $t_{1/2}$ obtained with a single C(6)-overhead *Ter-lock* (*cf* "single O/H C" and F5-*TerB*(G5) in Figure S1 in Mulcair et al., (2006)).

Taken together the quantitative data herein and the base substitution analysis by Coskun-Ari et al. (1997) also provides a simple explanation on why the recently identified *TerK*, *TerL*, *TerY* and *TerZ* are only marginally pausing replication forks. It is obvious from Table 13 that the large number of base substitutions present in their core sequences (*cf TerF* and Figure 14B) would result in a K_D of these species for Tus comparable to that measured for non-specific DNA sequences.

4.4.3 The *Ter* sites and TT-lock formation in the replication termination fork trap

Six *Ter* sites were found to form a strong TT-lock (*TerA*, *TerB*, *TerC*, *TerD*, *TerE*, *TerG*). In the group of moderate binders (*TerH-J*), a significant difference was observed in their ability or not to form a TT-lock (Figure 16). Indeed, upon binding to Tus, *Ter-lockI* and *J* were able to form moderate TT-locks whose $t_{1/2}$ were comparable to the one observed for a strong *Ter*, and *Ter-lockH* did not produce a relevant TT-lock. The GFP-Basta profile (Figure 15) and SPR data (Figure 16) were compared with the plasmid fork pausing data of Duggin and Bell (*cf* Figure 6 p22). It is immediately evident that the affinity and kinetic data obtained for *TerA-J* do not fit with the fork pausing efficiency profile obtained by Duggin and Bell but the data obtained for the *Ter-lock* fits well (Figure 15D,F and Figure 6 p22). Indeed, taking only the three last *TerH-J* sites into account, *TerJ* is the weakest binder of this group (Figure 15C,E and Figure 16B) but comparatively the strongest TT-lock-forming site (Figure 16B). These findings correlate perfectly with the higher fork pausing efficiency observed for *TerJ* compared with *TerH-I* (Duggin and Bell, 2009). Thus, fork pausing efficiency data obtained by Duggin and Bell is best explained by the formation of the TT-lock *in vivo*.

The results presented here provide essential information about the efficiency of binding, strength of the TT-lock and the importance of bases in the *Ter* and *Ter-lock*

sequences, but they do not necessarily reflect the chromosomal situation, which is influenced by the location of *Ter* sites, Tus occupancy and the frequency of forks approaching each side of the *Ter*-Tus complex. Duggin and Bell observed significant pausing at *TerA* (0.19%), *TerB* (0.14%) and *TerC* (0.85%) in the chromosomal wild type context (Duggin and Bell, 2009), in accord with the thermodynamic and kinetic data for these sites and their ability to form a strong TT-lock (Figure 15 and Figure 16). In principle, overexpression of Tus should result in higher Tus occupancy on *Ter* sites if they are not already fully occupied, creating an even tighter fork trap. The innermost *TerA* and *TerC* were expected to be fully occupied, but upon overexpression of Tus, fork pausing increased significantly at *TerA* (0.64%) and to a small extent at *TerC* (1.01%; Duggin, 2006). Fork pausing at *TerB* (0.12%), which can only occur if the replisome breaks through *TerC* (~15% of the time), was unaltered and consistent with the notion that *TerC* is already fully occupied by Tus. Interestingly, they observed only weak or no pausing at the remaining strong T*erD,E* and *G* (Duggin and Bell, 2009).

Taken together, *TerA-D* are the strongest and innermost *Ter* sites of the fork trap able to form the tightest TT-locks. They are located in the central part of the termination region. Within these sites, *TerB* and *C* are the strongest TT-lock forming sites; they are both located in the cluster as the two first sites capable of blocking clockwise moving forks. Not surprisingly, *TerA* and *TerD* are also located in a similar configuration in the opposite cluster stopping anti-clockwise moving replication forks. Considering *TerF* is probably not involved in replication fork arrest, the two remaining strong *TerE* and *G* and the moderate *TerH-I* and *J* are positioned towards the middle and the extremities of each cluster respectively with rather intriguing symmetry. Thus, with the exception of the weak *TerF*, the more distal the *Ter* sites are from the centre of the termination region, the weaker their binding with Tus. Indeed, it is quite striking that *TerH,* which forms the weakest TT-lock, and *TerJ,* which is the weakest *Ter* site, are the outermost *Ter* sites in each cluster. This could suggest that the

more distal *Ter* sites, which are rarely used in fork arrest, have probably devolved from their original function as more *Ter* sites appeared in the chromosome during evolution. The data obtained by Duggin and Bell seem to support this theory (Duggin and Bell, 2009).

4.4.4 Fork arrest: a three-step model

Based on the above analysis, it appears that a series of essential steps are required to arrest fork progression, including the non-specific binding of Tus to DNA followed by its correct docking to a strong *Ter* site and finally, the formation of a strong TT-lock induced by the unzipping action of DnaB helicase (Figure 18A-C). Here, the formation of the TT-lock involves the proper docking of C(6) in the C(6) binding pocket of Tus, which is dependent on the helicase activity of DnaB, i.e. the rate of unwinding. Recently, Bastia et al. showed that DnaB helicase could translocate over short stretches of double stranded DNA and that removal of Tus from a *Ter* site was easier when DnaB was moving towards the permissive face than the non-permissive face, and proposed that formation of the TT-lock is only a failsafe mechanism (Bastia et al., 2008). These findings could also suggest that when Tus docks on its *Ter* site it functions as a linear ratchet on the DNA to resist the "pushing action" from DNA binding proteins such as RNA polymerases (Mohanty et al., 1998) – i.e. Tus can be pushed and dislodged from the permissive face but less from the non-permissive face (Figure 18A-B).
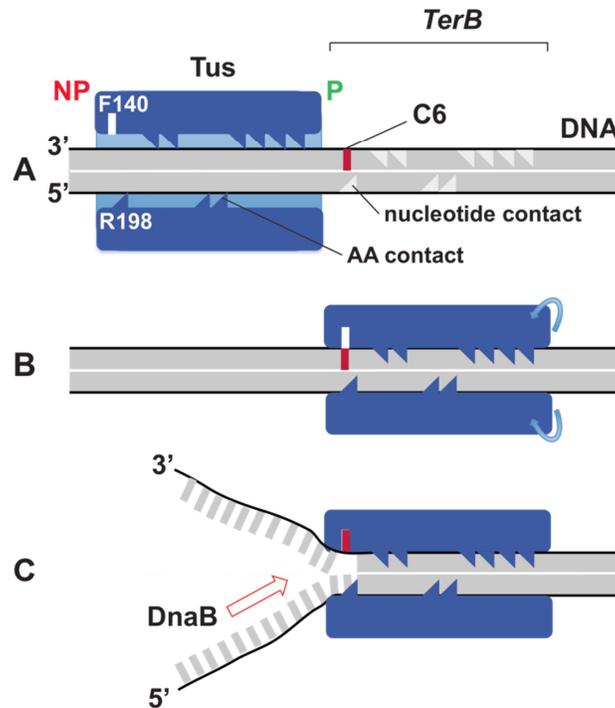
Figure 18: A three-step model for fork-arrest. (A)Non-specific binding of Tus to DNA mediated by cooperative electrostatic interactions allows sliding. (B) Proper docking of Tus to its *Ter* site upon correct alignment of nucleotide and amino acid (AA) contacts results in a linear ratchet. (C) Unzipping of the DNA by the action of DnaB at the non-permissive (NP) face leads to formation of the TT-lock through docking of C(6) in the C(6) binding-pocket of Tus.

The fact that all functional *Ter* sites with the exception of the outermost *TerH* were able to form a TT-lock and that C(6) and the C(6) binding-pocket have been maintained during evolution (see multi-alignment in Figure 19) demonstrates the biological importance of this dynamic process.

Finally, if we assume that fork pausing at *TerB* is the result of a fork breaking through *TerC* and that break-through is mostly due to failure to form a TT-lock, then Duggin and Bell's chromosomal data would suggest that, for *TerC*, the TT-lock fails to form only 15% of the time.

A

```
1ECR_A      88 SKAAVRLPGVLCYQVDNlsQa--ALVSHIQHINKLKTTFEHIVTVESElPTAARFEWVHRHLPGLITLNAYRTLTVLHDP 165
gi 166229880 88 SKAAVRLPGAVCFSVDD--AqheALHALADKINALKTQLEEIITVESGlPSTQRFEWVHRYLPGLITLNAYRRLTVLDNP 165
gi 166229881 88 SKAAVRLPGAICLQVTA--AeqqDLLARIQHINALKATFEKIVTVDSGlPPTARFEWVHRHLPGLITLSAYRTLTPLVDP 165
gi 166991204 88 SKAAVRLPGVICLQCDG--AlreEIESQIAHINTLKMAFEKIITEESGlPPAARFEWVHRQLPGLITLNAYRSLTVLRQP 165
gi 88943805  88 SKAAVRLPGVLCYQVDN--LsqaALVSHIQHINKLKTTFEHIVTVESElPTAARFEWVHRHLPGLITLNAYRSLTVLHDP 165
gi 81360679  88 SKAAVRLPGVLCYQVDN--AtqlDLENQIQRINQLKTTFEQMVTVESGlPSAARFEWVHRHLPGLITLNAYRTLTLINNP 165
gi 123519298 88 SKAAIRLPGALCFAVDN--DqyaAAVSLISEINDLKTHLSRLITVESQlPPEQRFDFVHRHLRGLKTLSAYRGLQVISDP 165
gi 81645039  88 SKSAIRLPGALCFAVNA--AqyqQAHTLIAEINALKKELEQIITVESGlEPEQRFEFVHSHLKGLITLSAYRSLTLITNP 165
gi 166991205 88 SKAALRLPGVLCFSVTD--KeliECQDIIKKTNQLKAELEHIITVESGlPSEQRFEFVHTHLHGLITLNTYRTITPLINP 165
gi 81419370  88 SKAAVRLPGVLCFSVNQ--QqyrSALLLIEEINKLKAELEHIVTVQSGvDKEQRFEFVHSQLHGLITLNAYRTITPLLNP 165
```

B

```
1ECR_A       85 -----NRSSKAAVRLPGVLCYQVDNl-sQa--ALVSHIQHINKLKTTFEHIVTVESElPTAARFEWVHRHLPGLITLNA 155
gi 75488391  75 -----DFSQKSARRASGVLWFPSED---NaftvELVRLLSQINALKKSIETHIITTYQ-TRSARFEALHNQCAGVLTLHL 145
gi 75486913  75 -----GLSQKSTRRTAGVLWYDDET---PefaaEVMALITSINAIKLSIQEHIITGFS-SQNARFSALHNACPGVMTLHL 145
gi 123068549 89 -----HDSTKVTTKYPGLIIVPERA---D----LLESLITSINEAKNDFAAAMRRIDN-KKNVRFDKVHKKLPGLVAMHS 155
gi 88857234  79 -----DISGRVLTRHPGILAVNTDN---Pd----ALIERINQINKAKVDFKNLILGLGN--NDARFNAVHEAIPNVITLAV 145
gi 118602013 78 -----DYSQKSARRTVGVLWFSPSRigvAd---EIAATVERINAAKAGIEEFIISTYP-TRQERFEALRADCPGVMTLHL 148
gi 149910878 114 -----PFSQKSARRYVGIIHLHPEN---Qns-aKIVTSIEDINTIKGAIEHAVVSENE-SRQARFSAIHDNCPGVMTVHL 183
gi 126668475 86 -----QYSTKVVNRVPGALVLANVD---Es---QVLDRMAEINRSKTAIESLIPSLGN--RDDRFELLHGHFPWLILAQL 152
gi 10957374  78 idvlsRNSTKVTTKYPGLIIVPEGA---Eq----LSGIICDINDAKSDFASAMKRVNS-EKHIRFKEVHRKLPGLVTAHS 149
gi 146343954 84 -----TQNAKETLRSPGVVGLPR---------EVIDKIIETNELRSKLEMLIGKIKK-PKDRRM--VWRKFEAISPKQA 145
```

Figure 19: Multi-alignment of Tus with Tus from diverse bacteria. The alignment was generated with Tus from *E.coli* (1ECR_A) as the query sequence using the Conserved Domains Database (accessed from the NCBI website)[4]. (A) Top listed sequences (most conserved sequences excluding *E.coli* strains). Bacteria include: *Cronobacter sakazakii* ATCC BAA-894, *Klebsiella pneumoniae* subsp. *pneumoniae* MGH 78578, *Enterobacter sp*. 638, *Shigella dysenteriae* Sd197, *Salmonella enterica* subsp. *enterica* serovar Paratyphi A str. ATCC 9150, Sodalis glossinidius str. 'morsitans', Pectobacterium atrosepticum, *Yersinia pseudotuberculosis* IP 31758, *Photorhabdus luminescens* subsp. *laumondii*. (B). Most diverse sequences included: IncT plasmid R394, *Proteus vulgaris*, *Pseudoalteromonas tunicata* D2, *Photobacterium damselae* subsp. *Piscicida*, *Moritella* sp. PE36, *Marinobacter sp*. ELB17, *Salmonella enterica* subsp. *enterica serovar* Typhi, *Pseudomonas fluorescens* SBW25. The shaded box highlights residues F140-G149 in the C(6) binding pocket domain. Residues with direct contacts to the C(6) are asterisked.

4.5 Conclusion

The quantitative analysis of the ten *Ter* sites and their respective *Ter-locks* allowed the definitive classification of *Ter* sites from strong to weak Tus-binders and provided essential information on base-residue interactions – i.e. their importance for formation of Tus-*Ter* complexes and TT-lock. Within the ten *Ter* sites studied here only *TerF* was found to be too

---

[4] A. Marchler-Bauer, S. Lu, J. B. Anderson, F. Chitsaz, M. K. Derbyshire, C. DeWeese-Scott, J. H. Fong, L. Y. Geer, R. C. Geer, N. R. Gonzales, M. Gwadz, D. I. Hurwitz, J. D. Jackson, Z. Ke, C. J. Lanczycki, F. Lu, G. H. Marchler, M. Mullokandov, M. V. Omelchenko, C. L. Robertson, J. S. Song, N. Thanki, R. A. Yamashita, D. Zhang, N. Zhang, C. Zheng and S. H. Bryant, Nucleic Acids Res., 2011, 39, D225-229.

weak to possibly be involved in fork pausing. The data demonstrate that Tus's resistance to dissociation mediated by the formation of a lock, whether weak or strong, correlates best with fork pausing *in vivo*. Further studies providing structural information on unbound Tus will be required to fully describe the complex dynamics of this system. The location of Tus and how it finds a *Ter* site remain also mysteries, which could be solved by a combination of genomewide localization, single-molecule and *in vivo* imaging studies. Finally, recent applications of Tus in proteomics (Askin et al., 2011, Chatterjee et al., 2008, Kaczmarczyk et al., 2010, Moreau et al., 2010, Sitaraman and Chatterjee, 2011) and its potential use as a connector between DNA and antitarget proteins in multiplex immuno-PCR diagnostics (Dahdah et al., 2009, Morin et al., 2011, Morin et al., 2010, Schaeffer and Dixon, 2009) is currently driving the search for even stronger *Ter* and *Ter-lock* sequences capable of never dissociating from Tus. The methods and data described herein will undoubtedly be invaluable for this purpose.

Chapter 5: Salt dependence of the Tus-*Ter* complex by differential scanning fluorimetry of GFP-tagged proteins (DSF-GTP)

Parts of this chapter are taken from a previously published manuscript in RSC Advances (Moreau, M. J., Morin, I. Askin, S. P. Cooper, A, Moreland. N. J,. Vasudevan, S.G. & Schaeffer, P. M) [5].

5.1 Introduction

The affinity and stability of a particular protein-DNA complex are usually characterized by their binding constant and Gibbs free energy of binding, but these parameters are not sufficient for elucidating the nature of the physical forces acting between the two molecules. Ionic contacts play a major role in the interaction between highly charged DNA and protein DNA-binding domains. The fine tuning of highly specific protein-DNA interactions can be further understood by separating the overall binding energy into its electrostatic and non-electrostatic components. The relative magnitude and importance of each component modulates the functional specificity and activity of a protein at various binding sites (Minton,

---

[5] The validation described in the introduction of this chapter has been published in Moreau, M. J., Morin, I. Askin, S. P. Cooper, A, Moreland. N. J., Vasudevan, S.G. & Schaeffer, P. M. 2012. Rapid Determination of Protein Stability and Ligand Binding by Differential Scanning Fluorimetry of GFP-Tagged Proteins**,** RSC Advances, 2, 11892-11900.

2001). The electrostatic component of an interaction results solely from the entropy of mixing the displaced DNA counter-ions with ions in bulk solution according to the counter-ion condensation (CC) concept (Manning, 1978, Record et al., 1978, Record et al., 1991, Record et al., 1976, Waldron et al., 2005). The contribution of electrostatic interactions to complex formation can therefore be determined from the salt dependence of the association equilibrium constant. Salt resistance of protein-DNA complex indicates a dominant contribution of non-ionic interactions.

The *E. coli* termination protein Tus binds to 14 termination sites (*Ter*) scattered on the chromosome. Ten of these sites (primary *Ter* sites *TerA-J*) were characterized in terms of binding affinity by SPR at two salt concentrations (*cf* Chapter 4, Figure 16 p90; Moreau and Schaeffer, 2012a). The salt dependence of Tus interacting with *TerB* has previously been studied (Neylon et al., 2000) and revealed a large effect of salt concentration change on association kinetics. However, there is no study on the salt dependence of Tus in complex with the other *Ter* sites or with their TT-lock analogue. Notably, *TerF* was shown to be only marginally more specific than a non-specific *oriC* DNA fragment (*cf* Chapter 4, Figure 15 p86; Moreau and Schaeffer, 2012a). Kinetic studies showed that *TerF* bound weakly to Tus at 150 mM KCl but not at 250 mM KCl (*cf* Figure 16 p90), suggesting that *TerF* forms mainly electrostatic interactions with Tus (Moreau and Schaeffer, 2012a).

The recent development of differential scanning fluorimetry (DSF) and the high-throughput (HT) capability of the protein stability and ligand binding assay Thermofluor have vastly facilitated research in the field of macromolecular interactions (Ericsson et al., 2006, Lavinder et al., 2009, Magliery et al., 2011, Niesen et al., 2007, Pantoliano et al., 2001, Senisterra and Finerty, 2009, Vedadi et al., 2006). The previously developed method GFP-Basta (*cf* Chapter 3 p60) was developed as an alternative method to DSF since it could be used with samples containing substantial amounts of additional proteins, i.e. ligands or

contaminants (Moreau et al., 2010). However the high throughput potential of GFP-Basta to generate large amount of data on protein stability suffers from the fact that it requires a separation step. Although this issue could be alleviated by using adequate robotics, a higher throughput method was developed to study the salt-resistance of all Tus-*Ter* complexes.

It was noticed during the development of GFP-Basta that the initial fluorescence intensity of *TerB*-bound Tus-GFP was systematically increased compared to free Tus-GFP at the same concentration (unpublished data). As both protein domains unfold independently in the fusion protein (*cf* Figure 9 p63; Moreau et al., 2010), this phenomenon could only be the result of changes in the proximal environment of the GFP upon binding of *TerB* to the Tus domain, i.e. GFP acts as a sensor and reporter of its proximal environment. In this case, a change in fluorescence should also be measurable in real-time by DSF when Tus unfolds and aggregates (**Error! Reference source not found.**A).

To test this hypothesis, a solution of Tus-GFP or GFP (control) was gradually heated using the melting curve protocol of a real-time thermal cycler and a transition in the melting curve profile was clearly observed for Tus-GFP before loss of fluorescence of GFP at ~80 °C. This transition was absent in the profile obtained for GFP alone (control) confirming that the effect was induced by the Tus domain in Tus-GFP. The curves were mathematically transformed to the first derivative (**Error! Reference source not found.**B) resulting in the transition in the melting curve being visualized as a peak with its tip representing the transition midpoint (T_m). It was shown that the T_m peak was indeed a reflection of the aggregation stage of the Tus-GFP by measuring the residual fraction of folded proteins (F_{fold}) present in reactions stopped before, at the maximum and at the end of the peak using GFP-Basta (**Error! Reference source not found.**C). The F_{fold} values at these temperatures indicated that the protein is unfolding and aggregating at the peak obtained with DSF-GTP.

This confirmed that the peak and associated T_m corresponded to the transition midpoint of aggregation (T_{agg}) of Tus-GFP (**Error! Reference source not found.**C).
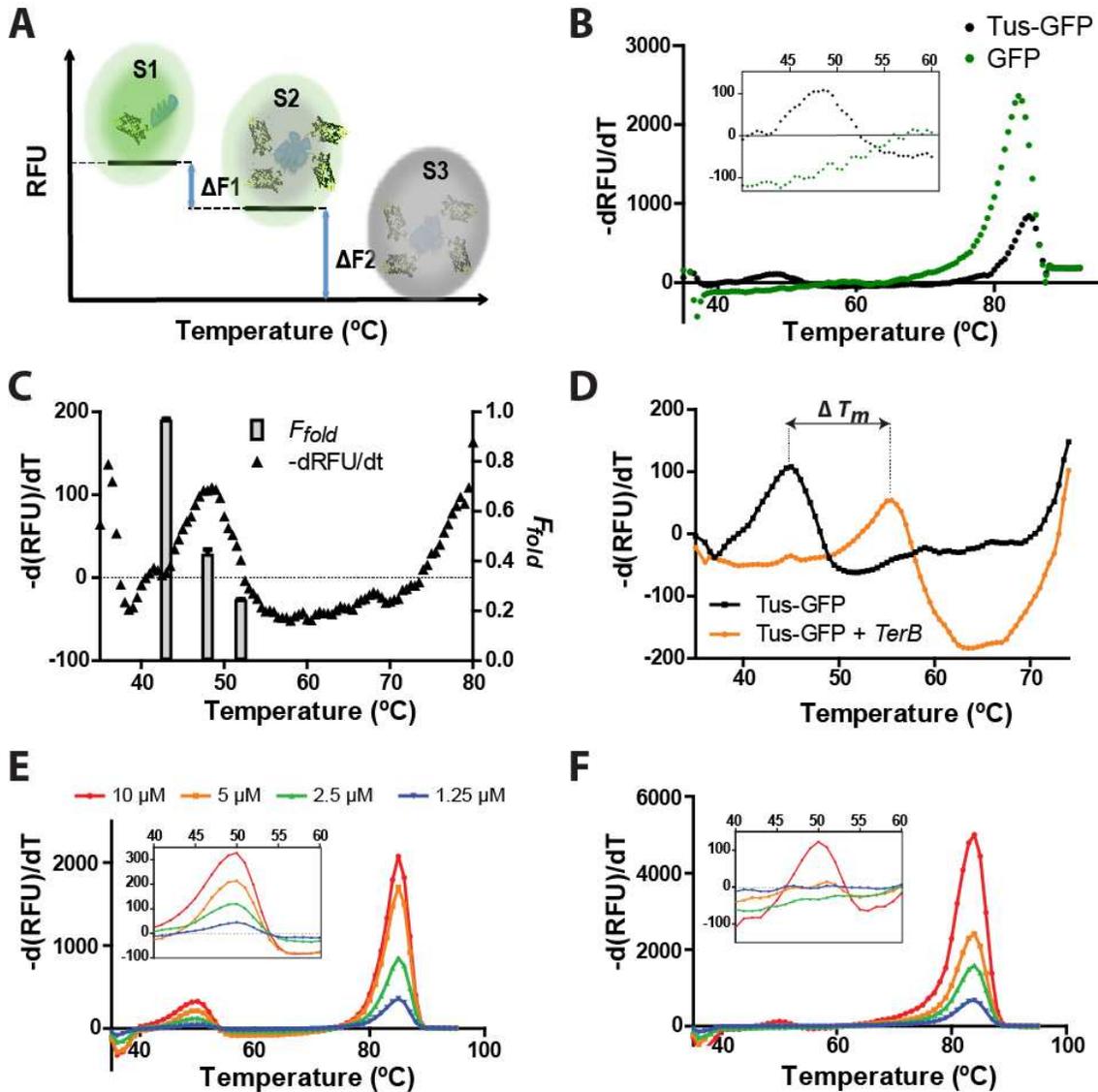
Figure 20: Concept and validation of DSF-GTP. The fluorescence of a POI-GFP (protein of interest fused to GFP) is monitored in real-time over a temperature range. At low temperature POI-GFP is in the fully folded state S1. POI-GFP switches to the lower fluorescent state S2 when POI unfolds and aggregates. Fluorescence is lost when GFP unfolds (S3 state). ΔF: difference in fluorescence. (B) Melting curves obtained with Tus-GFP and GFP at 2.5 µM each in phosphate buffer (45 mM Na$_2$HPO$_4$, 5 mM NaH$_2$PO$_4$, 10% glycerol (v/v), 2 mM β-mercaptoethanol) (0.5°C/cycle, 10 s dwell time). (C) Correlation between the T_m obtained with DSF-GTP and the residual fraction of folded Tus-GFP (F_{fold}) at the temperature immediately before, at the midpoint, and at the end of the T_m peak determined by DSF-GTP. (D) Thermal shift of Tus-GFP upon binding to *TerB* in SPR$_{250}$ buffer (0.5°C/cycle, 30 s dwell time). (E) Concentration effect on the melting curve of Tus-GFP in phosphate buffer (1°C/cycle, 10 s dwell time). (F) Melting curve

of a mixture of equimolar Tus and GFP in phosphate buffer at concentrations from 1.25 to 10 μM. The inset is a close up look of the peak obtained for Tus. Melt curve settings and curves color match pannel (E).

The phenomenon was best explained by a fluorescence quenching mechanism resulting from shielding of the fluorophore by the proximity of the Tus-aggregates. This was further supported by the fact that a small T_m peak was apparent for a mixture of Tus and GFP at 10 μM each corresponding to the T_{agg} of Tus-GFP (**Error! Reference source not found.**F). Thus, the fluorescence quenching is highly enhanced by the physical linking of Tus with GFP as it becomes only apparent at the highest concentrations tested for a mixture of Tus and GFP (**Error! Reference source not found.**E-F).

As expected, the binding of *TerB* to Tus-GFP at 250 mM KCl resulted in a significant shift in T_m of about 10°C (**Error! Reference source not found.**D). The analysis of the melting temperature dependence on the protein and ligand concentrations has been used to determine dissociations constants (Zubriene et al., 2009) and the concentration dependence of *TerB* on T_m of Tus-GFP was investigated using DSF-GTP (Moreau and Schaeffer, 2012a). The so obtained K_D value of 9 nM in SPR$_{250}$ buffer (pH 7.5) correlated well with the K_D of 11 nM obtained by SPR in the same conditions demonstrating that DSF-GTP can be used to quantitatively monitor protein-DNA interactions. The method was also validated with eleven other proteins for which large quantities of protein stability data as well as ligand-induced stabilization effects were generated (Moreau et al., 2012). DSF-GTP could therefore be used to determine the specificity of each *Ter* and *Ter*-lock for Tus by comparing the contribution of electrostatic interactions to the stability of the ten Tus-*Ter* and Tus-*Ter-lock* complexes. The effect of increasing potassium chloride on the overall stability of these complexes (T_m) was determined simultaneously and compared to the stability induced by a non-specific DNA.

5.2 Material and Methods

5.2.1 Protein expression and purification

His$_6$-Tus-GFP (refered as Tus-GFP) was expressed and purified as previously described (Dahdah et al., 2009) except that here proteins were resuspended in SPR$_{250}$ buffer (50 mM Tris pH 7.6, 250 mM KCl, 0.1 mM EDTA and 0.2 mM β-mercapoethanol) and dialysed twice against the same buffer at 4°C.

5.2.2 DSF-GTP

Melting curves were obtained with an IQ5 iCycler (Bio-rad). The temperature range was 35-75°C at 0.5 or 1 °C/cycle and 30 s dwell time. Reactions were performed in triplicate with 60 µl of proteins at 2.5 µM yielding initial RFUs between ~4000-6000. The first derivative of the fluorescence curve was used to determine the T_m at the maximum change in GFP fluorescence by visual inspection of the curves or by automatic peak recognition.

5.2.2.1   Automatic peak recognition using RStudio

The following program was adapted from Thermal Shift Assays – Xtalwiki[6]. After a DSF-GTP run, the raw data were exported to Excel. The RFU and -d(RFU) sheets were each saved as a CSV file. The following script commands RStudio to read the CSV files (the characters in red are to be adapted for each user, run or file name).

raw_data <- read.csv("C:/Documents and settings/path to file/RFUfile.csv")

grad_data <- read.csv("C:/Documents and settings/ path to file /-dRFUfile.csv")

---

[6] http://thermofluor.org/resources/Thermal-Shift-Assays---Xtalwiki.pdf. Retrieved the 1.11.2012

The following script commands RStudio to scale plots of RFUs and –dRFUs on a single graph and choose T_m at the maximum of the derivative function between 35°C and 71.5°C to avoid taking into account the peak corresponding to GFP unfolding. It then generates individual plots for each well in a pdf file.

```
find.tm <- function(temp=temp, I=I, grad=grad, well=well) {
Igrad <- matrix(1:154), nc=2)
Igrad[,1]=I
Igrad[,2]=grad
scaled_data<-scale(Igrad)
plot(x=temp, y=scaled_data[,1], type='p', col='red', xlab="", ylab="",ylim=c(-5,5))
lines(x=temp, y=scaled_data[,2], type='l', col='blue', lwd=2,xlab="", ylab="")
title(main=well)
tm.s=temp[which.max(Igrad[1:73,2])]
title(sub=sprintf("Tm = %4.1f", tm.s, cex.sub=1.2))
return(tm.s)}
pdf(file="C:/Documents and settings/username/path to file/Thermographs.pdf", width=30, height=21,pointsize=9)
layout(matrix(data=1:96, nrow=8, ncol=12, byrow=TRUE))
tma<-matrix(nrow=12,ncol=8)
for(i in 2:97) {try(expr=tma[i-1]<-find.tm(temp=raw_data[,1],I=raw_data[,i], grad=grad_data[,i], well=names(raw_data)[i]))}
dev.off()
```

It has to be noted that in line 2 of the above script, the matrix scale was adapted for the range of temperature tested in a particular experiment (i.e. ramping speed determines the number of rows in the database) and can be generally calculated as follow: (Total number of rows in the dataset – 1) x 2. In the above example, the temperature range was 35°C to 73.5°C with 0.5°C increment. In line 9, the red numbers indicate the range of rows corresponding to the temperatures over which the program identifies the highest value on the y-axis as T_m. Since the peak corresponding to the unfolding of GFP starts around 72°C and rapidly increases, the range of temperature used for T_m determination was set at 71.5°C (row 73 in the data set) to avoid false peak identification.

The following script generates a 2-D heat map of the 96-well plate with a color gradient code from red (low T_m) to yellow (high T_m):

```
pdf(file="C:/Documents and settings/username/Desktop/2Dheatmap.pdf", width=6, height=5,
paper="a4", pointsize=8)
tmaplot<- matrix(nrow=12,ncol=8,data=0)
for(i in 1:8) {tmaplot[,9-i]=tma[,i]}
image(tmaplot)
dev.off()
```

The correlation between T_m values obtained with the automatic peak recognition system described above and the visual curve analysis was tested using the Pearson r test in GraphPad Prism.

5.2.2.2   Effect of ionic strength on Tus-*Ter* complexes stability

Tus-GFP was incubated with each *Ter* site in the presence of eight KCl concentrations and subjected to the melt curve program of a real-time thermal cycler. Stock solutions of Tus-GFP, *Ter* oligonucleotides and KCl were prepared at three times the desired final concentration in SPR buffer (50 mM Tris (pH 7.7), 0.1 mM EDTA, 0.2 mM β-mercaptoethanol). An equal volume of each was mixed in a qPCR 96-well plate (Bio-Rad). The mixture was left 10 minutes at room temperature to reach equilibrium. *Ter* were in slight excess (3 µM) compared to Tus-GFP (2.5 µM) and KCl concentrations ranged from 8.4 to 351.5 mM. The IQ5 iCycler (Bio-Rad) was set on the melt curve program from 35°C to 75°C, 0.5°C/cycle, 30 s dwell time. The T_m values were determined by graphical analysis of the first derivative of RFU signal, or using the automatic peak recognition program developed with the free RStudio interface as described above.

5.3 Results and discussion

5.3.1 Automatic determination of T_m values and 2D-heat map screen

The first step was to increase the throughput of the method to be able to handle the volume of data generated in this study. A universal automatic peak recognition program was developed for the RStudio interface that produces a 2D-heatmap of a 96-well plate directly from raw data. This script provides individual thermoplots of normalized RFU and –dRFU/dT variables and reports the temperature at the maximum value of the derivative as T_m (see script in section 5.3.1). Figure 21A shows the thermoplots obtained for free and *TerC*-bound Tus-GFP at increasing KCl concentrations. The peak of free Tus-GFP shifted to slightly higher temperatures with increasing salt due to the stabilizing effect of KCl on Tus-GFP as previously observed (*cf* Chapter 3, Figure 11 p69).
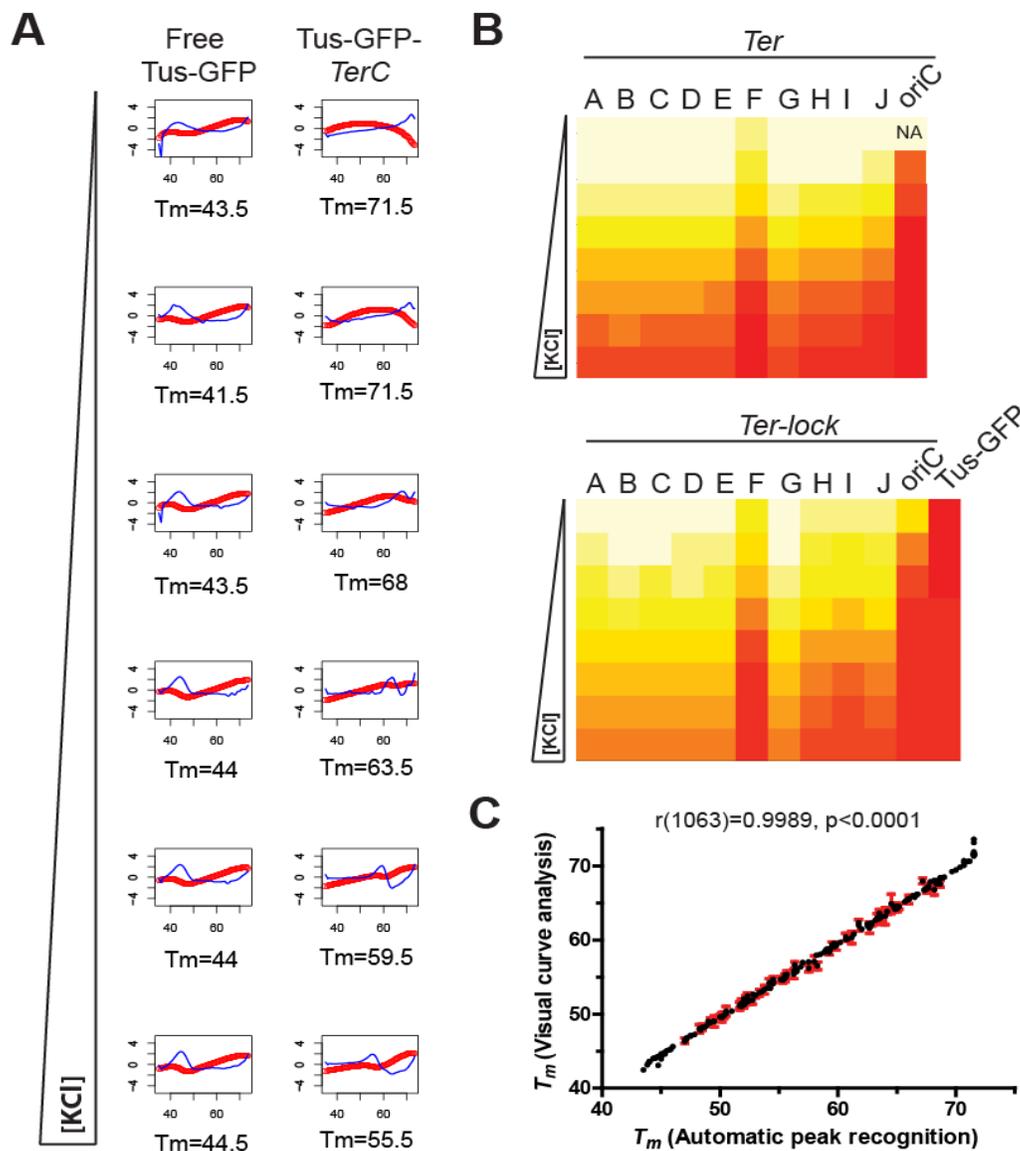
Figure 21: Automatic peak recognition of melt curves using RStudio interface. (A) Example of thermoplots showing RFU signal (red) and its derivative –dRFU/dT (blue) obtained for free or *TerC*-bound Tus-GFP at increasing KCl concentrations from 8.4 to 351.5 mM KCl. (B) 2D-heat map representing the T_m values of all Tus-*Ter* and Tus-*Ter-lock* complexes at increasing KCl concentration. The T_m was determined from the average of three –dRFU/dT curves. Tus-GFP was at 2.5 μM and *Ter* or *oriC* DNA was at 3 μM in SPR buffer (50 mM Tris pH 7.6, 0.1 mM EDTA and 0.2 mM β-mercapoethanol) supplemented with KCl at the above mentioned concentration. NA: not available due to false T_m determination as the peak corresponding to the unfolding and aggregation of Tus was below the peak of GFP at 71.5 °C. T_m values ranged from 45°C (red) to 71.5°C (pale yellow). (C) Pearson r correlation between T_m values obtained by visual analysis of DSF curves and by automatic peak recognition (RStudio) for each data point (*Ter*, *oriC*, their lock-forming analogues and free Tus-GFP data were obtained in triplicate for each salt concentration, n=183). The red error bars represent the 95 % confidence interval.

When bound to *TerC*, the Tus-GFP peak initially shifted by 28°C at the lowest salt concentration and gradually shifted to lower temperature with increasing salt due to the weakening of electrostatic interactions. Due to the prominent GFP peak between 75 and 80°C depending on the conditions used (**Error! Reference source not found.**B), the melt curve derivatives were only analyzed up to a temperature where the GFP signal was minimal in order to reduce false peak recognition rate. Here, DSF curves were analyzed until 71.5°C to detect peaks as small as 100 –dRFU/dT. An arbitrary T_m value of 71.5°C was automatically assigned to peaks at or above this temperature. This was only the case for *Ter* sites at the lowest KCl concentration (high affinity condition) with the exception of *TerF* which had a T_m value below 71.5°C. The T_m values for the remaining *Ter* sites could be visually determined and were generally within 0.5°C of the arbitrary value. Only for *TerG* and *TerI,* peaks at 73.6°C and 73.1°C respectively were missed by this method but could be obtained by visual examination (Figure 22C). Overall the data from the automatic peak recognition program correlated well with those obtained by visual determination of T_m for each curve (Figure 21D). Out of 552 T_m peaks analyzed with the program, only 24 were misidentified corresponding to an error rate of 4.3 %. Out of these errors, 50 % were due to the GFP peak being higher at 71.5°C than the POI peak and the remaining errors were from the unresolved peak on the original curves. This error rate could be further decreased by increasing the protein concentration and thus increasing the signal (**Error! Reference source not found.**E).

The 2D-heatmap of the 96-well plate provides a more explicit representation of the experimental screen by transforming T_m values across the plate by a 2-colors gradient code with the lowest T_m shown in dark red and the highest T_m shown in pale yellow (Figure 21B). The profile obtained for *Ter* and *Ter-lock* binding sites at 150 and 250 mM KCl were in good agreement with GFP-Basta (*cf* Chapter 4, Figure 15 p86 and Table 11 p88). This

demonstrated the accuracy of both visual and automatic methods and the significant advantage of the program for the analysis of high volumes of data.

5.3.2 Effect of ionic strength on Tus-*Ter* complexes stability

The salt dependence of binding of Tus-GFP to all ten *Ter* sites, their *Ter-lock* analogues and *oriC* was examined in the presence of 8.4 to 351.5 mM KCl (Figure 22). The oligonucleotides used in this study were the same as the ones used in the GFP-Basta experiments with the locked species having a 6 nucleotide single stranded tail at the non-permissive face that allow the C(6) to bind into the cytosine-binding pocket of Tus (Figure 22A).

The increase in T_m of Tus-GFP with increasing KCl concentration (Figure 22A-B) indicates that ions bind and stabilize the protein (Waldron et al., 2005). The non-specific *oriC* and *oriC-lock* conferred an increase in stability to Tus-GFP only below 150 mM KCl and resulted in a ΔT_m with free Tus-GFP of more than 10°C at the lowest salt concentration highlighting the significant contribution of electrostatic interactions in these conditions. The salt-dependent profiles obtained for the *Ter* and *Ter-lock* species correlate well with the data obtained with GFP-Basta at 150 and 250 mM KCl (*cf* Figure 15 p86; Moreau and Schaeffer, 2012a). The strong binders (*TerA-E* and *TerG*) induced a larger thermal shift than the moderate binders (*TerH-J*) at almost all salt concentrations reflecting the higher affinity of Tus for the strong binders (Figure 22A). Only *TerI* was as stabilizing as the strong binders at the lowest ionic strength (8.4 mM KCl).

The profiles obtained for all strong binders had similar slopes and amplitudes (Figure 22A and Table 14) with *TerG* having the highest T_m at the lowest salt concentration. These *Ter* sites (and their analogue) responded similarly to ionic strength suggesting that essentially the same ionic bonds are broken in these complexes and that additional or stronger

electrostatic interactions may occur in the Tus-*TerG* complex (Figure 22E). In accordance with previous data (*cf* Figure 15 p86; Moreau and Schaeffer, 2012a), *Ter* species were more stabilizing than their *Ter-lock* analogues in low salt due to the missing base specific and electrostatic interactions between G(6)/A(5) and R198 in the partially single-stranded *Ter-locks* compared to the *Ter* species. The R198 residue forms polar and Van der Walls contacts with A(5) and G(6) sugars in *Ter* species and a water-mediated ionic interaction with the phosphate group of G(6) (Kamada et al., 1996). It also forms specific hydrogen bonds with these two bases (Figure 23). Therefore R198 contributes significantly to the overall affinity of the complex (Neylon et al., 2000). The R198A mutant has a 150-fold reduced binding affinity for *TerB* (mainly affecting k_a) and also has a lower affinity for non-specific DNA demonstrating the significant contribution of this residue to complex formation rather than complex stability (Neylon et al., 2000). The higher stability induced by *Ter* compared to *Ter-lock* species in low salt is therefore mainly the result of a decreased association rate constant (k_a) with the Tus-*Ter-lock* complexes. In high salt, all ionic contacts are broken, reducing the difference in k_a between the two species and making the effect of the TT-lock apparent, at least for the strong TT-lock forming sites.
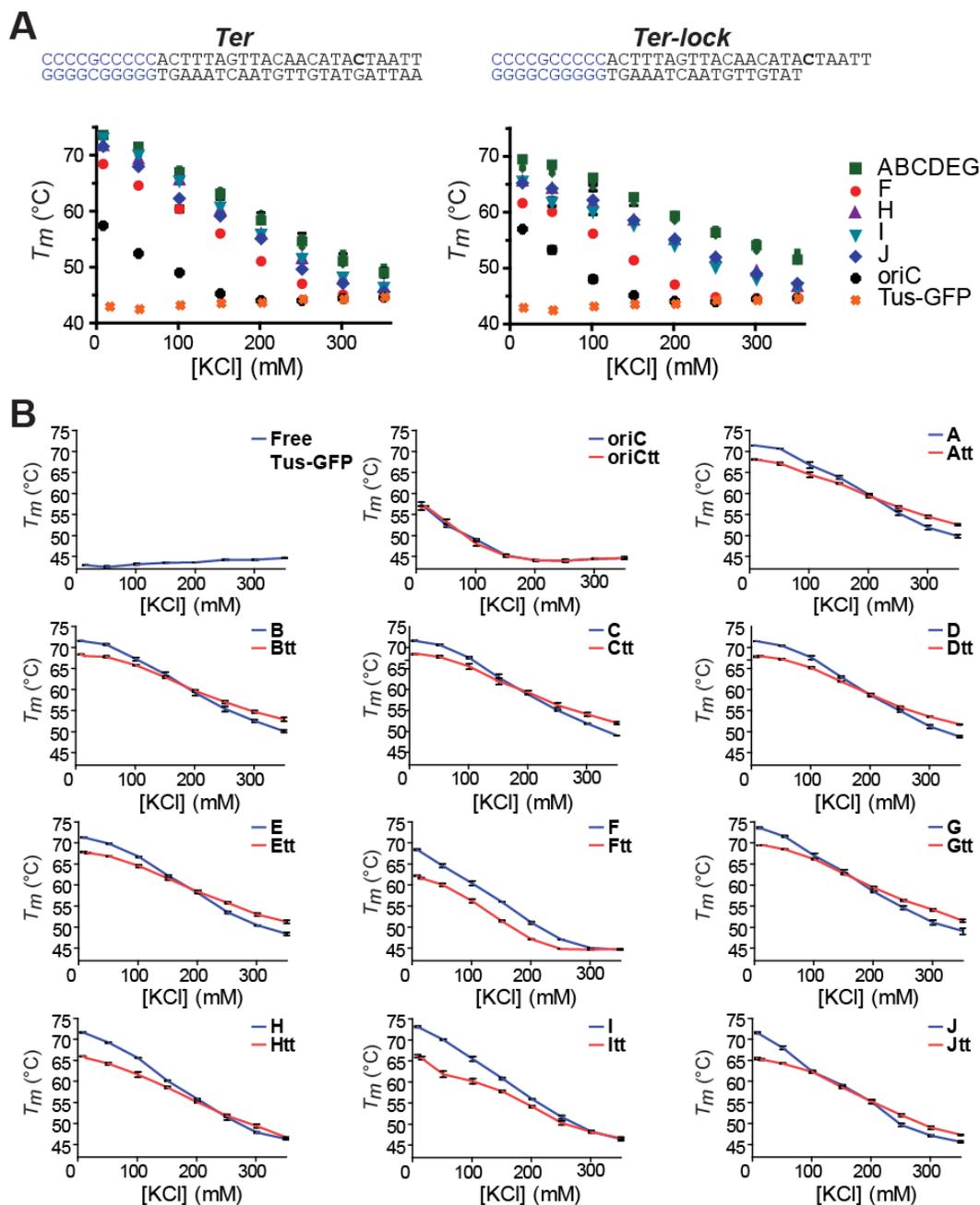
Figure 22: Effect of ionic strength on Tus-GFP in complex with *Ter* sites or their *Ter-lock* analogue. (A) KCl-dependent stability of Tus-GFP-*Ter* or Tus-GFP-*Ter-lock* complexes. *TerA* and *Ter-lockA* sequences and structures are shown as an example of the oligonucleotide design used for all *Ter* sites. The *Ter-lock* species are partially single stranded at the non-permissive end allowing C(6) to bind into the cytosine binding pocket at the surface of Tus. **(B)** *Ter* (blue) and *Ter-lock* (red)-induced stability at increasing KCl concentrations (8.4 mM to 350 mM). Tus-GFP (2.5 μM) and oligonucleotides (3 μM) were assembled in the presence of KCl and incubated for 10 minutes at room temperature prior to DSF-GTP.

As the KCl concentration increased, a shouldering effect was observed that was more prominent for the strong *Ter* sites (*TerA-E* and *TerG*) than the remaining *Ter* sites (Figure 22B) indicating a stronger contribution of specific interactions in these complexes that could not be outcompeted at low KCl concentrations.

As the KCl concentration increased further, a steeper negative slope and amplitude was observed for *Ter* species than for their *Ter-lock* analogue (Table 14), resulting in the crossing of *Ter* and *Ter-lock* profiles around 150 mM KCl for most *Ter* sites (Figure 22B). This trend indicates a larger contribution of electrostatic interactions in *Ter* species and the presence of additional specific interactions with *Ter-lock* species that reduce salt sensitivity. Ionic strength affects both site-specific DNA binding (K_{sp}) and non-specific DNA binding (K_{ns}). However, K_{sp} varies less severely with salt that does K_{ns} for non-specific binding (Engler et al., 1997, Garner and Rau, 1995, Record et al., 1991, Saecker, 2001, Sidorova and Rau, 1996). In *Ter-lock* species K_{ns} varies with less amplitude than in *Ter* species due to the missing interactions with G(6) compared to *Ter* species (Figure 23). On the other hand K_{sp} is negatively affected by the loss of specific interaction with A(5) and G(6) but positively affected by the formation of the TT-lock (Figure 23C-D). Therefore the only factor contributing to an increase in resistance to denaturation in high salt is the formation of the TT-lock.

Above 150 mM KCl, where only specific interactions contribute to the stability of the complexes (i.e. no binding to *oriC*), the contribution of the TT-lock was sufficient to overcome and/or surpass the loss of both electrostatic and specific interactions with the nucleotides missing at the non-permissive face of the *Ter-lock(A-E and G)* species. The magnitude of the contribution of each component to binding (TT-lock, specific and electrostatic) is reflected by the KCl concentration at the crossing point between *Ter* and *Ter-lock* profiles.

Table 14: Effect of ionic strength on Tus-GFP in complex with *Ter* sites or their *Ter-lock* analogue.

Ligand	*A*	*B*	*C*	*D*	*E*	*F**	*G*	*H*	*I*	*J*	*oriC***
Ter slope	-0.085 ± .005	-0.082 ± .007	-0.079 ± .006	-0.081 ± .004	-0.087 ± .003	-0.094 ± .005	-0.085 ± .006	-0.087 ± .005	-0.092 ± .004	-0.095 ± .006	-0.092 ± .008
Max T_m	71.4	71.5	71.6	71.5	71.3	68.4	73.6	71.7	73.1	71.6	57.4
Min T_m	49.8	50.1	49	48.8	48.4	44.6	49	46.3	46.4	45.6	44.6
Amplitude (°C)	21.6	21.4	22.6	22.7	22.9	23.8	24.6	25.4	26.7	26.0	12.8
Ter-lock slope	-0.057 ± .004	-0.058 ± .004	-0.058 ± .007	-0.061 ± .003	-0.057 ± .004	-0.091 ± .004	-0.063 ± .003	-0.067 ± .003	-0.075 ± .005	-0.066 ± .005	-0.10 ± .007
Max T_m	68.0	68.0	68.5	67.9	67.7	61.6	69.4	65.5	65.6	65.2	57.0
Min T_m	52.5	52.9	52.0	51.7	51.2	44.8	51.5	46.7	46.6	47.3	44.6
Amplitude (°C)	15.5	15.1	16.5	16.2	16.5	16.8	17.9	18.8	19.0	17.9	12.4

Slope (± SD) were obtained from linear regression of T_m values in the linear portion of the curve. The slope of all *Ter* and *Ter-lock* curves were taken between 150 and 250 mM KCl exept for *TerF*, *oriC* and their respective lock analogues. **TerF* and *Ter-lockF* were analysed between 100 and 200 mM KCl, and ***oriC* and *oriC-lock* were analyzed between 8.4 to 100 mM KCl. The max and min T_m values are the T_m at the lowest and highest KCl concentration respectively and the amplitude is the difference between these two values.

TerH-J differ in their electrostatic and specific contributions to binding. *TerH* has a less steep negative slope (Table 14) than *TerI* and *J* as a result of the maintained specific interaction with T(9) which is mutated in the other two species to an adenine. In accordance with previous data (*cf* Figure 15 p86 and Figure 16 p90; Moreau and Schaeffer, 2012a), although *TerJ* was the more susceptible to ionic strength suggesting that it forms less specific interactions, *Ter-lockJ* was able to strengthen the complex and confer stronger resistance to ionic strength than *Ter-lockI*. Like *TerG*, *TerI* has the highest T_m value at the lowest salt concentration compared to *TerH* and *TerJ*, however *Ter-lockI* was strongly affected at low KCl concentrations (Figure 22B). This data suggests that *TerI* forms numerous small cooperative electrostatic interactions. The salt dependent profile of *Ter-lockH* crosses *TerH* profile around 250 mM KCl whereas the profiles obtained for *TerI* and *Ter-lockI* did not cross. *Ter-lockI* had the steepest slope of all *Ter-lock* species. This was surprising given that *TerH* was shown to not form a TT-lock whereas *TerI* could form a moderate TT-lock (*cf* Figure 16 p90).
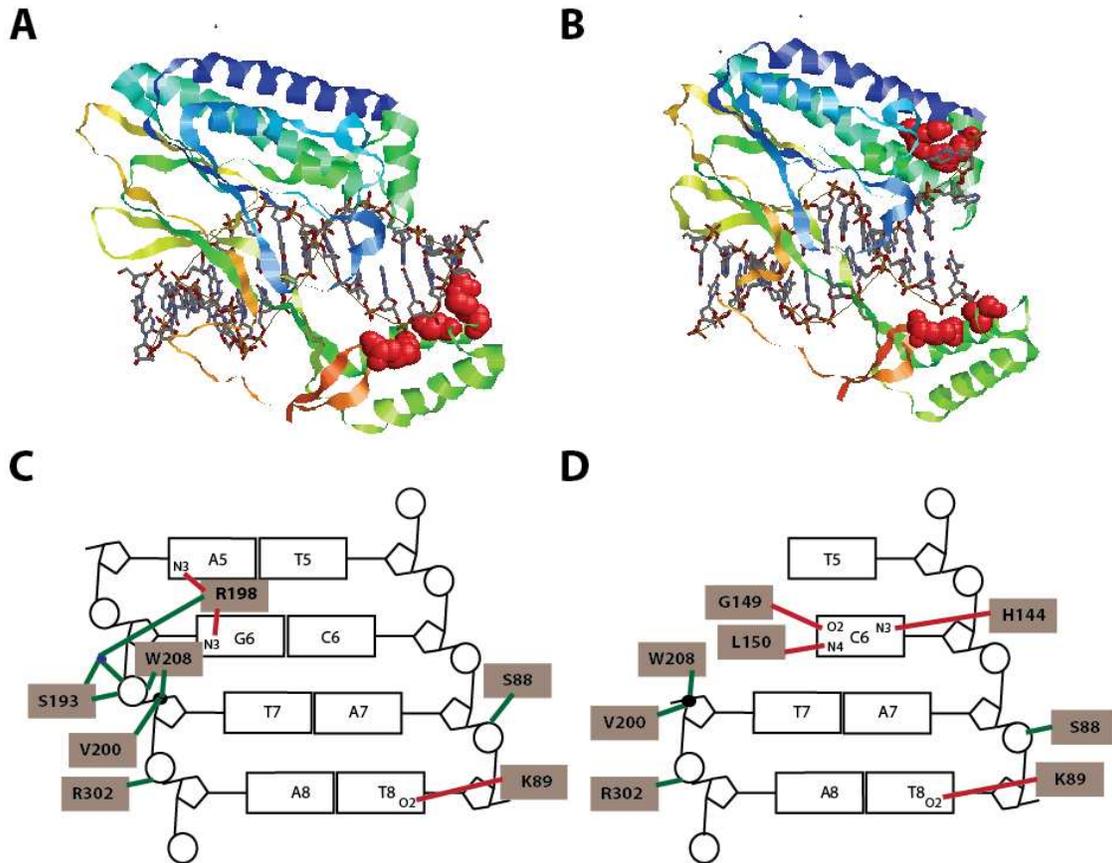
Figure 23: Tus-*Ter* interactions at the non-permissive face. Crystal structure of Tus in complex with (A) *TerA* (PDB 2I05) and (B) with *Ter-lockA* (PDB 2I06) with mismatched T(5) and C(6). In (A), amino-acids R198, S193, V200 and R302 and in (B) the cytosine binding pocket (G149, L150 , H144) and V200 and R302 are shown in red (spacefill of Van der Waals radius). Contacts between Tus and (C) *TerA* and (D) *Ter-lockA* used in this study. Sequence specific interactions are in red and non-specific interactions with the backbone DNA are in green. The black dot represents hydrophobic interactions. R198, with the nonpolar W208 and V200, is holding the C-carboxyl domain, namely α-helices VI and VII at the non-permissive face of *Ter*. In the *Ter-lock* species used in this study, residue (1-6) were not included to allow C(6) to flip out into the Tus cytosine specific binding pocket, therefore both R198 and S193 mediated contacts can not form. The domain is likely to be stabilised only by the nonpolar V200 and W208 residues and R302. Additional specific contacts occur with C(6) in *Ter-lockA*. The contacts with the remaining bases of the core sequence are maintained in both *TerA* and *Ter-lockA* (Mulcair et al., 2006).

This could be explained by the fact that *TerI* has reduced specific interactions compared to *TerH* and is therefore more strongly affected by ionic strength. The effect of the weak TT-lock is masked by a reduced k_a (and increased K_D) compared to the other moderate binders.

These results also suggest that some specific interactions might still occur in the *Ter-lockH* structure.

As expected from previous results (*cf* Figure 15 p86 and Figure 16 p90; Moreau and Schaeffer, 2012a), the *TerF* and *Ter-lockF* curves were essentially parallel (slopes of -0.94 ± 0.005 and -0.091 ± 0.004 respectively) confirming that *TerF* does not form a TT-lock (Table 14 and Figure 22). The difference in *TerF* and *Ter-lockF* is presumably the direct result of the effect of the loss of non-specific binding of R198 with the phosphate groups of G(6). The low affinity of *TerF* for Tus determined by SPR raised concerns about *TerF* being a real *Ter* site (*cf* section 4.4.1 p95; Moreau and Schaeffer, 2012a). The comparison of *TerF* curves and *oriC* curves showed that *TerF* maintained specific interactions with Tus up to 250 mM KCl (Figure 22A-B). At physiological concentrations (~150 mM), *oriC* did not stabilize Tus-GFP, whereas the low affinity *TerF* induced a thermal shift of 12.5°C. This result shows that *TerF* has maintained some specificity for Tus and could still act as a pausing site despite its low affinity and its inability to form a TT-lock.

It has to be noted that *in vivo*, Tus recognizes all *Ter* sites as fully double stranded DNA whereas here, Tus bound to partially stranded DNA, giving therefore only a partial measure of the TT-lock contribution as explained above (i.e. masked by a reduced k_a) and therefore subtle lock formation might be missed with this method like for *TerI*.

5.4 Conclusion

Taken together these results suggest that electrostatic interactions play an important role in Tus-*Ter* complex formation and stability. All Tus-*Ter* complexes reacted similarly to ionic strength, their differences in stability being mainly attributed to the number of specific interactions occurring between Tus and each *Ter* site and their respective TT-lock forming

capacity. *TerF* was the weakest termination site followed by *TerJ*, *TerI* and *TerH*, although *TerJ* forms a stronger TT-lock. This was previously observed using GFP-Basta and SPR (*cf* Chapter 4; Moreau and Schaeffer, 2012a). *TerA, B* and *C* were the most specific sites. These data also confirmed the importance of R198 in complex formation. This residue is likely to be the first step in *Ter* recognition and binding, bringing the flexible L3 loop of the C-domain close to DNA where additional specific interactions subsequently form. The data obtained with DSF-GTP were in good agreement with the results obtained with both GFP-Basta and SPR, and enabled the acquisition of information on the specificity and salt dependence of all Tus-*Ter* and Tus-*Ter*-lock complexes.

Finally, DSF-GTP is simple, fast, robust and insensitive to variations in reaction volumes. The technology is well suited for the study of protein-ligand interactions as it doesn't require solvatochromic dyes, eliminating the risk of interferences with additives, ligands or the protein itself. An advantageous and unique feature of DSF-GTP is that melting curves provide information on the effect of additives and buffers on fluorescence and stability of the GFP reporter itself, providing an in-built quality-control measure for individual reactions. The expression and folding reporter function of the C-terminal GFP tag (Waldo et al., 1999), combined with its new function as a sensor for protein aggregation, equips DSF-GTP with all essential features to become a powerful comprehensive HT tool for monitoring protein expression, folding, stability and ligand binding.

Chapter 6: A new polyplex qPCR-based DNA-binding assay to determine the preferential binding of Tus to genomic *Ter* sites *in vitro*

The data included in this chapter have been published in Analyst[7] (Moreau, M. J & Schaeffer, P. M). The original manuscript was slightly adapted for coherence with the rest of this thesis. The binding of Tus for the ten primary Ter sites (TerA-J) has been characterized in Chapter 4 and the effect of ionic strength in Chapter 5. This chapter describes the development of a new qPCR-DNA binding assay enabling to study the effect of flanking sequences proximal to Ter sites.

6.1 Introduction

In *E. coli*, two replisomes proceed bidirectionally to replicate the circular chromosome until they meet a section of the chromosome containing a series of termination (*Ter*) sites. These sequences are 21 bp in length (Hidaka et al., 1988, Hill et al., 1988a) and arranged in two clusters that act in a polar manner to constrain replication termination opposite to *oriC*. Replisomes can proceed through the first cluster of *Ter* sites on their way to the terminus, but

---

[7]Moreau, M. J. & Schaeffer, P. M. 2012. A polyplex qPCR-based binding assay for protein-DNA interactions. Analyst, 137, 4111-3.

will be stopped by the second cluster containing *Ter* sites in the opposite orientation (*cf* Figure 14 p79).

The study of protein–DNA interactions is challenging and often involves the manipulation of radioisotope-labelled material such as in electrophoretic mobility shift assays or radioactive filter-binding assay (Forwood and Jans, 2006, Oehler et al., 1999). Surface plasmon resonance (SPR) and fluorescence based methods have also been widely used for the study of protein-DNA interactions (Favicchio et al., 2009, Moreau et al., 2010, Moreau et al., 2012, Mulcair et al., 2006). These techniques are usually useful and very sensitive in acquiring quantitative data on binding affinities, although with limited throughput. Recently a promising microfluidics platform called MITOMI was developed that can be used to determine binding affinities in high-throughput (Geertz and Maerkl, 2010, Maerkl and Quake, 2007). For all these methods, the level of technical difficulties is high due to the need for specialised facilities or equipment and training. Specifically, for ChIP-seq and ChIP-chip experiments, where genome-wide binding data are obtained, there is also a need for more quantitative and comparative methods to rapidly validate the newly identified genomic DNA regions containing putative targets, in a time and cost-efficient manner.

The aim was to develop a simple and fast method to obtain comparative Tus-*Ter* binding affinity data using genomic *Ter* sites of about 150 bp in a competition and polyplex format. The method presented here is a new qPCR-based DNA-binding assay that involves the immunoprecipitation of GFP-tagged Tus in complex with a stoichiometric mixture of genomic *Ter* sites, followed by qPCR quantification of the immunocaptured DNA targets (Figure 25A). This method was also used as a validation step for the immunoprecipitation protocol for the ChIP-qPCR experiment on Tus described in Chapter 7.

6.2 Material and methods

6.2.1 Protein expression and purification

The His$_6$-Tus-GFP proteins were expressed and purified as previously described (Dahdah et al., 2009, Moreau et al., 2010) except that here, and like in Chapter 4 and 5, the ammonium sulfate pellets were resuspended in buffer A (50 mM Tris, pH 7.6, 250 mM KCl, 0.1 mM EDTA and 0.2 mM β-mercaptoethanol) and dialysed (SnakeSkin pleated dialysis tubing 10,000 MWCO; Pierce) twice against 200 ml of buffer A at 4°C. The purity of proteins was assessed by SDS-PAGE (Next-gel, Amresco) and concentration was determined by standard Bradford assay.

6.2.2 Determination of Tus-*Ter* binding by qPCR DNA-binding assay

The method was developed to compare the affinity of Tus for each *Ter* site in polyplex. For this assay, genomic DNA regions (~150 bp) containing *Ter* sites or *oriC* sequences (non-specific binding control) were amplified from *E. coli* DH12S using a MyCycler (Bio-Rad) with Taq DNA polymerase (New England Biolabs). Oligonucleotides, used to amplify the individual genomic regions containing the *TerA-J* sites, were standardized for PCR and are described in Table 4, p44. The protocol consisted of a denaturation step of 30 s at 95°C, followed by 40 cycles of 15 s at 95°C, 10 s at 60°C and 20 s at 68°C and a final extension step of 2 minutes at 68°C. DNA amplicons were purified and quantified after electrophoresis on an agarose gel using the image analysis software ImageJ (http://rsbweb.nih.gov/ij/). All DNA amplicons were diluted to a final concentration of 6 nM in TBS (20 mM Tris pH 7.5, 150 mM NaCl and 0.005% Tween 20). A solution containing one volume of Tus-GFP (6 nM in TBS) and one volume of each DNA amplicon (6 nM in TBS) was diluted in TBS buffer to obtain a final concentration of 0.4 nM of Tus-GFP and 4.4 nM of combined DNA amplicons. The reaction mix was left 10 minutes at room temperature to allow Tus binding to *Ter* sites.

Streptavidin coated plates (Thermoscientific, Reacti-Bind^{TM} Streptavidin coated HBC black 96-well plates with SuperBlock blocking buffer) were coated overnight with 50 μl of 1 μg/ml biotinylated goat anti-GFP antibody (Ab 66858; Abcam) in TBS at 4°C. The antibody suspension was removed and the wells were washed with 200 μl of TBS. A volume of 50 μl of Tus-GFP-DNA reaction was bound to each well for 60 minutes at room temperature. The supernatant was removed and wells were washed 5 times with 200 μl of TBS. DNA amplicons were dissociated with 50 μl of TBS containing 0.5 M NaCl during 30 minutes at room temperature, transferred into a new tube and diluted 10 times with water to reduce the salt concentration (output). The salt concentration and dilution of the initial Tus-GFP-DNA reaction were adjusted (input) to match the output conditions. Background controls were obtained using the same protocol with the omission of the anti-GFP antibody binding step. The IQ5 iCycler (Bio-Rad) was used for qPCR. Briefly, reactions contained 2 μl of input or output, 8 μl of primer pair (0.5 μM) and 10 μl of SensiMix SYBR & fluorescein mastermix (Bioline). The protocol used included 10 minutes activation at 95°C followed by 40 cycles at 95°C, 10 s and at 60°C, 10 s. A melt-curve was carried out to verify that the correct regions were amplified. Standard curves were obtained for each primer set with a 10-fold serial dilution of input matching the output buffer conditions (10-, 100- and 1000-fold).
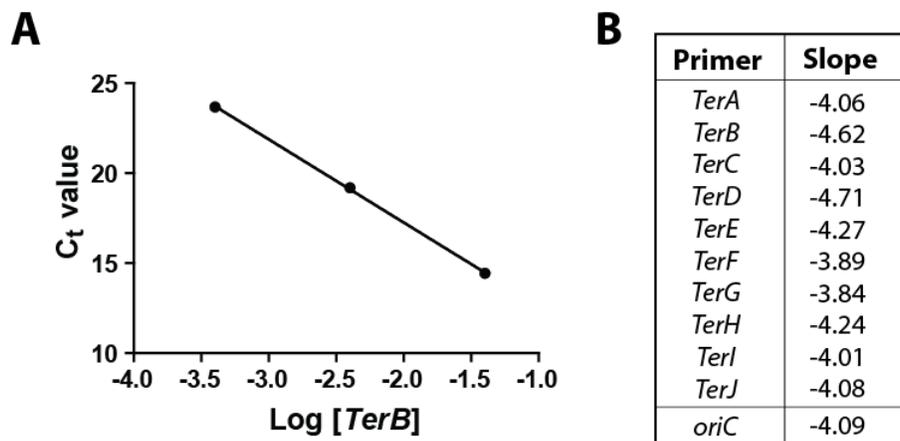
Figure 24: Determination of primer specific efficiency. (A) qPCR standard curve for *TerB*. **(B)** Slope values obtained for each primer.

ΔCt-values were obtained by subtracting background Ct-values (no Ab) from output Ct-values (with Ab). Slope values of the standard curves (for each primer set; see Figure 24) were used to obtain the enrichment factor using the relationship $10^{(\Delta Ct/slope)}$.

6.2.3 GFP-Basta

The thermal stability of Tus in complex with *Ter* and *Ter*-lock oligonucleotides at 150 mM KCl was carried out as described in section 4.2.2 p82, in buffer B (50 mM Tris, pH 7.6, 150 mM KCl, 0.1 mM EDTA and 0.2 mM β-mercaptoethanol). Oligonucleotides are described in Table 5, p45. Briefly, Tus-GFP (1.6 µM) was mixed with an equal volume of *Ter* DNA (2 µM). The reactions were left 10 minutes at room temperature to allow complex formation. Each reaction (70 µl) was heated at 58°C in a MyCycler (BioRad). Following cooling and centrifugation, the residual fluorescence in the supernatant (60 µl) was measured with the S method and aggregation rate constants (k_{agg}) were determined as described previously (*cf* section 3.2.4 p65). Aggregation half-lives ($t_{1/2-agg}$) were obtained as ln2/k_{agg}.

6.2.4 SPR

SPR experiments were carried out as described in Chapter 4 (*cf* section 4.2.3 p83) in buffer B (50 mM Tris, pH 7.6, 150 mM KCl, 0.1 mM EDTA and 0.2 mM β-mercaptoethanol) at 20°C.

6.3 Results and discussion

6.3.1 Preferential binding of Tus to the ten primary *Ter* sites

A new qPCR based DNA-binding assay was developed to compare the affinity of GFP-tagged Tus (Tus–GFP; Dahdah et al., 2009, Moreau et al., 2010, Moreau et al., 2012) for each *Ter* site in polyplex. For this assay, ten genomic DNA segments (~150 bp) containing the *TerA–J* sites and one region containing the *oriC* (non-specific binding) were amplified from *E. coli* DH12S (*cf* Table 4 p44 for primer sequences). Each of these DNA regions, as well as the Tus–GFP, was combined in a single reaction in equimolar amounts (each at 0.4 nM). The mixture of protein–DNA complexes was immuno-captured with an anti-GFP antibody and the relative amounts of each *Ter-* and *oriC*-containing DNA region were quantified by qPCR and converted into enrichment factor (EF) values (Figure 25B-C). It is important to note that the assay was performed in a competition format (all eleven DNA regions were present in the same reaction in equimolecular amounts during the binding step), so the EF values should be directly proportional to the relative differences in affinity of Tus–GFP for the different DNA regions (see section 6.2.2 for detailed experimental procedure).

A

Ter-containing genomic DNA

Tus-GFP fusion protein

Anti-GFP antibody

qPCR

B

qPCR DNA-binding assay

x-fold enrichment

A B C D E F G H I J oriC

C

x-fold enrichment

Site	Mean	SEM
A	7287	3507
B	4087	994
C	3347	182
D	2910	1632
E	1435	35
F	8	2
G	3637	921
H	513	210
I	321	159
J	153	109
oriC	1	0.3

D

GFP-Basta

$t_{1/2\text{-}agg}$

A B C D E F G H I J oriC Tus-GFP

E

$t_{1/2\text{-}agg}$

Site	Mean	SEM
A	7445	2653
B	4064	362
C	2484	130
D	1848	69
E	799	31
F	68	8
G	2044	519
H	294	17
I	355	45
J	244	8
oriC	10	1
-	8	0.2

Figure 25: Binding of Tus–GFP to *TerA–J*-containing genomic DNA regions determined by the qPCR DNA-binding assay. (A) Schematic representation of the assay. Ten genomic regions comprising each individual *TerA–J* site were amplified from *E. coli* DH12S genomic DNA by PCR and pooled in equimolar ratio with Tus–GFP. The DNA-Tus–GFP complexes were immunocaptured with goat polyclonal anti-GFP IgG and the DNA was eluted and quantified by qPCR. (B) Comparison of enrichment factors of each individual *Ter* genomic region. (C) Mean enrichment factor and SEM values (n=2) are given. Enrichment factor values were obtained by dividing the relative abundance of each genomic region by their relative background value. Background values were obtained by omitting the anti-GFP capture IgG. (D) Measurement of DNA-induced thermal stabilization of Tus–GFP using GFP-Basta. First order aggregation kinetics were measured at 58°C in 150 mM KCl and transformed into half-life values ($t_{1/2\text{-}agg}$). The error bars represent the upper and lower limits of the 95% CI of the mean obtained from $t_{1/2\text{-}agg}$ in duplicate. (E) Mean $t_{1/2\text{-}agg}$ values and SEM values (n=2).

To confirm that the qPCR DNA-binding assay data were reflecting real differences in affinities the binding of the ten *TerA–J* sites was compared to the half lives of aggregation obtained with GFP-Basta (Moreau et al., 2010). Here an affinity profile was generated using 30 bp sequences containing *Ter* sites (Figure 25D). The profile obtained with GFP-Basta was almost identical to that obtained with the new qPCR DNA-binding assay (Figure 25B). *TerA–J* could be ranked into strong, moderate and low affinity binders. *TerA–E*, and *TerG* had similar strong affinities for Tus–GFP whereas *TerH–J* had moderate affinity for Tus. *TerF* was binding only slightly better than the non-specific *oriC* region which correlated well with the GFP-Basta data and a recent study by Duggin and Bell (2009). Taken together the differences in binding seen for *TerA-J*-containing genomic DNA fragments to Tus confirmed the data obtained with the GFP-Basta for these species, and that no other significant effect on Tus-binding is conferred by bases adjacent to any of the *Ter* sites.

6.3.2 Estimation of K_D values from enrichment factor obtained by qPCR

The fraction of bound *Ter* is defined by the equilibrium dissociation constant (K_D). Hence, the comparative analysis of ten *Ter* sites using the qPCR DNA-binding assay enabled to rank all primary *Ter* sites with regards to their affinity for Tus-GFP. Furthermore, if a reference K_D value (rK_D) is obtained or known for one of the ligands in the conditions tested, then the K_D values for the remaining sites can be inferred directly from the difference in enrichment factors between the reference ligand ($_rEF$) and the unknown ligand ($_uEF$) from the relationship:

$$cK_D = rK_D * \left(\frac{_rEF}{_uEF}\right)$$

This relationship was tested with one $_rK_D$ (*TerF*) and three control K_D (*TerH–J*) values obtained by SPR using the untagged His$_6$-Tus (*cf* methods section 6.2.4 and Table 15 for values obtained by SPR).

Table 15: Determination of cK_D values for all *Ter* sites.

Ter	EF	SPR K_D (nM)	cK_D (nM)
F	8[a]	88[b]	88
H	513	1.4	1.4
I	321	1.3	2.2
J	153	6.1	4.6
A	7287	ND	0.10
B	4087	ND	0.17
C	3347	ND	0.21
D	2910	ND	0.24
E	1435	ND	0.49
G	3637	ND	0.19

[a] Reference EF value. [b] Reference K_D values obtained by SPR. [c] Control values obtained by SPR. ND: could not be determined by SPR. cK_D:calculated values using SPR K_D of *TerF* as a reference value.

In the tested condition (150 mM KCl), K_D values could only be obtained for *TerF,H–J* by SPR because the remaining Tus–*Ter* interactions were outside the range of the SPR machine. Using *TerF* as a reference, the K_D values were calculated (cK_D) for all *Ter* sites using the previous equation. The cK_D values for *TerH–J* were almost identical to the values obtained by SPR (Table 15). The system allowed the determination of the cK_D values for the remaining six *Ter* sites with confidence (Table 15). This would not have been possible using the well established SPR method only. Here, K_D values in the picomolar range could be determined but the limits of the method for high-affinity binders were clearly not reached. For low affinity binders of Tus–GFP, K_D values as low as 4×10^{-7} M can be determined based on three times the SEM of the EF of *oriC*. This value is only indicative as the

sensitivity of the method for low-affinity binders will be affected by the nature of the proteins, DNA, and their individual concentrations.

6.4 Conclusion

 In conclusion, the new qPCR DNA-binding assay was very sensitive, fast, convenient and simple-to-use, when compared to traditional methods, with the additional benefit that multiple DNA sequences can be analyzed in polyplex using one reaction. The system can be used in combination with GFP-Basta and SPR, and can also in principle be extended to determine dissociation rate constants. The enrichment factors obtained for 150 bp long *Ter* sites in competition format correlated well with *Ter* induced stability of Tus-GFP previously determined for the 21-bp *Ter* sites and confirmed that no nucleotide sequence other than the core sequence was involved in Tus binding. The qPCR DNA-binding assay can easily be adapted to determine the binding specificity of virtually any soluble and functional epitope-tagged DNA-binding protein. Finally, the system will be a useful and cost-effective alternative to tiled microarrays for refining low resolution and qualitative ChIP-chip and ChIP-seq data.

Chapter 7: *In vivo* distribution of Tus in the replication fork trap

7.1 Introduction

The replication termination protein of *E. coli* Tus, binds to a 21 bp sequence including a 16 bp consensus sequence (5'-AGNATGTTGTAACTAA-3') that is repeated ten times in the chromosomal termination region (called primary *Ter* sites in this manuscript). In 2009, a more permissive consensus sequence (GNRNGTTGTAAYKA) identified four new *Ter* sites (*TerK*, *L*, *Y* and *Z*), one within the termination region and the other three being on the left part of the chromosome (Figure 26). Interestingly, two of them (*TerZ* and *Y*) are oriented to block origin-to-terminus replication forks and are located 490 and 1060 kbp away from the origin (Duggin and Bell, 2009). The ten primary *Ter*-sites (*TerA-J*) are arranged in two clusters of five sites, one on each chromosomal arm, oriented with opposite polarity. One cluster arrests the clockwise moving fork (*TerB*, *C*, *F*, *G* and *J*) and the other cluster arrests the anti-clockwise moving fork (*TerA*, *D*, *E*, *I*, *H*, Figure 26A). Tus binds to these sites with varying affinity as demonstrated in Chapter 4, 5 and 6. Each cluster is composed of three high affinity and two moderate-to-low affinity *Ter* sites (Figure 26A; Moreau and Schaeffer, 2012b, Moreau and Schaeffer, 2012a). The affinity of Tus for the remaining four *Ter* sites (*TerK*, *L*, *Y* and *Z*) has not been determined but is likely to be very weak based on their respective

sequences (Figure 26B and Table 13 p96) and on their intrinsic fork arrest efficiency measured by Duggin and Bell (2009).
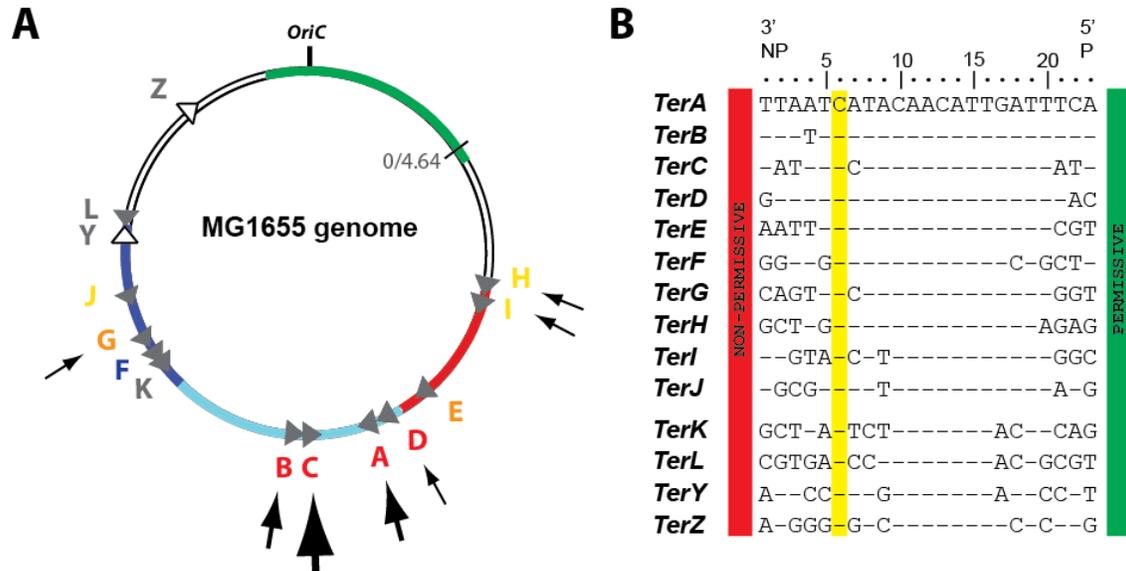
Figure 26: Distribution and sequences of *Ter* sites and their affinity for Tus in *E. coli*. (A) The colored circle represents the chromosomal macrodomains (green: ori domain, dark blue: left domain, red: right domain, light blue: termination domain, white: non-structured domain) according to Scolari et al. (2011) and Valens et al. (2004). The ten primary *Ter*-sites (*TerA-J*) are color coded from blue to red as a function of decreasing dissociation rate constant (k_d) determined for double-stranded *Ter* sites at 250 mM KCl. The grey triangles represent the orientation of the *Ter* sites with the tip corresponding to the non-permissive face of the complex. The grey labeled *Ter* sites have no affinity data available and are the least conserved *Ter* sites. *TerY* and *Z* are oriented to block a fork moving in the origin-to-terminus direction (white triangles). The outer black arrows pointing towards *Ter* sites show where paused fork has been detected *in vivo* under wild type conditions, with the size of the arrow indicating the frequency of pausing, the larger, the more frequent according to Duggin and Bell (2009). (B) *Ter* site sequences with the conserved C(6) highlighted in yellow. NP: non-permissive face, P: permissive face.

The well studied polarity of the Tus-*Ter* complex is mediated by the unusual asymmetric binding mode of Tus to *Ter* DNA and by the unwinding action of the DnaB helicase at the non-permissive face (Bastia et al., 2008, Duggin and Bell, 2009, Kaplan, 2006, Mulcair et al., 2006, Neylon et al., 2000, Neylon et al., 2005, Schaeffer et al., 2005, Moreau and Schaeffer, 2012a). Tus binds tightly to *Ter*, bending the double helix and precisely docking Tus on the

chromosome for its subsequent locking when DnaB unwinds DNA at the non-permissive face of the complex (Kamada et al., 1996, Mulcair et al., 2006, Moreau and Schaeffer, 2012a). The formation of forked DNA on this specific side of the Tus-*Ter* complex results in the flipping and locking of the C(6) base of the *Ter* core sequence into a specific cytosine binding pocket on the surface of Tus. This Tus-*Ter*-lock mechanism prevents Tus dissociation and inhibits DnaB translocation (Kaplan, 2006, Moreau and Schaeffer, 2012a, Mulcair et al., 2006). It has also been proposed that a specific interaction between DnaB and the non-permissive face of Tus has a role in the polarity of fork arrest (Bastia et al., 2008, Mulugu et al., 2001). All *Ter* sites where shown to have some DNA replication arrest activity but their use and efficiency varied dramatically (Duggin and Bell, 2009). This variation in DNA replication arrest efficiency was best correlated to their ability to form a TT-lock whether strong or moderate (*cf* Chapter 4; Moreau and Schaeffer, 2012a).

In Chapter 4, Tus was shown to bind with varying affinity to the ten primary *Ter* sites (*TerA-J*) and these differences were mostly due to a 10-fold variation in dissociation rates k_d between the strong *Ter* sites (*TerA-E* and *TerG*) and the moderate affinity sites (*TerH-I*) at 250 mM KCl (*cf* Figure 16 p90; Moreau and Schaeffer, 2012a). A ~10-fold difference in calculated dissociation constants (cK_D) was also obtained at 150 mM KCl between strong and moderate binders (*cf* Table 15 p35; Moreau and Schaeffer, 2012b). The weakest affinity site, *TerF,* had a 20 to 880 times higher dissociation constant cK_D at 150 mM KCl compared to the weak *TerJ* and strong *TerA* and was not able to form a TT-lock. All the strong *Ter* sites (*TerA-E* and *TerG*) were shown to form a strong TT-lock whereas more distal and weaker *Ter* sites (*TerH-J*) produced a weaker lock or were not able to form a lock (*cf* Figure 16 p90; Moreau and Schaeffer, 2012a). *TerH* was unable to form a TT-lock due to the T to G substitution at position 5 in the core sequence, a base important for DNA recognition and TT-lock formation (*cf* 4.4.2 p96; Moreau and Schaeffer, 2012a). Interestingly, Duggin and Bell

(2009) showed that no pausing occurred at *TerE* whereas some occurred at the outer *TerH* and *I in vivo* (Duggin and Bell, 2009). To be arrested at *TerH*, the fork has to break through the stronger *TerE* and moderate *TerI*, but no fork pausing was observed at *TerE* and little at *TerI* (Duggin and Bell, 2009). Nevertheless, *TerE* could arrest forks in a unidirectional replication plasmid assay with an efficiency proportional to its affinity and lock strength (*cf* Chapter 4, Duggin and Bell, 2009, Moreau and Schaeffer, 2012a). The low probability of the anti-clockwise fork to reach *TerH*, the absence of pausing at the strong *TerE* and the non-TT-lock forming characteristic of *TerH*, suggest that the pausing observed at *TerH* could be either due to the pausing of the clockwise moving fork at the permissive face of Tus-*TerH* or to recombination events at *TerH* (Horiuchi et al., 1995, Mohanty et al., 2009, Rothstein et al., 2000). Duggin and Bell (2009) showed that pausing was abolished at *TerC* in a *tus* null strain, confirming that the Y-shaped DNA intermediates were indeed due to the blocking effect of the Tus-*TerC* complex. Nevertheless, they did not verify if the pausing observed at the outermost *TerH-I* sites was also strictly due to Tus binding.

The presence of the distal *Ter* sites and their involvement in DNA termination remains unclear. Forks most frequently meet at *TerC* and to some extend at *TerA* as a result of different rates of accumulation of paused forked at each site (Duggin and Bell, 2009). Assuming the two forks progress at equivalent rates, forks are most likely to meet at *TerC* than at *TerA* since *TerC* is almost perfectly located directly opposite to *oriC* whereas the anti-clockwise moving fork must travel an additional ~259 kb to encounter the non-permissive face of the Tus-*TerA* complex. Despite the strength of the Tus-*Ter-lockC* (*cf* Figure 16 p90) significant pausing still occurred at *TerB* and to some extend at *TerG* (Duggin and Bell, 2009). A three step model has been proposed for the polar fork arrest involving the non-specific binding of Tus to DNA followed by the precise docking of Tus to a strong *Ter* sites where it acts like a linear ratchet that becomes locked when the DnaB unzips *Ter* at the non-

permissive face and induce the TT-lock (Moreau and Schaeffer, 2012a). One explanation for pausing at *TerB* is that in some cases, the ratchet-lock mechanism fails to form and the next site serves as a backup for DNA replication arrest.

The low affinity and usage of distal *Ter* sites raises the question whether or not Tus, expressed at low level (Roecklein et al., 1991, Roecklein and Kuempel, 1992) is bound to all *Ter* sites *in vivo*? The affinity and kinetic parameters of the ten primary *Ter* sites have been described in Chapter 4 but there is no data on how their binding properties relate to their occupancy by Tus *in vivo*.

Chromosome-immunoprecipitation (ChIP) techniques are powerful tools to study protein-DNA interactions *in vivo*. In *E. coli,* ChIP has been mainly used for the identification of transcription factors (CRP, MelR, FNR) and nucleoid associated proteins (FIS, N-HS, IHF) targets or to study the effect of chromosome domain organization on gene expression and replication (Grainger et al., 2007, Grainger et al., 2005, Grainger et al., 2004, Oshima et al., 2006). Amongst the DNA replication associated factors, SeqA (replication initiation and chromosome dynamics) has been extensively studied by ChIP as well as SlmA (chromosome segregation with FtsZ) and MatP (chromosome segregation and cell division; reviewed in Dame et al., 2011). With regards to replisomal proteins, the chromosomal distribution of DnaA binding sites, helicase loading factors and helicase were characterized in *B. subtilis* (Breier and Grossman, 2009, Ishikawa et al., 2007, Smits et al., 2011) but no study on the chromosomal distribution of replisomal proteins, including Tus, was ever reported in *E. coli*.

In this chapter, I attempted to determine the occupancy of *Ter* sites by Tus in exponentially growing cells by ChIP-qPCR using two alternative approaches. One approach was to raise antibodies against Tus, DnaA and SSB but this approach was unsuccessful. A second successful approach was to use GFP-tagged Tus and DnaA proteins and a commercial anti-GFP IgG for their immunoprecipitation. This chapter presents the first comparative and

quantitative study on the binding of Tus to the ten primary chromosomal *Ter* sites *in vivo* and the effect of the orientation of ectopic *Ter* sites on cellular growth rate.

7.2 Material and methods

7.2.1 Expression and purification of DnaA-GFP

His$_6$-DnaA-GFP proteins were expressed and affinity purified as described in section 2.2.3 p51. Following ammonium sulphate precipitation on protein elutions, protein pellets were resuspended in DnaA buffer (50 mM HEPES-KOH (pH 7.6), 1 mM EDTA, 1 mM β-mercaptoethanol, 20 % sucrose (w/v)).

7.2.2 Strains and plasmids used for ChIP

E. coli KRX (K12 derivative) was used to induce moderate levels of plasmid-born DnaA-GFP and Tus-GFP as it carries the T7 RNA polymerase gene under the tight control of the rhamnose promoter (*rha*P$_{BAD}$). The Tus-GFP plasmid pPS1259 was previously described (Dahdah et al., 2009). The construction of plasmid pMM220 encoding DnaA-GFP is described in section 2.2.1.5, p48.

7.2.3 ChIP-qPCR analysis

7.2.3.1   Protein induction and crosslinking

The *de novo* development of the ChIP-qPCR protocol presented here was influenced by previous work by Regev et al. (2012) and Ishikawa et al. (2007). Plasmids pPS1259 (Tus-GFP) and pMM220 (DnaA-GFP) were transformed into competent KRX cells and grown overnight at 37°C. For Tus-GFP cells, colonies were resuspended and diluted to an OD$_{600}$ of

0.1 in 12 ml of LB broth supplemented with ampicillin (100 µg/ml). For DnaA-GFP cells, single colonies were first streaked on a master plate (LB agar plate supplemented with 100 µg/ml ampicillin and 0.4 % glucose to avoid toxicity) and incubated for 24 hours at 37°C. These colonies were then resuspended in 12 ml of LB broth supplemented with ampicillin (100 µg/ml) at an OD_{600} of 0.1. All cultures were grown for 45 minutes at 37°C before inducing low expression levels of GFP-tagged proteins with 0.02 % Rhamnose (final concentration). Cells were incubated for 2 hours at 37°C, followed by 2 hours at 16°C. A 9 ml culture aliquot was transferred on ice for 30 minute and bacterial nucleoproteins were crosslinked by the addition of 36 % formaldehyde to yield a final concentration of 1 %. After 20 minutes at room temperature, crosslinking was quenched by the addition of solid glycine (0.5 M final concentration) for 5 minutes at room temperature followed by 5 minutes on ice. Cells were then centrifuged 5 minutes at 800 *g* at 4°C and washed twice with 4 ml and 10 ml of cold TCS buffer (50 mM Tris (pH 7.5), 150 mM NaCl and 2 mM KCl). KRX cells without plasmid were subjected to the same protocol in parallel (control). Cell pellets were stored at -80°C until required.

7.2.3.2   Detection and quantitation of overproduced GFP-tagged proteins

An aliquot of each culture was taken prior to crosslinking, centrifuged at 1,000 g for 1 minute and resuspended in SDS-PAGE loading buffer at a concentration of 7.8 x 10^9 cells.ml^{-1}. The mixture was heated for 10 minutes at 90°C and 5 µl of total proteins sample (corresponding to total proteins of 3.95 x 10^7 cells) were separated in 10 % SDS-PAGE (*cf* section 2.2.4.1 p54) alongside known amounts of purified Tus-GFP (0.5 µg) or DnaA-GFP (0.5 and 1 µg) proteins as standards. Proteins were transferred to immuno-blot PVDF membrane as described in section 2.2.4.4, p56. The membrane was blocked with 5 % skim milk in PBST (10 mM sodium phosphate (pH 7.4), 150 mM NaCl, 0.05% Tween) for 1 hour at room

temperature with gentle agitation. After three washes with PBST for 5 minutes, the membrane was incubated in a sealed bag with 10 ml of 1µg/ml chicken anti-GFP IgY (Abcam ab92456) in PBST with 1 % skim milk for 1 hour at room temperature with gentle agitation. Following three washes as described above, 10 ml HRP-conjugated goat anti-IgY (Jackson 103-035-155) was applied at 0.16 µg/ml (1/5000 dilution) in PBST with 1 % skim milk for 1 hour at RT. Following three washes, fast 3,3'-diaminobenzidine tablets set (DAB, Sigma) were dissolved in 5 ml $_{dd}H_2O$ and applied to the membrane in the dark for 10 minutes. The membrane was rinsed in PBS and allowed to dry overnight in the dark prior scanning. Bands were quantified using imageJ (http://rsbweb.nih.gov/ij/) and intracelullar concentrations were estimated based on the intensity of bands of known protein concentration and using cell parameters determined by Volkmer and Heinemann (2011) for cell volume (4.4 fL) and cell concentration at a given OD_{600} in LB (7.8×10^8 cells.ml^{-1}.OD^{-1}).

7.2.3.3   Immunoprecipitation and qPCR

Cell pellets were resuspended in lysis buffer (10 mM Tris (pH 8), 20 % sucrose, 50 mM NaCl, 10 mM EDTA, 1 mg/ml lsozyme and 10 µg/ml RNase) in 1/10 of initial culture volume (adjusted between replicates to reach same suspension concentration). Following a 30 minutes incubation period at 37°C, the lysate was diluted 5 times in IP buffer (50 mM HEPES-KOH (pH7.5), 150 mM NaCl, 1 mM EDTA) and passed three times in a French press at 12,000 psi to ensure maximum and reproducible cell lysis and DNA shearing. The Tus-GFP lysates were heated for 10 minutes at 50°C to denature free Tus-GFP. Control KRX lysates were treated accordingly in parallel for each replicate. After centrifugation at 30,000 *g* for 20 minutes at 4°C, a 50 µl-aliquot of cleared lysate (input) was incubated for 90 minutes at room temperature in a 96-well MAXISORB plate coated overnight at 4°C with 0.5 µg of goat anti-GFP IgG (Abcam; Ab6673) in 50 mM phosphate buffer (pH 7.5) supplemented

with 10 % glycerol. Wells were washed once with 200 µl of TCS buffer prior to immunoprecipitation. An immunoprecipitation experiment without antibody was performed in parallel as a background control. After 90 minutes, wells were washed three times with 200 µl of TCS buffer. Immunocaptured DNA was released by adding 50 µl of elution and de-crosslinking buffer (2 mM Tris, 50 mM NaCl, 0.005 % tween and 300 µg/ml proteinase K) to each well for 1 hour at 37°C (output). In parallel, the input was diluted 10,000 times in elution buffer (2 mM Tris, 50 mM NaCl, 0.005 % Tween) and 50 µl was transferred to a tube containing proteinase K at the same final concentration (300 µg/ml) to de-crosslink input DNA.

Samples (inputs and outputs) were incubated 15 minutes at 95°C to denature proteinase K and residual crosslinked proteins. After 5 minutes incubation on ice, samples were centrifuged at maximum speed for 5 minutes at 4°C and the supernatant was used as template for qPCR. qPCR reactions contained 2 µl of input or output, 8 µl of primer pairs at 0.5 µM each and 10 µl of SensiMix SYBR & fluorescein mastermix (Bioline). The 'no template controls' were run in parallel. The protocol used included 10 minutes activation at 95°C followed by 40 cycles at 95°C, 10 s and 60°C, 15 s.

7.2.3.4   Standard curves

A standard curve was performed on purified *Ter* and *oriC* amplicons and diluted 10-, 100- and 1000-fold matching output buffer conditions of ChIP experiment as described in section 6.2.2, p126. The average slope of triplicate standard curves was used to determine the primer specific efficiency (E_{amp}) as follow:

$$E_{amp} = 10^{\left(-\frac{1}{slope}\right)}$$

An E_{amp} value of 2 indicates that the primer amplify with 100 % efficiency, doubling the quantity of starting material every cycle.

7.2.3.5   Data analysis

Ct-values were obtained at the same threshold Ct-value for all experiments. A melt-curve was performed to verify that the correct regions were amplified. ChIP-qPCR data were analysed by comparative quantitation as follow:

$Ct_{(input)}$-values obtained by qPCR were corrected for the dilution factor ($_cCt_{(input)}$) according to the following equation:

$$_cCt_{(input)} = Ct_{input} - log_{E_{amp}}(dilution\ factor)$$

The immunoprecipitation efficiency of each specific target DNA region relative to a non-specific DNA region (IP efficiency$_{(ns)}$) was calculated as follow:

$$IP\ efficiency_{(nS)} = \frac{E_{amp}^{\left(_cCt_{(input)sp} - Ct_{(output)sp}\right)}}{E_{amp}^{\left(_cCt_{(input)ns} - Ct_{(output)ns}\right)}}$$

where $_cCt_{(input)}$ and $Ct_{(output)}$ are the Ct-values obtained for each DNA target before (input) and after ChIP (output). Specific DNA target (i.e. binding sites) and non-specific control DNA region are indicated with "sp" and "ns" subscripts respectively.

The enrichment factor relative to the no antibody control (EF$_{(No\ Ab)}$) was calculated as follow:

$$EF_{(No\ Ab)} = E_{amp}^{\left(Ct_{(No\ Ab)} - Ct_{(IP)}\right)}$$

where $Ct_{(IP)}$ and $Ct_{(No\ Ab)}$ are the Ct-values obtained with output samples from wells coated with or without anti-GFP antibody respectively.

The enrichment factor relative to control KRX cells (EF$_{(KRX)}$) lacking the GFP-tagged protein was calculated as follow:

$$EF_{(KRX)} = E_{amp}^{\left(\left(_cCt_{(POI-GFP^+\ input)} - Ct_{(POI-GFP^+\ output)}\right) - \left(_cCt_{(POI-GFP^-\ input)} - Ct_{(POI-GFP^-\ output)}\right)\right)}$$

where POI-GFP$^+$ and POI-GFP$^-$ refer to the strain expressing or lacking the GFP-tagged protein (i.e. Tus-GFP or DnaA-GFP). Enrichment factors were calculated for each biological

replicate and presented in a floating bar graph to show the minimum, maximum and mean enrichment factor for each amplicon.

7.2.4 Effect of ectopic *Ter* sites

7.2.4.1 Strain construction

E. coli strains containing ectopic *TerB*, *TerH* or *TerI* in the permissive (P) or non-permissive (NP) orientation were created by Jiri Perutka, Savitri Mandapati and Peter Enyeart in Prof. Andrew Ellington's laboratory (University of Texas at Austin, USA). *Ter* sites were introduced into the chromosome of *E. coli* BL21(*DE3*) (accession number AM946981) by producing *Ter*-targetrons (mobile group II introns carrying *Ter* sequences). Targetrons were designed to insert *Ter* sites in the safe insertion region SIR.5.6 defined by Isaacs et al. (2011) located in the right non-structured chromosome domain using a retrotransposition-activated marker (RAM) constructed in the EcI5 plasmid.

Table 16: Sequence of ectopic *Ter* sites and SIR.5.6.

	Sequence	Position in BL21(*DE3*)	Distance from *oriC* (kbp)
TerB (P)	5'-ACTTTAGTTACAACATACTTATT TGAAATCAATGTTGTATGAATAA-5'	185,367-185,389	929.7
TerH (P)	5'-GAGATAGTTACAACATACGATCG CTCTATCAATGTTGTATGCTAGC-5'	184,460-184,482	928.8
TerH (NP)	5'-CGATCGTATGTNGTAACTATCTC GCTAGCATACANCATTGATAGAG-5'	184,460-184,482	928.8
TerJ (P)	5'-GCATTAGTTACAACTTACTGCGT CGTAATCAATGTTGAATGACGCA-5'	184,460-184,482	928.8
TerJ (NP)	5'-ACNCAGTAAGTTGTAACTAATGC TGNGTCATTCAACATTGATTACG-5'	185,367-184,489	929.7
Sir5.6	ATTGTGCAAATGCCTAAAGGATGATGAAGATGTATGGAGTTGTGG	185,211-185,255	929.6

The SIR.5.6 is located about 930 kbp downstream of *oriC* (right chromosome arm, Table 16). Insertion of *TerB* in the non-permissive orientation was also attempted using the Lambda Red recombination system. Insertion of each *Ter* site was detected by colony PCR and verified by

144

sequencing (Enyeart, P., unpublished). The strains were checked by sequencing and some scrambling was observed for *TerJ* (NP) and *TerH* (NP), i.e. an N at position 3 and 12 in the core sequence respectively (Table 16).

7.2.4.2   Growth curve analysis

The growth curve analysis was performed by Savitri Mandapati in the Ellington laboratory (University of Texas at Austin, USA). BL21(*DE3*) cells carrying ectopic *Ter* sites were grown in LB broth supplemented with chloramphenicol at 37°C and OD_{600} was measured every 5 minutes for 12 hours. The results were plotted as $\log_2(OD_{600})$ versus time (minute). In order to select the linear region of the curve, each point was assigned a correlation coefficient R^2 corresponding to the value of R^2 for the line consisting of that point and the five points before and after. The variance was lower when the same time window was used for all three replicates so the resulting R^2 values were averaged for all three replicates at each time point. The longest stretch in which all these averaged R^2 values were equal to or greater than 0.99 was taken as the linear range. The slope of the least-squares linear fit of the $\log_2(OD_{600})$ curve of each replicate in that time range was then taken as the growth rate and the doubling time was calculated as 1/growth rate.

7.3 Results

7.3.1 Strategy

To determine the *in vivo* distribution of Tus to *Ter* sites, the initial strategy was to immunoprecipitate the endogenous nucleoprotein complexes with chicken IgYs raised against the replisomal proteins Tus, DnaA and SSB. DnaA was to be used as an experimental control and SSB as a marker of replisome dynamics. Unfortunately these IgYs showed high

levels of cross-reactivity and could not be validated for ChIP experiment (data not shown). Consequently, Tus occupancy on the *Ter* sites was analysed using vector-encoded and rhamnose induced expression of Tus-GFP during exponential growth of *E. coli* KRX cells (Figure 27). It was hypothesized that endogenous expression of Tus would be downregulated by the overproduced Tus-GFP binding to the Tus promoter. This in turn would allow the exchange of Tus by Tus-GFP on *Ter* sites. The binding of DnaA-GFP to the *oriC* region was also examined and used as a control experiment since the DnaA binding profile to the chromosome is well characterized (Ozaki and Katayama, 2012). Over-expression of DnaA was shown to repress transcription of chromosome born DnaA (Kucherer et al., 1986) suggesting that here again overexpressed DnaA-GFP could replace endogenous DnaA on the chromosome assuming a comparable activity for both proteins.

For this study, a new ChIP-qPCR method was designed involving formaldehyde-mediated crosslinking of GFP-tagged Tus and DnaA expressed at moderate levels using formaldehyde and glycine, followed by French Press lysis and immunoprecipitation on anti-GFP IgG coated 96-well plates. Following crosslinking reversal, immunocaptured DNA fragments were quantified by qPCR using primers for the ten primary *Ter* and *oriC* regions. The strategy is illustrated in Figure 27. This method by-passes the need for specific beads, overnight incubation or DNA purification prior to qPCR analysis.
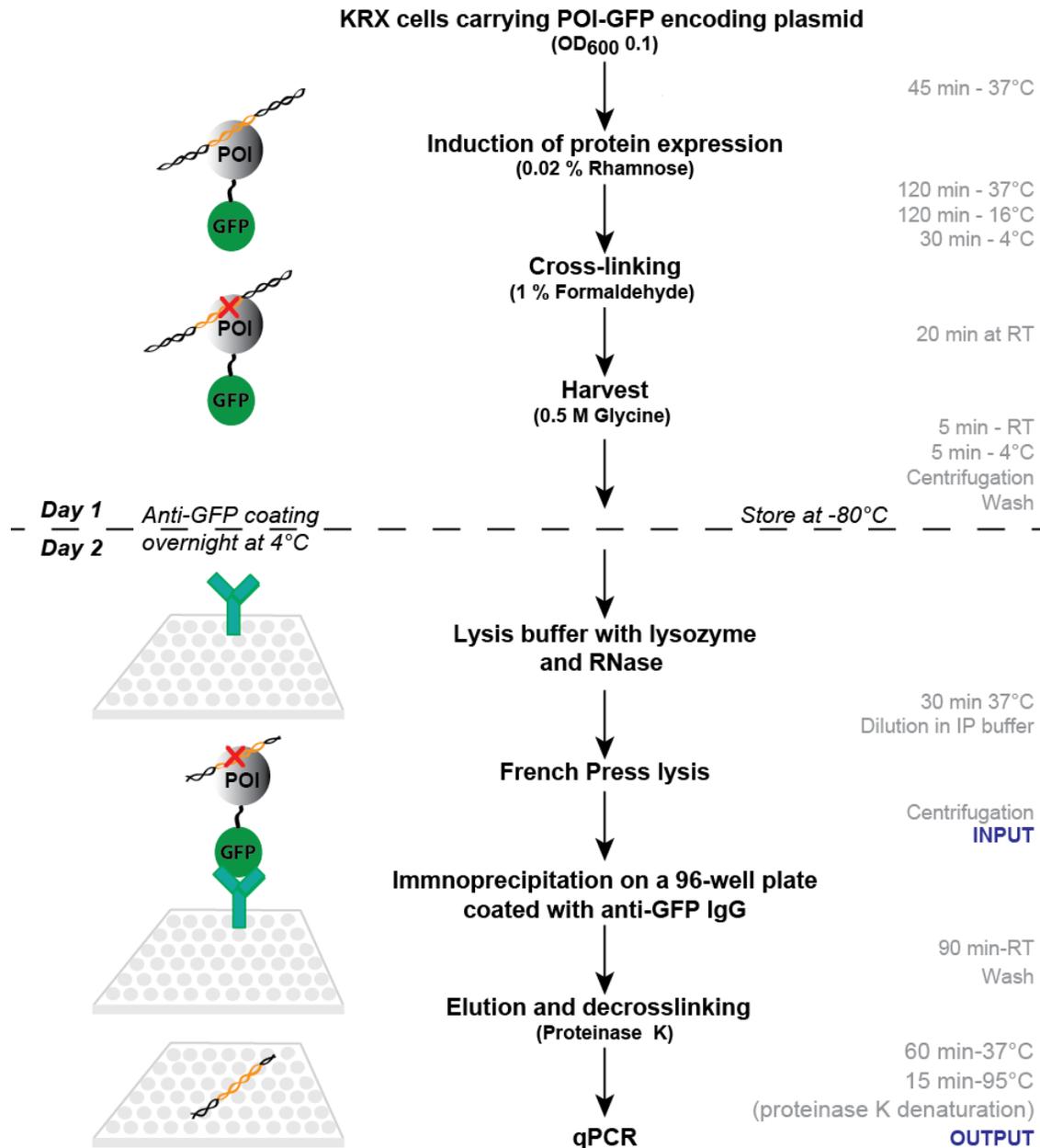
Figure 27: ChIP-qPCR protocol for GFP-tagged proteins. Following crosslinking, cells were resuspended and incubated for 30 minutes in lysis buffer containing lysozyme (10 mM Tris (pH 8), 20 % sucrose, 50 mM NaCl, 10 mM EDTA 1 mg/ml lysozyme and 10μg/ml RNase) and diluted in IP buffer (50 mM HEPES (pH 7.5), 150 mM NaCl and 1 mM EDTA) prior to French Press lysis. Protein-DNA complexes were captured using a commercial anti-GFP antibody (light green) coated on a 96-well plate. Following washes in TCS (50 mM Tris (pH 7.5), 150 mM NaCl and 2 mM KCl), DNA was released in elution and decrosslinking buffer (2 mM Tris, 50 mM NaCl, 0.005 % tween, 300 μg/ml proteinase K). Following proteinase K denaturation (15 minutes at 95°C), co-purified DNA fragments were quantified by qPCR. See section 7.2.3 for further details on the procedure and volumes used.

The use of the same anti-GFP IgG to immunocapture Tus-GFP and DnaA-GFP bound to chromosomal DNA ensures that both protein-DNA complexes were captured with the same efficiency, allowing the comparative analysis of their binding.

7.3.2 Validation

To validate this strategy, the binding activity of Tus-GFP and DnaA-GFP needed to be verified *in vitro* to demonstrate that these epitope-tagged proteins were functional and could potentially bind their targets on the chromosome *in vivo*. The *in vitro* binding activity of Tus-GFP has already been demonstrated several times before, using various methods (*cf* Chapter 4, 5 and 6; Moreau and Schaeffer, 2012b, Moreau and Schaeffer, 2012a). The binding of Tus-GFP to *Ter-lock* sites was in good agreement with *Ter* sites' intrinsic *in vivo* pausing efficiencies observed by Duggin and Bell (2009) suggesting that the GFP domain in the fusion protein has no effect on Tus' functionality. The DnaA domain that was amplified from *E. coli* DH12S genomic DNA, carried a missense mutation (P18S) compared to the MG1655 sequence. *E. coli* DH12S is a derivative of the DH10B strain which has many missense mutations compared to MG1655, including the same mutation in *dnaA* (see Table S2 in Durfee et al. (2008). This mutation is located in the domain I of the protein responsible for oligomerisation and DnaB interactions at *oriC* (Messer, 2002, Weigel et al., 1999). This residue is not conserved across bacterial DnaAs (Felczak et al., 2005, Sutton and Kaguni, 1997) and is therefore not critical for replication initiation activity. The binding activity of DnaA-GFP to ATP, ADP and *oriC* was previously demonstrated using the thermal shift assay DSF-GTP (*cf* Appendix B; Moreau et al., 2012). All three ligands increased DnaA-GFP stability upon binding in the presence of $MgCl_2$ and their combination resulted in cumulative stabilizing effects demonstrating the ability of DnaA-GFP to bind both cofactors and the *oriC* fragment analyzed in this study. However, the *in vivo* activity and efficiency of the GFP-

tagged-proteins Tus and DnaA compared to the endogenous proteins was not verified. ChAP-chip experiments have been performed previously on His$_{12}$-tagged DnaA in *B. subtilis* (Ishikawa et al., 2007) suggesting that the N-terminal His$_6$-tag in the DnaA-GFP fusion protein used in this study should not affect its activity. Felczak et al. (2005) have shown that His$_6$-DnaA functions as wild type DnaA in *E. coli*. However the effect of the C-terminal GFP-tag on DnaA activity *in vivo* has not been verified.

The ChIP-qPCR protocol was first validated with purified and non-crosslinked protein-DNA complexes at low concentration using the qPCR binding assay developed in Chapter 6 (*cf* Figure 25A p130 for the principle of the method). Tus-GFP-*Ter* complexes were successfully captured at 0.4 nM in an affinity dependent manner (*cf* Figure 25 p130; Moreau and Schaeffer, 2012b). The same technique was applied to DnaA-GFP pre-assembled with equimolar amounts of *oriC-* or *TerC-* (negative control) containing DNA at final concentrations of 4 or 0.4 nM each (K_d DnaA-R1 = 10^{-9} M; Schaper and Messer, 1995). The protein-DNA mixture (input) was applied to Maxisorb wells coated with anti-GFP IgG and the fraction of DnaA-GFP bound DNA was quantified by qPCR (Figure 28B). The *oriC* and *TerC* DNA used in this experiment were obtained by PCR amplification of purified DH12S genomic DNA using the same primers used for detection (*cf* Table 4 p44 for primer sequences). The DnaA amplicon spans 115 bp out of the 245 bp-long full origin sequence (Leonard and Grimwade, 2005) and contains the R1 and R5 DnaA-box as well as five ATP-DnaA boxes. They are the minimal elements of DnaA-assembly region (DAR) required for origin unwinding (Figure 28A; Ozaki and Katayama, 2012).
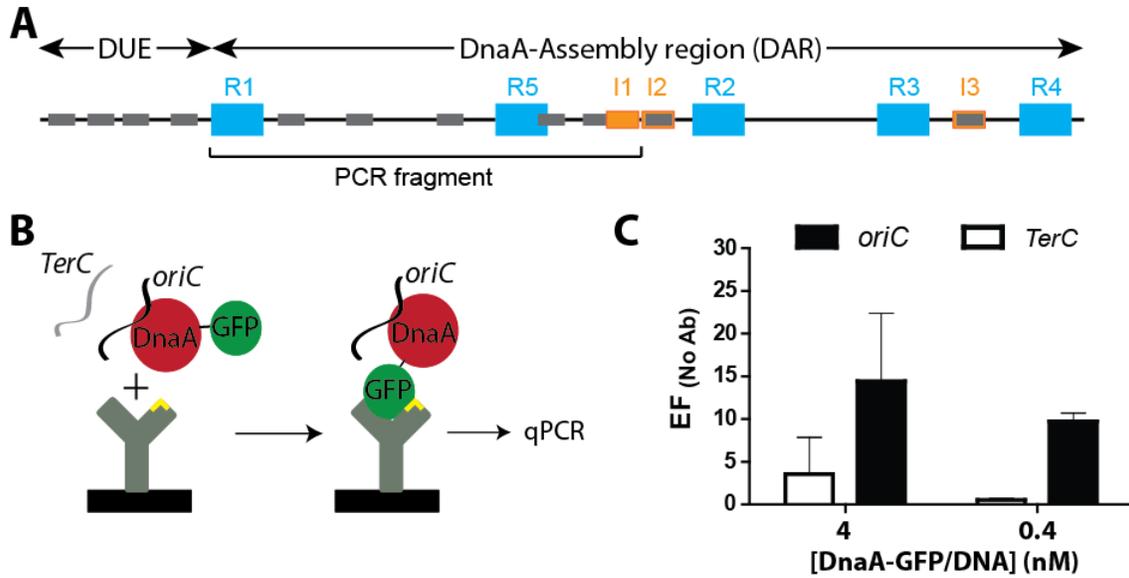
Figure 28: Map of *oriC* and immunoprecipitation of DnaA-GFP-*oriC* complexes. (A) Map of *E. coli* chromosomal origin. Blue boxes are DnaA box sequences recognized by DnaA, smaller orange boxes represent I sites bound to ATP-DnaA. The small grey bars represent GATC sequences recognized by DNA adenine methyltransferase. The *oriC* fragment used for qPCR binding assay and ChIP-qPCR analysis is shown by the black bracket. (B) A mixture of DnaA-GFP, *oriC* and *TerC* (negative control) containing DNA fragments was applied at a final concentration of 4 nM or 400 pM onto anti-GFP coated wells. Following washes, bound DNA was eluted and quantified by qPCR. (C) Enrichment factors relative to the no antibody control ($EF_{(No\ Ab)} = E_{amp}^{(Ct(No\ Ab)\ -Ct(IP))}$ with E_{amp} being the primer specific efficiency obtained from the serial dilution of the input ($E_{amp}=10^{(-1/slope)}$)). E_{amp} for *TerC* and *oriC* were obtained for each replicate experiment. Error bars represent SD (n=2).

The DnaA-GFP bound DNA was captured by anti-GFP IgGs, eluted in high salt and quantified by qPCR. The ΔC_t method was used to determine enrichment factors (EF) relative to the background signal obtained in absence of antibodies ($EF_{(No\ Ab)}$, Figure 28C). Specific binding of DnaA to *oriC* was detected at both 4 and 0.4 nM with 14.4 (± 7.9) and 9.75 (± 0.95) fold enrichment (± SD) at the respective concentrations whereas *TerC* was enriched only by 3.58 (± 4.3) and 0.55 (± 0.15) fold (± SD) at the same concentrations. In this experiment, neither MgCl$_2$ nor ATP was added to the binding reaction implying that DnaA could only bind to high affinity R1 and weaker affinity R5 (K_D > 50 nM; Schaper and Messer, 1995, Weigel et al., 1997) and not to ATP-DnaA boxes (shown in orange in Figure

28A). A higher occupancy of *oriC* by DnaA-GFP is likely to occur *in vivo*. The enrichment of *TerC* at the highest concentration could be the result of non-specific binding of DnaA due to the absence of ATP or ADP (Makise et al., 2002). Overall, these results confirmed the preferential binding of DnaA-GFP to *oriC* and that DnaA-GFP could be immunoprecipitated at picomolar concentrations.

For ChIP experiments, Tus-GFP and DnaA-GFP were expressed separately in two different *E. coli* KRX cultures. In order to achieve near endogenous and comparable levels of protein expression, Tus-GFP and DnaA-GFP expression was induced in KRX cells for 2 hours at 37°C with 0.02 % rhamnose followed by 2 hours at 16°C to increase the proportion of folded proteins (Figure 29A). Control KRX cells were treated similarly. Under these conditions, cells expressing DnaA-GFP or Tus-GFP underwent about 3 and 4 cell divisions respectively (Figure 29A). Proteins were detected by Western blot analysis with a commercially available chicken anti-GFP antibody (Figure 29B-C) and the cellular level of GFP-tagged proteins was estimated from band intensities of known amounts of pure proteins and the cell parameters determined by Volkmer and Heinemann (2011) in LB broth (7.8×10^8 cells.ml^{-1}.OD^{-1} and 4.4 fL cell volume). The cellular concentration of Tus-GFP was roughly estimated to be between 30 and 45 µM corresponding to $70 - 120.10^3$ molecules per cell which is a 1000-fold higher than the endogenous level of Tus (fewer than 100 molecules) reported by Natarajan et al. (1993) and at least 4 orders of magnitude above the K_D of the moderate binders at 150 mM KCl (*cf* Table 12 p91). Because Tus expression is autoregulated, an increase in cellular concentration of plasmid born Tus-GFP above K_D will most certainly repress expression of Tus from the endogenous promoter and will result in saturation of chromosomal *Ter* sites by Tus-GFP. DnaA-GFP was expressed at the estimated cellular concentration of 14 µM corresponding to 37.10^3 molecules per cell which is about 37-fold higher than the level reported by Sekimizu et al. (1988) of 1000 molecules per cell.
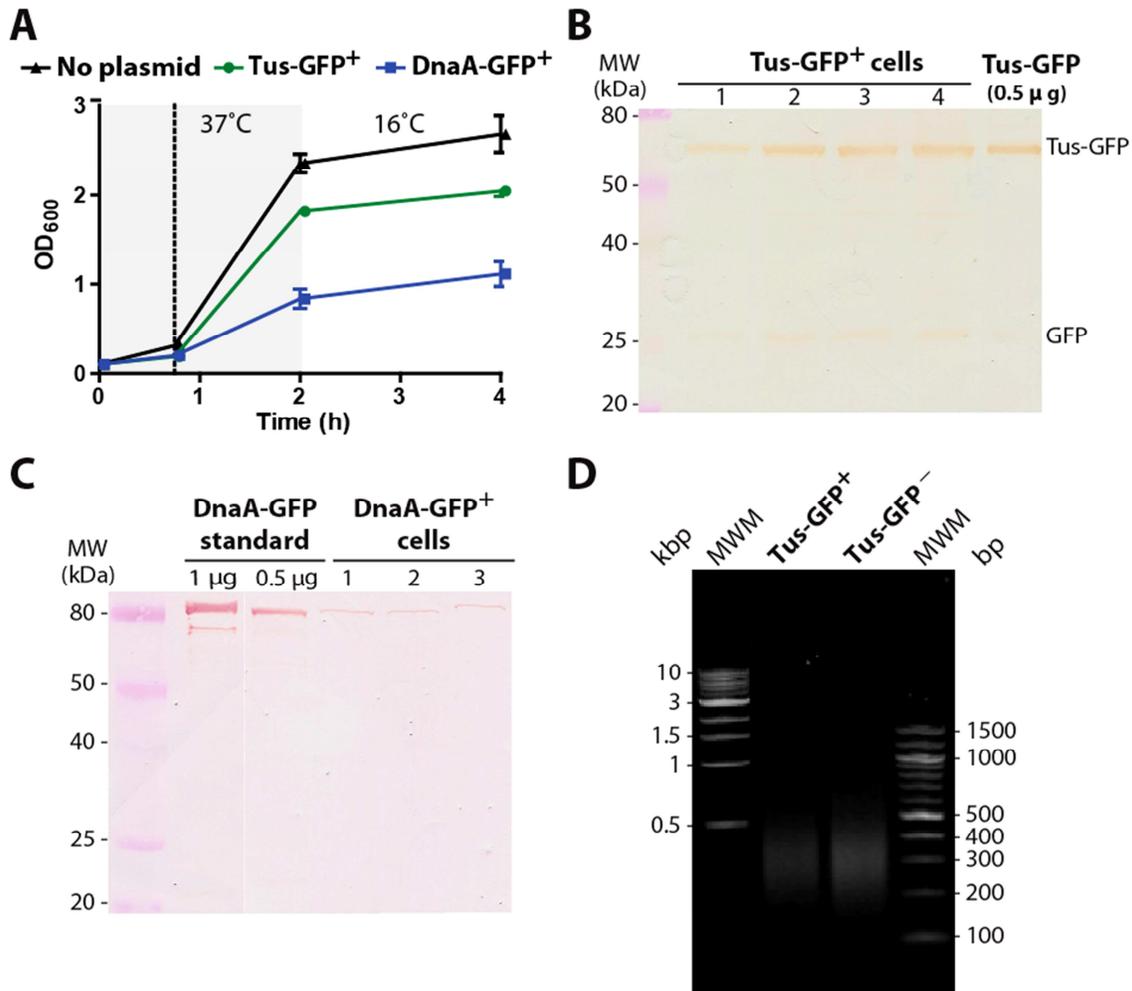
Figure 29: Expression of GFP-tagged DnaA and Tus proteins in *E. coli* KRX cells. (A) Growth curves of KRX control cells (no plasmid), Tus- and DnaA-GFP expressing cells (Tus-GFP⁺ and DnaA-GFP⁺ respectively). At t=0, pre-cultures were used to inoculate fresh media at OD$_{600}$ of 0.1 and grown at 37°C for 45 minutes prior to induction of GFP-tagged protein expression with 0.02 % rhamnose (represented by the dashed line). Cultures were incubated at 37°C for 2 hours (dark grey area) and at 16°C for an additional 2 hours (light grey area). (B) Detection of Tus-GFP proteins in the four independent replicate cultures (1-4) used for ChIP-qPCR by Western blot analysis using a chicken anti-GFP IgY. The last lane contains 0.5 µg of purified Tus-GFP. (C) Detection of DnaA-GFP proteins in the three independent cultures (1-3) used for ChIP-qPCR by Western blot analysis using the same anti-GFP IgY as for Tus-GFP. A DnaA-GFP standard (1 and 0.5 µg of purified proteins) was loaded alongside to estimate the DnaA-GFP expression levels. (D) Distribution of DNA fragments size after French press lysis of crosslinked cells. Example of Tus-GFP expressing cells (Tus-GFP⁺) and KRX control cells (Tus-GFP⁻) used for ChIP. Lysates (inputs) were de-crosslinked with proteinase K (300 µg/ml) for 1 hour at 37°C. Samples were heated for 15 minutes at 95°C and centrifuged to remove denatured proteins. 35 µl of lysate was loaded onto 1 % agarose gel. MWM: molecular weight marker; Quick-Load 1 kb DNA Ladder in lane 1 and 100 bp DNA ladder in lane 4 (NEB).

DnaA-GFP overexpression caused cell elongation and filament formation (data not shown) which has been observed previously as a result of cell division inhibition (Grigorian et al., 2003). Cells with successful expression of GFP-tagged proteins were selected for ChIP-qPCR analysis (Figure 29B-C). Selected cultures were lysed and the DNA was sheared by three passages in a French Press at 12,000 psi. The size distribution of DNA fragments was assessed by examination of de-crosslinked lysates in 1% agarose gel (Figure 29D). This method was highly reproducible and yielded DNA fragments ranging from 150 to 1000 bp in length with a maximum distribution of fragments being ~300 bp in length.

7.3.3 Determination of DnaA-GFP binding to *oriC* by ChIP-qPCR

In *E. coli*, DnaA is known to be bound to *oriC* throughout the cell cycle to high affinity sites like the R1 DnaA box *in vivo* (Cassler et al., 1995, Miller et al., 2009, Nievera et al., 2006, Samitt et al., 1989). Therefore, to verify the sensitivity of the immunoprecipitation protocol, the binding of DnaA-GFP to *oriC* was examined first. KRX cells expressing DnaA-GFP (*DnaA-GFP$^+$*) or lacking DnaAGFP (*DnaA-GFP$^-$*; control) were crosslinked with formaldehyde and homogenized with a French press. The protein-DNA complexes were captured on an anti-GFP coated microplate for 90 minutes and the crosslinking was reversed using proteinase K and heat treatment (Figure 27). DNA fragments co-purified with DnaA-GFP were analysed by qPCR using primers specific for *oriC* and *TerC* (negative control) regions (Table 17 and Figure 30).

Table 17: Mean Ct-values and enrichment factors obtained for *oriC* and *TerC* in *DnaA-GFP*⁺ and *DnaA-GFP*⁻ cells by ChIP-qPCR.

Amplicon (E_{amp})	DnaA-GFP	$_cCt_{(input)}$ (SD)	$Ct_{(output)}$ (SD)		$Ct_{(No\ Ab)}-Ct_{(IP\ sample)}$ (SD)	IP efficiency$_{(TerC)}$ (SD)		$EF_{(No\ AB)}$ (SD)	$EF_{(KRX)}$ (SD)	
			IP	No Ab		IP	No Ab		IP	No Ab
oriC	+	12.43 (0.59)	24.46 (1.31)	28.13 (1.14)	3.67 (0.56)	5.4 (2.9)	1.7 (1.1)	7.6 (2.2)	24.9 (26.4)	1.5 (0.6)
(1.73)	—	11.95 (0.1)	28.86 (1.07)	28.26 (1.13)	-0.60 (2.33)	1.7 (0.1)	1.4 (1.1)	1.0 (1.3)		
TerC	+	17.56 (0.81)	31.98 (1.43)	33.31 (1.59)	1.33 (0.93)			2.3 (1.3)	9.4 (12.1)	2.1 (2.4)
(1.76)	—	16.39 (0.61)	33.65 (1.24)	32.57 (0.65)	-1.08 (2.4)			1.3 (1.6)		

E_{amp} is the primer specific amplification efficiency ($E_{amp}=10^{(-1/slope)}$). Ct-values obtained for *oriC* and *TerC* in the diluted input were corrected for the dilution factor using primer specific efficiencies (E_{amp}) to yield $_cCt_{(input)}$. IP: immunoprecipitated; No Ab: no antibody; IP efficiency$_{(TerC)}$: input fraction of *oriC* enriched relative to the non-specific *TerC* DNA region; $EF_{(No\ Ab)}$: enrichment factor relative to the no antibody control; $EF_{(KRX)}$: enrichment factor relative to *DnaA-GFP*⁻ control cells (n=3).

The use of the French Press allowed highly reproducible cell lysis as demonstrated by the small standard deviation of $_c$Ct-values obtained for each DNA fragment of interest in the input of replicate IP experiments (see SD in Table 17 and section 7.2.3.5 for calculation procedure). In replicating cells, the copy number of *oriC* should be larger than the copy number of *TerC* as a result of multiple initiations of DNA replication events. The difference in $_c$Ct-values obtained for *TerC* and *oriC* supported this hypothesis (see $_c$Ct$_{(input)}$ in Table 17). The *oriC* region was successfully enriched in *DnaA-GFP*$^+$ cells as demonstrated by an increase in Ct-values of ~3.7 cycles compared to the no antibody control sample and little to no enrichment of *TerC* (Figure 30A and Table 17). The fold enrichments of *oriC* and *TerC* were analyzed using three different normalization procedures to demonstrate that the same enrichment profiles were obtained regardless of the normalization method.

First, the IP efficiency of *oriC* was normalized against the IP efficiency of the *TerC*-containing region (Figure 30B) to highlight the fold enrichment relative to a non-specific DNA region. The IP efficiency is the fraction of a given DNA species present in the output relative to its concentration in the initial input and was calculated using the efficiency (E_{amp}) of each primer pair (Table 17). An E_{amp} value of 2 corresponds to an amplification efficiency of 100 %, and a value of 1.76 corresponds to an efficiency of 76 %. The immunoprecipitation of the *oriC* region was between 3.2 and 8.8-fold more efficient than the non-specific *TerC* region while background values were less than 2-fold on average (Figure 30B and Table 17). Second, the enrichment factors were normalized against the background obtained in the absence of anti-GFP IgG (EF$_{(No\ Ab)}$) and third, against the *DnaA-GFP*$^-$ strain (EF$_{(KRX)}$). The EF$_{(No\ Ab)}$ obtained for *oriC* was between 5.1 to 9.1-fold whereas *TerC* enrichment was close to background values (EF$_{(No\ Ab)}$ of *TerC* between 1.7 and 3.9-fold in *DnaA-GFP*$^+$ cells and between 0.2 and 2.6-fold in *DnaA-GFP*$^-$ cells, Figure 30C and Table 17).
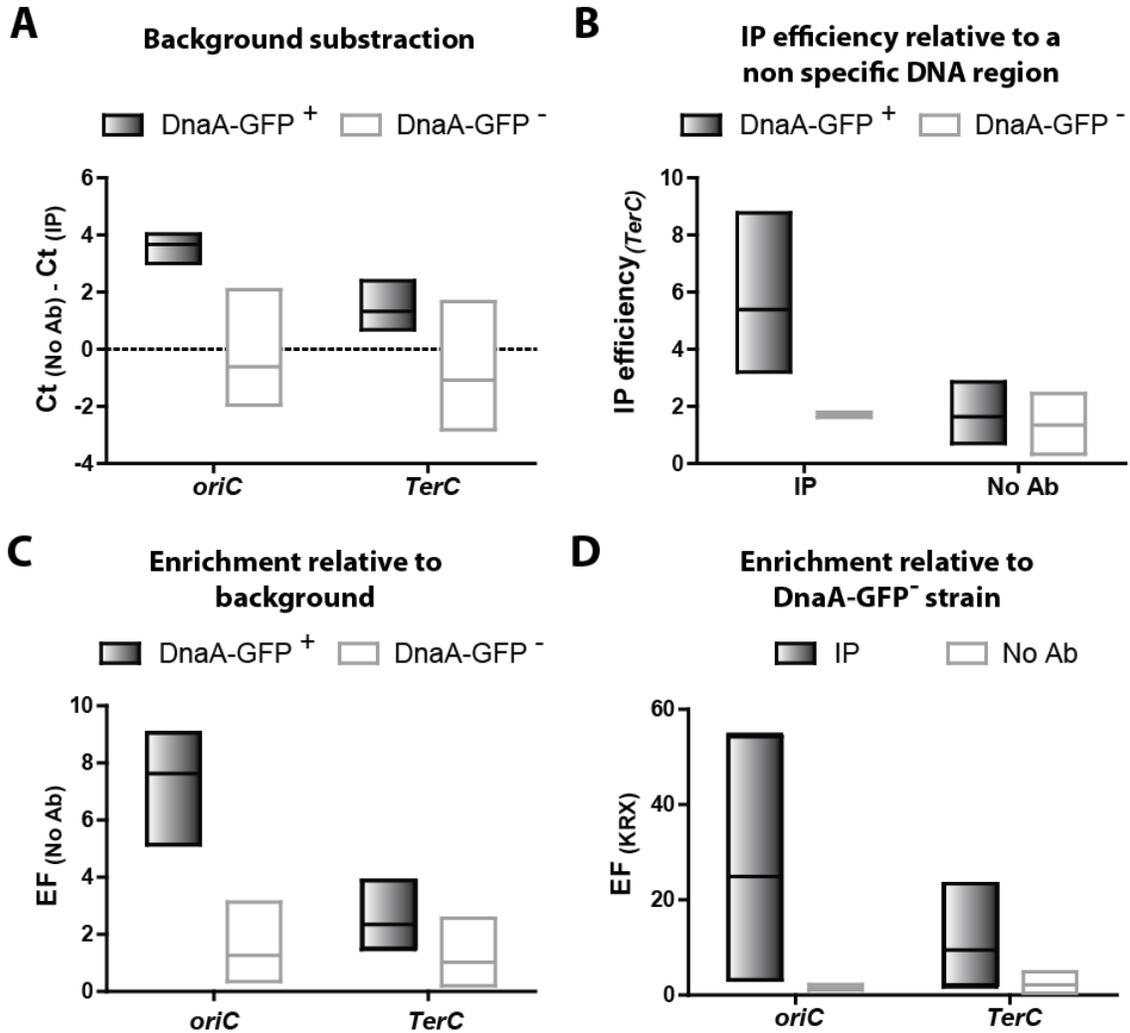
Figure 30: Binding of DnaA-GFP to *oriC* in *E. coli* KRX cells by ChIP-qPCR. The floating bars represent the minimum, maximum and mean values (mean values are reported in Table 17 along with SD, n=3). (A) Difference in Ct-values between immunoprecipitated DNA (IP) and background signal in absence of anti-GFP IgG (No Ab) obtained for DnaA-GFP expressing cells (DnaA-GFP$^+$) or KRX cells lacking the DnaA-GFP encoding plasmid (DnaA-GFP$^-$). (B) IP efficiency of *oriC* relative to the non-specific *TerC* region (IP efficiency$_{(TerC)}$=(E$_{amp}$$^{(cCt(input)-Ct(output))oriC}$/E$_{amp}$$^{(cCt(input)-Ct(output)TerC}$)) obtained for DnaA-GFP$^+$ and DnaA-GFP$^-$ cells in the presence (IP) or absence of anti-GFP IgG antibody (No Ab). (C) Enrichment factor relative to the no antibody control of *oriC* and *TerC* (EF$_{(No Ab)}$=E$_{amp}$$^{(Ct(No Ab)-Ct(IP))}$). (D) Enrichment factor of *TerC* and *oriC* relative to DnaA-GFP$^-$cells (EF$_{(KRX)}$=E$_{amp}$$^{((cCt(DnaA-GFP+ input)-Ct(DnaA-GFP+ output))-(cCt(DnaA-GFP- input)-Ct(DnaA-GFP- output))}$).

After subtracting the background obtained with *DnaA-GFP$^-$* cells (EF$_{(KRX)}$, Figure 30D), *oriC* was enriched by 3.1 to 54-fold more than in *DnaA-GFP$^-$* cells and *TerC* was enriched between 1.7 and 23.4-fold. With the last two normalization methods, *oriC* was enriched

between 2 to 6-fold more than *TerC* when comparing each biological replicates (data not shown). Despite low enrichment values, this method could detect the preferential binding of exogenous DnaA-GFP to *oriC* and suggested that the capture of Tus-*Ter* complexes would be possible with this method.

7.3.4 Distribution of Tus-GFP on *Ter* sites

The binding of Tus-GFP to chromosomal *Ter* sites was investigated by ChIP-qPCR in exponentially growing cells expressing plasmid-born Tus-GFP following the same protocol as for DnaA-GFP (Figure 27). An additional thermal denaturation step was included prior to immunoprecipitation at the temperature at which free Tus-GFP denatures (50°C) in order to remove unbound Tus-GFP. The Tus-GFP associated DNA fragments were quantified by qPCR using primer pairs specific for the ten primary *Ter* sites and *oriC* (negative control).

Most primary *Ter* sites were significantly enriched by the immunoprecipitation step as demonstrated by an increase in ΔCt-values of ~6 to 7.7 cycles compared to the no antibody control sample, with the exception of *TerF* and *TerJ* that had ΔCt-values of only 2.4 and 4.54 respectively (see values for $Ct_{(No\ Ab)}$-$Ct_{(IP)}$ in Table 18 and Figure 31A). As expected, no enrichment could be observed for the *Tus-GFP^{-}* control cells (Figure 31A). The data were analyzed using three normalization procedures as for DnaA-GFP ChIP experiment and demonstrated similar enrichment profiles although variable in amplitude (Figure 31B-D).

Table 18: Mean Ct-values and enrichment factors obtained for the ten primary *Ter* sites and *oriC* in *Tus-GFP*$^+$ and *Tus-GFP*$^-$ cells by ChIP-qPCR.

Amplicon (E$_{amp}$)	Tus-GFP	$_c$Ct$_{(input)}$ (SD)	Ct$_{(output)}$ (SD)		Ct$_{(No\ Ab)}$ −Ct$_{(IP\ sample)}$ (SD)	IP efficiency$_{(oriC)}$ (SD)		EF$_{(No\ Ab)}$ (SD)	EF$_{(KRX)}$ (SD)	
			IP	No Ab		IP	No Ab		IP	No Ab
TerA	+	15.71 (0.68)	24.13 (1.58)	31.72 (2.36)	7.59 (0.98)	32.3 (19.5)	0.9 (0.8)	129 (113)	249 (85.5)	2.9 (1.7)
(1.79)	—	15.48 (0.34)	33.27 (0.80)	32.84 (1.37)	-0.44 (0.57)	0.7 (0.3)	0.7 (0.5)	0.9 (0.7)		
TerB	+	16.40 (0.82)	26.20 (1.55)	33.57 (1.94)	7.37 (0.95)	27.2 (16.9)	0.7 (0.5)	67.9 (60.4)	256 (222)	4. 5 (3.1)
(1.69)	—	15.78 (0.31)	35.60 (0.96)	35.26 (1.41)	-0.34 (0.62)	0.9 (0.4)	0.6 (0.3)	1.0 (0. 8)		
TerC	+	16.80 (0.52)	26.21 (1.70)	33.25 (2.25)	7.05 (1.15)	21.7 (12.3)	0.6 (0.6)	95.9 (109)	260 (131)	3.7 (1.7)
(1.76)	—	15.92 (0.59)	34.97 (0.73)	34.48 (1.51)	-0.49 (0.64)	0.7 (0.3)	0.4 (0.3)	0.9 (0.8)		
TerD	+	14.71 (0.86)	26.00 (1.88)	32.99 (2.33)	7.00 (1.10)	17.6 (12.8)	0.7 (0.5)	48.3 (47.8)	150 (114)	3.3 (2.1)
(1.64)	—	14.25 (0.43)	35.15 (1.09)	34.41 (1.58)	-0.74 (0.62)	0.8 (0.1)	0.7 (0.4)	0.8 (0.6)		
TerE	+	17.71 (0.59)	27.44 (1.57)	35.10 (2.24)	7.67 (1.15)	23.3 (13.5)	0.6 (0.3)	106 (109)	289 (288)	3.5 (1.8)
(1.72)	—	16.89 (0.34)	36.35 (0.75)	36.34 (1.82)	-0.02 (0.66)	0.6 (0.2)	0.4 (0.2)	1.2 (0.7)		
TerF	+	13.59 (0.56)	27.79 (0.70)	30.22 (2.05)	2.43 (1.08)	1.0 (0.2)	0.4 (0.1)	6.9 (6.3)	18.4 (17.9)	2. 7 (1.6)
(1.79)	—	13.17 (0.58)	31.82 (1.18)	31.14 (1.70)	-0.68 (0.42)	0.4 (0.1)	0.4 (0.1)	0.7 (0.3)		

Amplicon (E$_{amp}$)	Tus-GFP	$_c$Ct$_{(input)}$ (SD)	Ct$_{(output)}$ (SD)		Ct$_{(No\ Ab)}$ −Ct$_{(IP\ sample)}$ (SD)	IP efficiency$_{(oriC)}$ (SD)		EF$_{(No\ Ab)}$ (SD)	EF$_{(KRX)}$ (SD)	
			IP	No Ab		IP	No Ab		IP	No Ab
TerG	+	13.19 (0.17)	23.14 (1.95)	29.73 (1.81)	6.60 (0.82)	15.8 (10.3)	0.7 (0.4)	62.3 (69.5)	155 (111)	3.1 (1.4)
(1.77)	—	12.54 (0.67)	30.97 (0.74)	30.86 (1.04)	-0.11 (0.55)	0.7 (0.4)	0.5 (0.3)	1.1 (0.7)		
TerH	+	12.29 (0.51)	24.68 (1.47)	31.07 (1.83)	6.40 (0.86)	6.4 (2.0)	0.5 (0.2)	38.2 (36.2)	89.7 (78.6)	3.3 (2.8)
(1.68)	—	11.75 (0.67)	32.35 (1.12)	32.30 (1.99)	-0.04 (0.78)	0.4 (0.2)	0.3 (0.1)	1.3 (1.3)		
TerI	+	12.10 (0.50)	24.11 (1.46)	30.13 (1.67)	6.02 (0.69)	5.7 (2.1)	0.5 (0.3)	33.8 (27.2)	82.9 (65.5)	2.4 (1.9)
(1.73)	—	11.63 (0.63)	31.26 (1.72)	30.90 (1.51)	-0.37 (0.73)	0.4 (0.2)	0.4 (0.1)	1.04 (0.9)		
TerJ	+	11.91 (0.64)	25.74 (1.65)	30.28 (1.90)	4.54 (1.07)	1.9 (1.6)	0.5 (0.2)	16.7 (15.1)	58.0 (68.7)	3.3 (2.1)
(1.70)	—	11.87 (1.50)	31.86 (1.09)	31.47 (1.33)	-0.38 (0.42)	0.4 (0.3)	0.5 (0.2)	0.9 (0.3)		
oriC	+	13.01 (0.70)	28.14 (0.67)	29.27 (2.15)	1.13 (0.97)			2.7 (2.3)	7.7 (8.7)	2.8 (2.1)
(1.73)	—	12.06 (0.77)	30.20 (1.56)	29.56 (2.14)	-0.64 (0.46)			0.8 (0.4)		

E$_{amp}$ is the primer specific amplification efficiency (E$_{amp}$=10$^{(-1/slope)}$). Ct-values obtained for *Ter* sites and *oriC* in the diluted input were corrected for the dilution factor using primer specific efficiencies (E$_{amp}$) to yield $_c$Ct$_{(input)}$. IP: immunoprecipitated; No Ab: no antibody; IP efficiency$_{(oriC)}$: input fraction of *TerC* enriched relative to the non-specific *oriC* DNA region; EF$_{(No\ Ab)}$: enrichment factor relative to the no antibody control; EF$_{(KRX)}$: enrichment factor relative to *Tus-GFP$^-$* cells (n=4).

The ratio of IP efficiency of *Ter* sites relative to the IP efficiency of a non-specific DNA (i.e. *oriC*) is likely to be the most accurate normalization method as it takes into account differences in concentrations in the input between the different DNA regions analyzed. Indeed, the concentration of each analyzed DNA region in the input sample decreases with its distance from *oriC* (Table 18). With the $EF_{(No\ Ab)}$ method, the background signal varies for each DNA region due to their variation in concentration in the input sample, thus affecting the resulting enrichment factors. The $EF_{(KRX)}$ method does take into account the differences in the initial concentration of each DNA region but some variability may arise from the handling of two different cell cultures. A floating bar diagram was chosen to show the range of enrichment factor obtained with each method (Figure 31). Despite the variation in amplitude, the three binding profiles of Tus-GFP to the ten primary *Ter* sites were all relatively similar to the binding affinity profile previously obtained for fully double-stranded *Ter* sites (*cf* Figure 15 p86; Moreau and Schaeffer, 2012a). A higher enrichment of the strong binders (*TerA-E* and *TerG*) was obtained compared to the outermost moderate binders, with *TerF* and *J* being systematically the least enriched *Ter* sites (Figure 31B-D). The strong binders were enriched between 10 and 60-fold above the non-specific *oriC* region with *TerD* and *TerG* being the least enriched in this group (Figure 31B, Table 18) as also observed in Figure 31C and D.
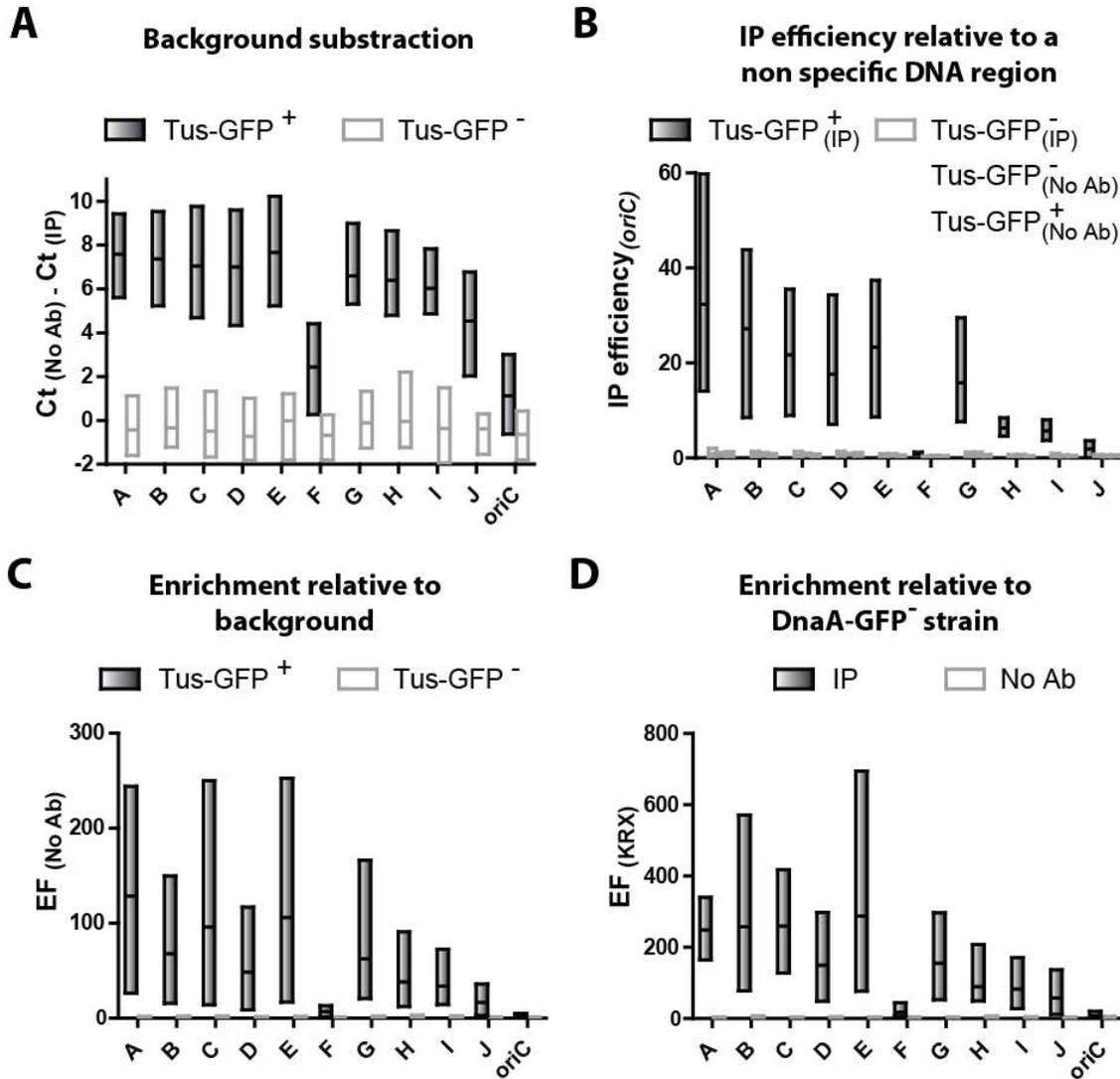
Figure 31: Binding of Tus-GFP to *Ter* sites in *E. coli* KRX cells by ChIP-qPCR. The floating bars represent the minimum, maximum and mean values (mean values are reported in Table 18 along with SD, n=4). (A) Difference in Ct-values between immunoprecipitated DNA (IP) and background signal in absence of anti-GFP IgG (No Ab) obtained for Tus-GFP expressing cells (Tus-GFP⁺) or Tus-GFP⁻ control cells. (B) IP efficiency of *Ter* sites relative to the non-specific *oriC* DNA region (IP efficiency$_{(oriC)}$=(E$_{amp}$$^{(cCt(input)-Ct(output))Ter}$/E$_{amp}$$^{(cCt(input)-Ct(output)oriC)}$)) obtained for Tus-GFP⁺ and Tus-GFP⁻ cells in the presence (IP) or absence of anti-GFP IgG antibody (No Ab). (C) Enrichment factor relative to the no antibody control of *Ter* sites and *oriC* (EF$_{(No\ Ab)}$=E$_{amp}$$^{(Ct(No\ Ab)-Ct(IP))}$). (D) Enrichment factor relative to KRX control cells (EF$_{(KRX)}$=E$_{amp}$$^{(Ct(Tus-GFP+\ output)-cCt(input))-(Ct(Tus-GFP-\ output)-cCt(Tus-GFP-\ input))}$) of *Ter* sites and *oriC*.

TerH and *TerI* were enriched with comparable efficiencies with only 6.4 and 5.7-fold above

oriC) which was ~3-fold higher than the enrichment level of *TerJ* of 1.9-fold (Figure 31B and

161

Table 18). *TerF* could not be enriched compared to *oriC* but $EF_{(No\ Ab)}$ and $EF_{(KRX)}$ values were between 7 to 9-fold above background and 2.5-fold above *oriC* (Figure 31C-D and Table 18). The $EF_{(No\ Ab)}$ values obtained for *TerH* and *TerI* were only slightly lower than *TerD* but between 1.8 and 3.8-fold lower than the remaining strong *Ter* sites. *TerJ* was enriched ~2-fold less than *TerH* and *TerI* (Figure 31C). A similar profile was obtained for $EF_{(KRX)}$ values, only the amplitude of the signal-to–noise ratio increased by a factor of 2 due to a 2-fold lower background signal obtained with strains lacking the target protein.

7.3.5 Effect of ectopic *TerB*, *TerH* and *TerJ*

To determine the effect of TT-lock strength on replication dynamics, a strong TT-lock forming site (*TerB*), a non-TT-lock forming site (*TerH*) and a moderate TT-lock forming site (*TerJ*) were inserted in the right chromosome arm of *E. coli* strain BL21(*DE3*), 930 kbp downstream to *oriC* (right arm) in the permissive (P) or non-permissive (NP) orientations. Although most *Ter* sites could be inserted in both orientations with good efficiencies, *TerB* could not be inserted in the non-permissive orientation using either a retrotransposon activated marker (RAM) or the Lambda Red recombination system (Poteete, 2001). This indicates that the presence of an ectopic *TerB* in the non-permissive orientation in a strain carrying the wild type *tus* gene is lethal for the bacteria as a result of fork blockage 930 kbp downstream to *oriC*. The effect of the insertion direction of the remaining ectopic *Ter* sites was investigated further on *E.coli* growth rates (Figure 32).
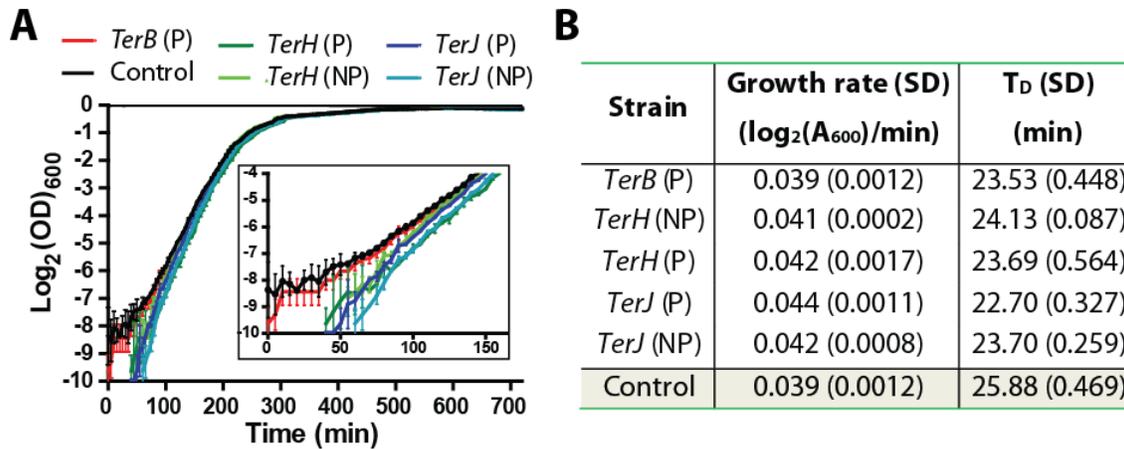
Figure 32: Effect of ectopic *Ter* sites on the growth rate of *E. coli* BL21(*DE3*). *TerB*, *TerH* and *TerJ* were inserted ~ 930 kpb downstream to *oriC* in the permissive (P) or non-permissive (NP) orientation. (A) Growth rates were measured in independent triplicates. Error bars represent SD. A culture of wild type BL21(*DE3*) was grown as a control. (B) Averaged growth rates obtained from the slope of the linear regression performed between 100 and 210 minutes (see section 7.2.4.2) and averaged doubling time (T_D) were obtained as 1/growth rate (n=3, except for *TerH* (NP), n=2).

No difference in growth rates could be observed in the exponential growth phase between these strains. Furthermore, all strains reached the same plateau as the control strain (Figure 32A-B) suggesting that *TerH* and the TT-lock *TerJ* are not sufficient to induce fork arrest.

7.4 Discussion

7.4.1 A fast microplate-based ChIP assay

A modified microplate-based ChIP-qPCR method was developed to quantify the *in vivo* binding of two replisomal proteins, Tus and DnaA, to their respective targets. As for Matrix-ChIP (Flanagin et al., 2008), the entire procedure from immunoprecipiation to PCR-ready ChIP DNA acquisition is performed in a 96-well plate reducing sample handling and transfers, hence increasing reproducibility. Capture and analysis of crosslinked protein-DNA complexes could be performed in less than 4 hours after cell lysis and DNA shearing. The use of French press lysis instead of sonication enabled very efficient cell lysis and DNA shearing,

yielding a majority of DNA fragments in the 300 bp range. Protein-DNA complexes were captured using surface-immobilized anti-GFP antibodies. The use of the anti-GFP IgG as capture antibody ensures that all complexes are immunoprecipitated with the same efficiency (i.e. the epitope is not hidden by ligand binding). An alternative method using streptavidin-coated plates with biotinylated anti-GFP IgGs was found to be equally efficient but much more costly and therefore abandoned (data not shown). Using moderate expression levels of GFP-tagged proteins in *E. coli* cells, it was possible to immunocapture their respective targets using this technique.

The ChIP-qPCR assay allowed the determination of the global distribution of Tus to *Ter* sites in exponentially growing cells. The enrichment factor data obtained for *Ter* sites reflected their binding frequency and affinity assuming that the crosslinking efficiency was the same for the ten complexes. Taking into account the average ΔCt-values between outputs and input, about 0.03 to 1 % of each target in the input DNA was specifically captured by Tus and 0.3 % by DnaA which was at least 10-fold above background. However, the low affinity *TerF* site was enriched just above background (similar to *oriC* in terms of IP efficiency) and was therefore close to the detection limit of this method. In total, 4.9 % of input DNA was captured by Tus-GFP when combining all the ten *Ter* sites and their enrichment profile was similar to the affinity profile previously determined (Figure 31 and Figure 15 p86) demonstrating the reliability and sensitivity of this method.

7.4.2 Distribution of Tus in the fork trap

Transcription and translation (coupled in bacteria) of Tus occurs once Tus has been dissociated from *TerB* by the anti-clockwise replisome and *TerB* has been replicated (Natarajan et al., 1991, Neylon et al., 2005, Roecklein et al., 1991, Roecklein and Kuempel, 1992). Transcription would rapidly cease once the newly synthesized Tus bind to old and

new *TerB*. One study determined the transcript level of Tus during the cell cycle in synchronized cells and could not detect a significant change of expression during the cell cycle (Zhou et al., 1997). It was suggested that the change in transcript level was too small to be detected. The negative auto-regulation of Tus expression coupled with the weakness of the Tus promoter and the strength of the *Tus-TerB* complex suggest that Tus expression may occur as a short burst when Tus is dissociated from *TerB*. In this condition, the innermost and strong *Ter* sites, especially *TerB*, are likely to be occupied first due to their proximity to the *tus* gene and their affinity. Thus, the remaining *Ter* sites would be occupied as a function of free Tus concentration. The endogenous Tus concentration is estimated to be between 20 and 100 nM (Natarajan et al., 1993, Roecklein and Kuempel, 1992) which is at least 3 to 15 times higher than the K_D of moderate binders determined at 150 mM KCl (*cf* Table 15 p132; Moreau and Schaeffer, 2012b). Hence, the moderate *Ter* sites should be occupied for a significant portion of the cell cycle ($t_{1/2}$ of *TerH-J* range from 20 to 6 minutes, Figure 16C). The termination sites where the forks most frequently merge (*TerC* and *TerA*; Duggin and Bell, 2009) should be bound during almost the entirety of the cell cycle due to the long $t_{1/2}$ and locking of Tus on these sites. It could reasonably be anticipated that *TerC*, *TerA* and *TerB* would be the most enriched *Ter* sites compared to the remaining *Ter* sites by ChIP-qPCR analysis.

The strong binders (*TerA-E*, *G*) were bound to similar levels and only 3 to 6 times more than the remaining moderate and outermost *Ter* sites (Figure 31B). Tus-GFP was only marginally bound to *TerF* if at all (Figure 31B-D). The distribution of Tus-GFP on the ten chromosomal primary *Ter* sites, determined in conditions of moderate expression of Tus-GFP, matched the affinity profile for double-stranded *Ter* sites that has previously been obtained by GFP-Basta and qPCR binding assay (*cf* Figure 25 p130; Moreau and Schaeffer, 2012b, Moreau and Schaeffer, 2012a). However, the difference in EF between the strong and

moderate binders was significantly reduced compared to their difference in affinity (*cf* Table 15 p132 and Figure 31B; Moreau and Schaeffer, 2012b). This is not surprising in presence of high levels of Tus-GFP relative to endogenous Tus levels. Indeed, due to the fact that Tus-GFP is in competition with endogenous Tus, Tus-GFP is more likely to replace Tus on lower affinity and fast dissociating *Ter* sites than on the strong affinity sites. This would certainly lead to a reduction in the difference in EF expected between these two *Ter* groups. Nevertheless, the large excess of Tus-GFP compared to Tus should allow Tus-GFP to compete for binding on the strong *Ter* sites after removal of Tus by the replisome.

The data demonstrate that all *Ter* sites can be bound by Tus. The existence of additional factors contributing to the distribution of Tus on *Ter* sites such as the effect of adjacent sequences (*cf* Chapter 6; Moreau and Schaeffer, 2012b), DNA supercoiling (i.e. site availability) can be definitely ruled out. This raises the question, why *TerE* was enriched to similar levels than the other strong binders despite the fact that no pausing has been observed at this site *in vivo* (Duggin and Bell, 2009)? Two hypotheses had previously been proposed by Duggin and Bell (2009), i.e. either *TerE* is non-functional or it does not encounter a replisome at the non-permissive face. The data obtained here support the second hypothesis and therefore suggest that forks never break through *TerA* and *TerD* (see below for discussion on the pausing at *TerH* and *TerI* observed by Duggin and Bell (2009).

Overall, the data suggest that all *Ter* sites are functional although differently occupied and that under natural conditions, moderate *Ter* sites are likely to be significantly less occupied than strong binders. Their occupancy depends on free Tus concentration and Tus' ability to be recycled on these sites. Surprisingly, no correlation between occupancy and replication fork stalling or pausing could be inferred. More work is necessary to clarify the distribution of Tus on *Ter* sites in natural conditions and synchronized cells. This new assay will be invaluable for such experiments.

7.4.3 Is there a role for the weak *Ter* sites?

Of the ten primary *Ter* sites, *TerF* was again found to be the least bound by Tus. This confirms the low affinity of *TerF in vivo* as suggested by previous *in vitro* data (*cf* Chapter 4 and 6; Moreau and Schaeffer, 2012b, Moreau and Schaeffer, 2012a). Duggin and Bell (2009) observed Y-forked structures at *TerF* only when Tus was overproduced, further raising questions about its occupancy under endogenous levels of Tus expression. The ChIP-qPCR data obtained with a moderate expression level of Tus-GFP showed that, chromosomal *TerF* was enriched 18-fold less than the strong binders and just 2.5-fold more than the non-specific *oriC* (Figure 31C-D*)*. The small specific enrichment observed could be the result of a transient yet specific interaction of Tus for this site. This could explain why pausing was observed at *TerF* when Tus was overexpressed to 5 % of the cellular protein content. Given that Tus-GFP was expressed at concentrations above its K_D for *TerF* determined by SPR at 150 mM KCl (*cf* Figure 16 p90, (Moreau and Schaeffer, 2012a), *TerF* is likely to be only transiently if at all occupied by Tus in wild type cells and may not have a significant role in terms of fork pausing. The sequence similarity of all *Ter* sites does not strictly imply that they all perform the same role or function. A function of *TerF* could reside in the coordination of replication and cell division based on its position within the *rsc* gene whose product controls capsule synthesis and the cell division control gene *ftsZ*; but this remains to be demonstrated.

The most recently discovered *Ter* sites (*TerK*, *L, Y* and *Z*) share the deleterious mutations present in *TerF*, *H, I* and *J* and harbor additional mutations (Figure 26A). Their combination suggests that Tus will bind with even lower affinity to these sites based on the destabilizing effects of each of these mutations on binding kinetics (*cf* Table 13 p96). The occupancy of these sites is therefore suspected to be lower than the occupancy of *TerF*. This is further supported by the fact that two of these sites are oriented to oppose origin-to-terminus moving forks (*TerY* and *Z*) and must be weakly bound by Tus since the occurrence

of a similarly oriented strong TT-lock forming site so close to *oriC* would trigger homologous recombination and SOS induction for cell viability (Figure 32; Bidnenko et al., 2002, Bidnenko et al., 2006). Therefore, these weak *Ter* sites are unlikely to have a role in the inhibition of *oriC*-initiated replication forks (see below).

The surprising observation of Y-forked species at *TerH* but not at *TerE* under endogenous levels of Tus expression (Duggin and Bell, 2009) did not correlate with their expected occupancy by Tus *in vivo*. Therefore, the question still remains, how can *TerH* pause a replication fork although *TerE* that is positioned upstream of *TerH* can not?

7.4.4 Moderate TT-locks cannot stop DNA replication

TerB has previously been inserted ectopically at several positions in the genome either to determine a link between DNA replication termination and the cell cycle or to examine the structure and recombination events occurring at forks blocked at a Tus-*TerB* complex. In one study, two *TerB* have been introduced back to back between *TerC* and *TerA*, preventing forks to replicate a 2 kbp region (Sharma and Hill, 1995). In another study, two *Ter* sites were introduced so that both forks were arrested at midway (Bidnenko et al., 2002). In all cases the viability of mutant strains was strictly dependent on homologous-recombination pathways (Bidnenko et al., 2002, Esnault et al., 2007). It was shown that blocked forks are stable until the arrival of a second round of replication which then triggers the RecA and RecBC pathways for homologous-recombination (Bidnenko et al., 2002). In an attempt to determine the respective strength of *Ter* sites, Esnault et al. (2007) inverted large chromosomal domains (0.4 to 1 Mbp) containing either *TerE*, *TerHI*, *TerI* or *TerJ* and examined viability, cell size and nucleoid morphology (Esnault et al., 2007). It was shown that *TerE* severely impacted cell viability, and the rare successful recombinants were RecA-dependent and induced the SOS response. Thus, the ChIP-qPCR data obtained for *TerE* suggest that this site is bound by

Tus *in vivo* and is able to stop the replisome. This further confirms that the absence of pausing observed by Duggin and Bell (2009) at *TerE* is due to the fact that it never encounters a replisome approaching towards its blocking face.

The inversion of large chromosomal regions containing either *TerI* or *TerJ* had no effect on growth but the inversion of *TerHI* induced mild growth defects (Esnault et al., 2007). Since the *TerHI* inverted region encompasses the *oriC*-distal end of the right non-specific domain and the *oriC*-proximal end of the right domain (i.e. *TerHI* is at the interface of the two chromosomal domains, Figure 26A) the growth defect observed in the inverted *TerHI* strain could also be attributed to detrimental non-replication associated events such as the perturbation of chromosome organization. In order to rule out this possibility and evaluate the strength of the TT-lock *in vivo*, a representative of each *Ter* site category (i.e. strong (*TerB*), moderate (*TerJ*) and non-TT-lock forming sites (*TerH*) was inserted in the right chromosome arm of *E. coli recA$^+$* strain BL21(*DE3*), 930 kbp downstream to *oriC* in the permissive and non-permissive orientations. It was expected that the growth rate of these strains would be affected by the strength of TT-lock formation close to *oriC*. All *Ter* sites could be inserted in the permissive orientation without any observable effect on growth rate. The absence of growth defect in the strain carrying an ectopic *TerB* in the permissive orientation suggests that no significant pausing occurs at the permissive face of Tus-*Ter* complexes. *TerB* could not be inserted in the genome in the non-permissive orientation despite the fact that the comparatively strong *TerE* had previously been successfully inverted in the genome and the strain retained some, albeit low, viability (Esnault et al., 2007). It is important to note that *TerE* is one of the weakest of the strong *Ter* sites (Moreau and Schaeffer, 2012a). Two different targetron-based techniques were used to insert *TerB* in the non-permissive orientation but it was systematically mutated or truncated and therefore non functional in the few recombinant colonies obtained (personal communication of P. Enyeart).

This was not too surprising considering that previously, ectopic *TerB* could only be successfully inserted in the non-permissive orientation in a *Δtus* strain (Bidnenko et al., 2002). These results and previously published data suggest that replication forks can only break through the TT-locks of the weaker *TerE* and *TerD*. Indeed, The $t_{1/2}$ of dissociation measured for *TerA-C* were systematically longer than for *TerD*, *E* and *G* (*cf* Figure 16 p90). Duggin and Bell (2009) also showed that *TerA* and *TerB* had the same (and highest) intrinsic pausing efficiency while, *TerC* had a fork arrest efficiency only slightly superior to *TerE* (*cf* Figure 6 p22).

Overall, these data support that (i) *TerB* and probably *TerA* are sufficient to fully arrest replication forks; forks arrested at *TerB* might be the result of rare forks breaking through the slightly less efficient *TerC*; (ii) forks are unlikely to break through the innermost *Ter* sites (*TerA-D*) and reach the outer *Ter* sites; (iii) any *Ter* site weaker than *TerE* such as *TerH* is unable to significantly block a replication fork. It is anticipated that *TerK*, *L* as well as *TerY* and *Z*, oriented to block origin-to-terminus moving forks, cannot stop or pause a replication fork in a significant manner either.

The low probability of *oriC*-initiated forks to reach *TerH* or *TerI* and the fact that no pausing is likely to occur at their permissive face suggests that the pausing observed *in vivo* at these sites (Duggin and Bell, 2009) is due to the blocking of non-*oriC* initiated replication forks. The outer *Ter* sites may be needed to pause recombination-dependent replication forks travelling towards the origin and avoid collision with *oriC*-originated forks. This could explain the growth defect observed by inversion of *TerH* and *TerI* by Esnault et al. (2007) which may have induced recombination events and increased the proportion of fork progressing towards the origin, resulting in genome instability.

A systematic analysis of the effect of ectopically inserted wild-type and mutant *Ter* sites on bacterial survival, growth rates and morphologies could provide valuable information

on the minimal affinity and kinetic properties of a *Ter* site and its TT-lock required to stop a replication fork. It would also be interesting to test the pausing efficiency of various helicases (i.e. PriA and UvrD; Bidnenko et al., 2006) by the outermost *Ter* sites.

7.4.5 A new paradigm for the multiplicity of *Ter* sites?

The function of the outer *Ter* sites is still a question of debate but it is becoming clear that they are not involved in *oriC*-dependent DNA replication pausing. As suggested above, they may be involved in the control of non-*oriC*-initiated forks. However, another role for the Tus-*Ter* complex was envisaged that could also explain why the fork trap is so large. Several membrane associated proteins have been predicted to be functional Tus partners according to the String 9.0 database, such as rstB sensory histidine kinase (directly upstream to the *tus* gene), ftsQ (growth of wall septum), or yrfF (inner membrane protein, Figure 33). Most of these proteins were identified due to their co-occurence with *tus* in prokaryotic genomes (confidence > 0.7). This leads to the possibility that Tus bound to the outer binding sites could act to tether the chromosome to the membrane to help distributing the sister chromosomes in their individual cells by an, as yet unknown, mechanism. Indeed, when considering the symmetry and orientation of the primary *Ter* clusters within the chromosome (Figure 26), *Ter* sites could very well act as anchor points for each newly replicated chromosome to the poles of the bacteria to facilitate their segregation during cell division.
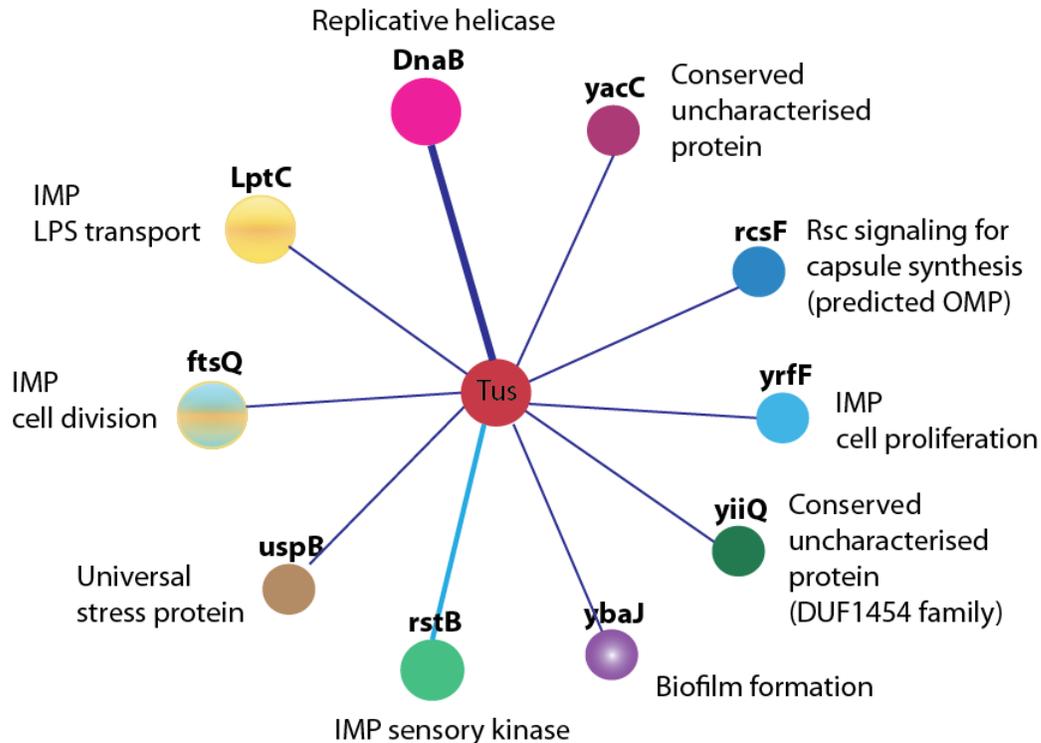
Figure 33: Predicted functional partners of Tus in *E. coli* K-12 MG1655 according to the String 9.0 database[8]. The association map was simplified to represent only the proteins associated with Tus and not the associations between the predicted partners. Circle color is used as a visual aid only. Circle size depends on whether or not there is structural information available for this protein; there is no structural information for proteins shown with a small circle. The biological pathway in which each protein is involved is described next to each circle. The thick dark blue line between Tus and DnaB means that this interaction has been proven by direct experimental evidence (Mulugu et al., 2001). The light blue line connecting Tus and rstB shows that this partner was identified as a neighbor (gene occurs repeatedly in close neighborhood to Tus in prokaryotic genomes (rstB is 76 bp upstream to Tus in *E. coli*) and was shown to be co-expressed with Tus. Thin dark blue lines connecting the other proteins to Tus mean that they are predicted partners based on their co-occurrence in prokaryotic genomes. IMP: inner membrane protein, OMP: outer membrane protein. UniprotKB numbers: yacC (P0AA95), rscF (P69411), yrfF (P45800), yiiQ (P32160), ybaJ (synonym TomB P0AAR0), rstB (P18392), uspB (P0A8S5), ftsQ (P06136), LptC (P0ADV9), DnaB (P0ACB0).

This could also explain why Tus is not an essential protein (Hill, 1992, Hill et al., 1987, Hill et al., 1989, Roecklein et al., 1991). It is very tempting to imagine that the almost perfect positional symmetry of the two primary clusters of *Ter* sites could act in such a simple yet

---

[8] http://string-db.org/

elegant manner to assure that each dividing cell ends with one chromosomal copy but this remains to be demonstrated.

7.5 Conclusion

The global distribution of Tus to its ten primary chromosomal *Ter* sites could be obtained using a modified ChIP-qPCR assay where Tus-*Ter* complexes were captured on a microplate coated with anti-GFP antibodies. It is anticipated that the use of anti-GFP antibodies for ChIP-qPCR will enable the comparative study of virtually any protein-DNA interaction even without the availability of a specific antibody, assuming that the GFP-tag does not impede protein function. Using this technique, *TerF* was shown to be a very weak binder *in vivo* and that no other factor contributes to Tus binding to *Ter* sites. *TerF* is likely to be only rarely occupied *in vivo* in natural cellular concentration of Tus. The distribution of Tus on *Ter* sites correlated well with their individual affinity. It was expected that *TerJ* might be crosslinked and immunoprecipitated in higher yield than *TerI* and *TerH* as a result of longer pausing induced by the stronger TT-lock that it forms, but this was not the case suggesting that the replisome probably never passes the inner *Ter* sites (*TerA-D*). The crosslinking of *Ter* sites was a reflection of Tus binding to *Ter* rather than TT-lock formation and suggested that moderate *Ter* sites are more transiently occupied than the strong *Ter* sites. Finally, it was shown that only the strong TT-lock forming sites are able to arrest replication forks and that the moderate outer *Ter* sites are likely to not be involved significantly in *oriC*-initiated replication termination.

Chapter 8: General Discussion

TT-lock or not TT-lock: that is the question!

Despite its importance, DNA replication termination is the stage of replication that is the least understood both in prokaryotes and eukaryotes. In *E. coli*, 14 termination sites (*TerA-J*, *TerK, L*, *TerY, Z*) are spread throughout the genome; nine *Ter* sites are on the left arm and five *Ter* sites are on the right arm of the chromosome (Duggin and Bell, 2009). Two of the newly identified *Ter* sites (*TerY, Z*) are oriented to block origin-to-terminus moving forks, increasing the complexity of the fork trap. The *Ter* sites are arranged with an intriguing symmetry that has puzzled scientists for decades. The significance of having maintained such a wide fork trap is still poorly understood. It was suggested that the presence of a series of termination sites served as back-up fork barriers in case fork arrest failed at the first *Ter* site of each cluster. In 2009, the pausing efficiency of all *Ter* sites was investigated *in vivo* (Duggin and Bell, 2009). Fork pausing was detected at seven *Ter* sites in wild type cells, namely *TerA-D*, *TerG*, *TerH* and *TerI* (Duggin and Bell, 2009). The remaining *Ter* sites (*TerE*, *TerF*, *TerJ*, *TerK*, *TerL*, *TerY, TerZ*) were qualified as *pseudo-Ter* sites (*pTer*) since they were either non-functional or did not encounter a replisome approaching the non-permissive face. Nevertheless, all *Ter* sites were able to arrest forks in a unidirectional replication plasmid assay, yet with varying efficiencies (Duggin and Bell, 2009).

174

Furthermore, replication forks generally meet at the innermost *Ter* sites directly opposite to *oriC* (i.e. *TerA*, *TerB* or *TerC*) prompting the question of whether or not the remaining *Ter* sites maintained their biological function.

The mechanism by which the Tus-*Ter* complex acts in polar fork arrest is still a matter of debate. One model proposes that fork arrest is mediated by a protein-protein interaction between the DnaB helicase and the non-permissive face of Tus (Bastia et al., 2008, Mulugu et al., 2001). This mechanism is based on a weak Tus-DnaB interaction that has been detected by yeast-two-hybrid analysis and ELISA (Mulugu et al., 2001). One mutant with an affinity for *Ter* similar to wild type Tus, E49Q, was defective in fork arrest but later findings showed that this residue is implicated in the alternative Tus-*Ter*-lock mechanism involved in fork arrest (Mulcair et al., 2006). Indeed, in 2006, Mulcair et al. showed that a specific protein-DNA interaction was formed at the non-permissive face of the Tus-*Ter* complex upon DNA unwinding. Here, the cytosine at position 6 in the *Ter* core is captured by Tus resulting in slower dissociation of Tus from the complex (Mulcair et al., 2006). Interestingly, E49 is involved in this mechanism through a water mediated hydrogen bond with the phosphate of the displaced A(7). Therefore the decrease in fork arrest by the E49Q Tus mutant observed by Mulugu et al. (2001) could have been the result of defective formation of the TT-lock. Mulcair et al. (2006) showed that the E49A Tus mutant had a two-fold decrease in dissociation half-life compared to wild-type Tus by SPR. In order to discriminate between the contribution of the TT-lock and the Tus-DnaB interaction, Bastia et al. (2008) designed an experiment capable of distinguishing DnaB helicase sliding activity from DNA melting activity and observed that Tus could block the sliding helicase without the need for TT-lock formation, but blocking efficiency increased upon lock formation. They also showed that the E49K was defective in stopping the helicase on a bubble substrate with paired C·G(6). They further dismissed the involvement of C(6) flipping and TT-lock

contribution in stopping the helicase by showing that the transversion of C·G(6) to A·T(6) did not affect Tus blocking efficiency. Nevertheless, since the assay was designed to prevent the helicase from unwinding DNA (i.e. C or A(6) can not flip out), it did not really measure the impact of this substitution on helicase arrest. Therefore, even though the TT-lock may not be the only mechanism responsible for the polarity of fork arrest it certainly contributes, justifying the need for further studies to determine if it occurs in all Tus-*Ter* complexes. Indeed, since its discovery in 2006 with the Tus-*TerB* complex, there has been no further investigation of its occurence in the other Tus-*Ter* complexes or *in vivo*. It was only assumed that since the C(6) is conserved in all *Ter* sequences, all Tus-*Ter* complexes could form a TT-lock. The work presented in this thesis provides the first *in vitro* as well as *in vivo* comparative study of the ten primary *Ter* sites (*TerA-J*) in terms of:

- the affinity of Tus for the ten primary *Ter* sites and their ability to form a TT-lock

- the correlation between affinity, TT-lock forming capacity and intrinsic efficiency in arresting a replisome *in vivo*.

- the relative distribution of Tus on the primary *Ter* sites in *vivo* and,

- the ability of the outer *Ter* sites to cause fork arrest

The Ter sites are not equal neither in their affinity for Tus nor in their ability to form a TT-lock

A new GFP-based stability assay (Moreau et al., 2010) was developed to provide a rapid method to determine and compare the affinity of Tus for each primary *Ter* and their *Ter*-lock analogue. GFP-Basta afforded a new quantitative method to correlate the affinity of various DNA ligands with their effect on the aggregation rate (k_{agg}) of Tus-GFP. Importantly it allowed the ranking of *Ter* ligands according to their reported K_D for Tus. Using a combination of GFP-basta and SPR, *Ter* sites could be classified from strong to weak

binders. Six strong affinity *Ter* sites (*TerA-E* and *TerG*) were identified for Tus, with K_D values in the sub-nanomolar range at 150 mM KCl. *TerH*, *TerI* and *TerJ* were classified as moderate binders. *TerF* was the weakest binder and demonstrated a binding affinity up to 880-fold weaker than the strongest binder (*TerA*) and could not bind to Tus in the presence of 250 mM KCl. The binding of Tus to *TerF* was only marginally stronger than to non-specific DNA, leading to the questioning of its role in the fork trap. All strong binders were able to form a strong TT-lock as observed by a decrease in k_d of more than 10-fold by SPR. Interestingly, *TerJ* was the weakest of the moderate binders but was the strongest TT-lock forming site of this group. *TerF* and *TerH* were unable to form significant locks. These data enabled the identification of a base - i.e. T(5) - important for TT-lock formation. T(5) is present in all strong TT-lock forming sites and is substituted to a G in the non TT-lock forming sites. This base seems to be important for the precise docking of C(6) into the cytosine binding pocket on the surface of Tus. Interestingly, the complementary base A(5) has previously been shown to form base-specific contacts with R198 and seems to be crucial for the association of Tus to *Ter*.

These data prompted the development of a higher-throughput method based on differential scanning fluorimetry of GFP-tagged proteins (DSF-GTP; Moreau et al., 2012) to study the effect of ionic strength on the Tus-*Ter* complexes. The specificity of each *Ter* site was determined by screening the effect of increasing potassium chloride concentration on the stability of Tus in complex with *Ter*, *Ter-lock* or non-specific DNA. The data revealed that ionic strength did affect *Ter* and *Ter*-lock induced stability differently. The strong *Ter* sites were more stabilizing in low salt than their respective *Ter-locks* but this trend reversed in high salt. This could be attributed to the formation of specific TT-lock interactions and some non-specific interactions missing in the Tus-*Ter*-lock complexes. These data confirmed the importance of the missing interactions between A(5), G(6) and the R198 residue as well as

the specific contacts occurring in the C(6) binding pocket in complex formation with the *Ter-lock* sequences. Moderate binders showed a variety of salt-dependent profiles indicating that the nature and strength of the interactions in these complexes are more diverse. *TerI* was the least specific of this group and showed a strong dependence on small cooperative electrostatic interactions. *TerH* was more specific than *TerI* and *TerJ* and surprisingly, some TT-lock specific interactions still occurred in the Tus-*Ter-lockH* complex. *TerF* showed some specificity compared to non-specific DNA only below 150 mM KCl.

The affinity of Tus for *Ter* sites in their genomic context was also investigated to verify that there are no additional sequences that account for binding. A new qPCR-based DNA-binding assay was specifically developed for this purpose (Moreau and Schaeffer, 2012b). Using a single binding reaction, enrichment factors were obtained for ~150 bp DNA containing the ten primary *Ter* sites. The data correlated well with the affinity data obtained by SPR and GFP-Basta confirming that no other nucleotide other than the core sequence was involved in Tus binding. This method could detect picomolar concentrations of DNA and was successfully employed as a platform to perform ChIP experiments.

These data, taken together with the latest published work (Bastia et al., 2008, Duggin and Bell, 2009), lead to the proposition of a three-step fork arrest model at TT-lock forming sites involving the non-specific binding of Tus to DNA followed by its correct docking to a strong *Ter* site using a linear ratchet-like binding mode and finally, the formation of a strong TT-lock induced by the unzipping action of DnaB helicase (Moreau and Schaeffer, 2012a). The fact that all functional *Ter* sites with the exception of the outermost *TerH* and *TerF* were able to form a TT-lock and that C(6) and the C(6) binding-pocket have been maintained during evolution, demonstrated the biological importance of this dynamic process.

The TT-lock is formed in vivo but is not solely responsible for fork pausing

The data presented in this thesis revealed that the fork pausing efficiency profile obtained by Duggin and Bell (2009) was best explained by the formation of the TT-lock *in vivo*. However, the formation of the TT-lock, although clearly important, is not the only factor controlling the polarity of fork arrest by Tus as highlighted by the facts that significant pausing was observed at *TerH in vivo* despite its inability to form a TT-lock and that some pausing was also detected at *TerF* when Tus was overexpressed. The pausing at these outer *Ter* sites must therefore be influenced by other factors (see below).

The distribution of Tus-GFP on Ter sites correlates with their affinity but cannot explain fork pausing events

Following the in-depth thermodynamic and kinetic characterization of the Tus-*Ter* complexes, the binding of Tus to the ten primary *Ter* sites was investigated *in vivo* using ChIP analysis to determine if the occupancy on individual *Ter* sites was affected by other factors than their intrinsic affinity (i.e. accessory factors, supercoiled DNA). The new ChIP assay incorporating a French Press step and a microplate-based GFP capture surface yielded highly reproducible data. Immunocapture and de-crosslinking steps were performed in a 96-well plate in three hours. The distribution profile of moderately expressed Tus-GFP on the ten primary *Ter* sites and *oriC* was similar to the binding affinity profile previously obtained for fully double-stranded *Ter* sites, confirming that Tus can bind to all *Ter* sites in an affinity-dependent manner. *TerF* was only marginally enriched compared to the non-specific *oriC* demonstrating that *TerF* is only very transiently bound *in vivo*. Under endogenous Tus expression levels, *TerF* is therefore unlikely to be occupied. This is supported by the fact that the cellular concentration of endogenous Tus has been estimated at 20-100 nM (Natarajan et al., 1993, Roecklein and Kuempel, 1992) which is around the $_cK_D$ of *TerF* (dissociation half

life of about 30 seconds at 150 mM KCl). At this Tus concentration, the strong affinity sites would be occupied first and the remaining *Ter* sites would be partially occupied as a function of free Tus concentration, leaving *TerF* only transiently occupied by Tus if at all.

When the ratio of fork pausing at each *Ter* site was investigated in the chromosomal context (i.e. fork pausing at a given *Ter* site depends on whether or not this site encounters a replication fork at the non-permissive face), *TerE* and *TerJ* were classified as *pTer* sites along with the last four *Ter* sites and *TerF*, based on the characteristic that no Y-shaped DNA intermediate was observed at these sites under endogenous Tus expression (Duggin and Bell, 2009). Using ChIP-qPCR, *TerE* was demonstrated to be a true Tus binding site that should be significantly occupied by Tus under natural conditions. Since *TerE* has been previously shown to be an efficient fork barrier (Duggin and Bell 2009, Esnault et al., 2007, Hidaka et al., 1991), it was concluded that *TerE* does not encounter replication forks moving towards its non-permissive face. This also suggested that *TerH* and *TerI* should not encounter *oriC*-initiated replication forks at their non-permissive face either, although some pausing was detected by Duggin and Bell (2009). The pausing at these outermost *Ter* sites observed by Duggin and Bell must therefore be the result of non-*oriC* initiated replication forks travelling towards the *oriC*. This is supported somewhat by the fact that pausing was increased at these sites upon Tus-overexpression which should result in a tighter inner fork trap and increased inhibition of fork progression towards the outermost *Ter* sites.

Only strong Ter sites can arrest the replisome

Finally, in order to determine the effect of TT-lock formation on replication dynamics *in vivo*, a representative of each category of *Ter* site, i.e. strong (*TerB*), moderate (*TerJ*) and non-TT-lock forming site (*TerH*) was inserted in the right non-structured domain of *E. coli*, 930 kbp downstream to *oriC* in permissive and non-permissive orientations. It was expected that the

formation of the TT-lock would delay DNA replication and cell division proportionally to its strength. *TerB* could not be inserted in the genome in the non-permissive but the remaining *Ter* sites (*TerH* and *J*) could without any noticeable effect on growth rate. These data demonstrate that a strong *Ter* site is able to fully block DNA replication without the need for backup *Ter* sites which strongly suggests that replication forks are unlikely to break through the four innermost *Ter* sites (*TerA-D*).

Table 19: Summary of *Ter* sites affinity and characteristics.

Ter	Affinity		Characteristics
	Ter	TT-lock (fold increase in $t_{1/2}$)[a]	
A	Strongest	Stronger (10.5)	Innermost *Ter* site, second most frequently used termination site
B	Strongest	Strongest (11.5)	Strongest *Ter* and TT-lock forming site, auto-regulation of Tus expression[c], sufficient to fully block replisomes
C	Stronger	Stronger (14)	Innermost *Ter* site, most frequently used termination sites[d]
D	Strong	Strong (12)	Very likely to cause fork arrest
E	Strong	Strong (22)	Never encounters replication forks at the non-permissive face[d] but can arrest replication forks[e] if reached
F	Weak	-	Too weak to be significantly occupied or have a role in replication termination of *oriC*-initiated forks
G	Strong	Strong (14.5)	Unlikely to encounter replication forks at the non-permissive face but very likely to cause fork arrest if reached
H	Moderate	-	Cannot arrest an *oriC*-initiated replication fork, likely to be involved in the control of recombination forks
I	Moderate	Weak (1.1)[b]	
J	Moderate	Yes (3.7)[b]	
K, L and Y, Z	Very weak	-	Predicted to be very weak, unable to generate a TT-lock or stop *oriC*-initiated replication forks

[a] TT-lock induced increase in dissociation half life $t_{1/2}$ relative to $t_{1/2}$ Tus-*Ter* determined by SPR at 250 mM KCl and [b] at 150 mM KCl; [c] (Natarajan et al., 1991); [d] (Duggin and Bell, 2009); [e] (Esnault et al., 2007).

In addition, the absence of a clear phenotype when an ectopic *TerB* was inserted in the permissive orientation suggested that no pausing occurs at the permissive face of Tus-*Ter*

complexes. The data also demonstrated that moderate or weak *Ter* sites, such as the newly identified *Ter* sites, cannot significantly arrest *oriC*-initiated replication forks and are therefore likely to be involved in the control of recombination-initiated forks.

The major outcomes and conclusions about the role of the primary *Ter* sites reached during this thesis are presented in Table 19. While the function of the inner *Ter* sites is relatively clear, the significance of the presence of the remaining *Ter* sites is still to be determined. Most of the outer *Ter* sites have non-canonical base substitutions in their core sequences that are affecting their binding to Tus. Is it due to random mutagenesis of a chromosome in constant evolution or has it been fine-tuned to serve a desirable function? In other words, did they devolve from their original function as stronger *Ter* sites appeared in the fork trap during evolution, and should they be classified as *pseudo-Ter* sites? Can they have another function than coordinating DNA replication termination? A new paradigm for an alternative use of Tus-*Ter* complexes and further work are suggested in the following sections to deepen our understanding of the evolutionary advantage of maintaining such a wide fork trap.

Can pausing at Ter sites vary depending on DNA supercoiling and the composition of the replisome?

The discrepancies between the affinity, TT-lock forming efficiency and the pausing observed at the outer *Ter* sites (i.e. pausing at *TerH* and *TerI* but not at *TerE*) suggested that the outer *Ter* sites may be involved in the pausing of non-*oriC* initiated forks, i.e. recombination forks or forks initiated at alternative origins. If that was the case, they may have a different pausing efficiency depending on the nature of the replisome approaching these *Ter* sites or the distance a fork has travelled (Bidnenko et al., 2002). The base substitutions in the outer *Ter* sites may cause Tus to bind in a different conformation

affecting the pausing at Tus-*Ter* complexes differently depending on the sequence (i.e. Tus conformation), orientation of *Ter* and nature of the approaching helicase. Furthermore, in *E. coli*, a topoisomerase I mutant resulting in increased negative supercoiling, abolished replication arrest at Tus-*Ter* (Valjavec-Gratian et al., 2005). It was suggested that topoisomerase-mediated supercoiling relaxation modulates the activity of Tus-*Ter* complex. Indeed, supercoiling could affect the permissive and non-permissive face of Tus-*Ter* complexes differently, generating tighter or weaker fork blocks depending on the state of the DNA and speed of the helicase. More studies are needed to determine if the variable sequences at the permissive side of *Ter* can differentially influence fork pausing *in vivo* as well as *in vitro*. It would also be very informative to determine the crystal structure of Tus in complex with the outer *Ter* sites to determine if Tus takes a different conformation on these sites.

Can Ter sites be involved in chromosome segregation during cell division?

A new paradigm on the role and significance of having maintained such a wide fork trap emerged from the following observations: (a) Tus is not essential for cell viability (Sharma and Hill, 1995, Hill, 1992, Hill et al., 1987, Hill et al., 1989, Roecklein et al., 1991); (b) the termination region is maintained at a given position with little movement until cell division (Bates and Kleckner, 2005); and finally, (c) several sensory and cell division-related membrane proteins have been predicted to be partners of Tus, leading to the question of whether Tus could facilitate cell division through a chromosome management mechanism during replication. When combined, these observations raise the possibility that the *Ter* sites could act as membrane anchoring points to ensure that each new cell ends up with one copy of the chromosome.
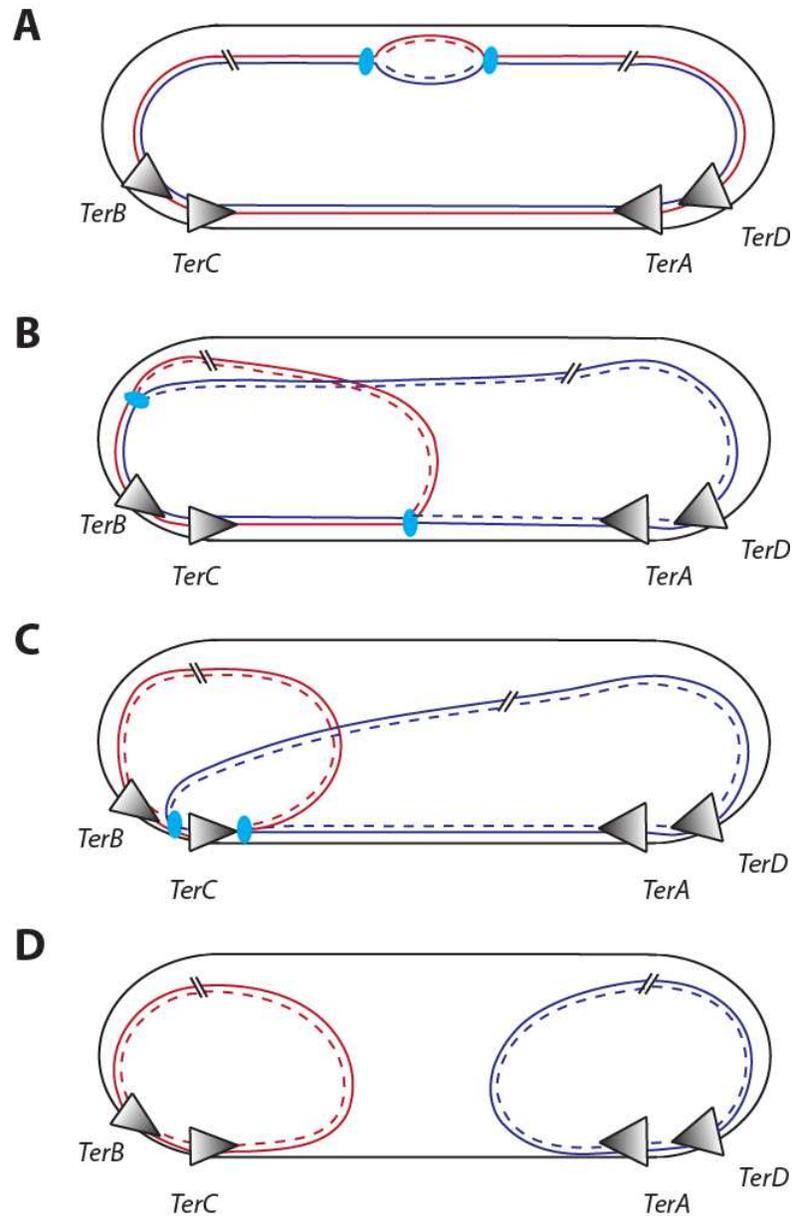
Figure 34: Hypothetical model on the role of Tus in chromosome segregation. Tus is represented by a grey triangle with the tip representing the non-permissive face and is tethered to the membrane by a yet unknown mechanism. The replisome is represented by a light blue circle. Parental and daughter chromosomes are represented by full and dashed lines respectively. The sister chromosomes are in red and blue respectively. (A) Following initiation the two replisomes proceed in opposite directions. (B) Following passage of the right replisome through *TerD* and *TerA*, Tus reassociated to these sites and tethers the right chromosome to the right membrane compartment. (C) While the right replisome is stalled at *TerC*, the left replisome passed *TerB* where Tus reassociated and tethered the left chromosome to the left membrane compartment. The left replisome dissociates Tus from *TerC* and the two replisomes merge. (D) Each sister chromosome is tethered to the opposite cell compartment prior to cell division.

Indeed, when taking into account the symmetry of each cluster of primary *Ter* sites and their opposite polarity within the chromosome, it seems possible that by tethering the chromosome to the membrane, the Tus-*Ter* complexes could help in the distribution of the sister chromosomes into the daughter cells during cell division. A simplified diagram is presented in Figure 34 to illustrate this theory. In this paradigm, Tus is associated with the membrane through an unidentified protein partner and can re-associate with *Ter* after it has been dissociated by the passage of a replication fork. The model requires that at least one Tus-*Ter* complex on each side of the chromosome must be tethered to their respective membrane compartment at any time (Figure 34). In this context, the orientation of Tus on *Ter* sites in each cluster would allow each leading strand to move towards the outer poles as observed by White et al. (2008). Furthermore, if the sister chromosomes are attached to the cell membrane and cell wall growth takes place in between the opposite attachment sites, DNA segregation could be achieved passively as a by-product of cell elongation as suggested early on by Jacob et al. (Ryter et al., 1968, Toro and Shapiro, 2010). Chromosome-anchoring proteins have been found in *Caulobacter* (ParB-PopZ) and *B. subtilis* (RacA-DivIVA) but not in *E. coli* (reviewed in Toro and Shapiro, 2010) although it is clear that *E. coli* maintains the origin region around midcell and left and right replichores in separate cell halves (Bates and Kleckner, 2005, Toro and Shapiro, 2010, White et al., 2008). Therefore segregation is nonrandom, and the leading strand replicated arms are shuttled to the outer edges of the cell (Toro and Shapiro, 2010, White et al., 2008). Early biochemical experiments in *E. coli* showed that multiple chromosomal sites, including not only the region around *oriC* but also the terminus and the site of ongoing replication, are found to associate with the membrane (Hendrickson et al., 1982, Leibowitz and Schaechter, 1975, Ryter et al., 1968, Toro and Shapiro, 2010). More recent experiments on chromosome segregation in *E.coli* showed that the origin region undergoes a short period of "sister chromatid cohesion," after which the

entire region is segregated as a single unit (Bates and Kleckner, 2005, Espeli et al., 2008). Other regions of the chromosome segregate progressively (Nielsen et al., 2006). Tus could be the long sought-after anchoring landmark. The origins of sister chromosomes would be segregated only once the outer *Ter* sites move apart upon cell wall growth and the rest of the chromosome would segregate progressively. This model would also explain why the fork trap is so large. How the chromosome is organized and segregated is still unclear (Dame et al., 2011), although there is evidence that segregation is linked to replication rather than cell division (Nielsen et al., 2007). Therefore it would be very informative to determine if a fraction of Tus is associated with the membrane and its localization within the cell. The Tus-GFP could undoubtedly be an invaluable tool to perform these interesting experiments.

Conclusion

This thesis has shed light on the structure of the fork trap in terms of biochemical characterization of Tus in complex with the ten primary *Ter* sites and their ability to form a TT-lock. The variation in affinity amongst *Ter* sites correlated with their intrinsic efficiency in pausing a replisome and suggested that the TT-lock does occur *in vivo*. A three-step fork arrest model was proposed. Although the TT-lock is clearly important, it is not the only factor responsible for replication fork arrest. While the role and use of the innermost *Ter* sites was consistent between *in vivo* and *in vitro* data, the role of the moderate and weak *Ter* sites in replication termination is still debatable as the nature of the pausing observed at these sites by Duggin and Bell (2009) remains uncertain. It was demonstrated that only strong TT-lock forming sites can efficiently arrest replication forks suggesting that replication forks can not break through the innermost *Ter* sites and that the outermost *Ter* sites (including the newly identified *Ter* sites) have no role in the termination of *oriC*-initiated DNA replication. These

findings negate the theory that a series of *Ter* sites act as "back-up" fork barriers in case of fork blocking failure at the first *Ter* site of the cluster. Finally the new assays, i.e. GFP-Basta, qPCR-based DNA-binding assay and DSF-GTP, developed to characterize the fork trap will undoubtedly have uses for the study of other protein-DNA complexes, and will certainly increase the pace of current and future genomics and proteomics programs.

References

ADACHI, S., KOHIYAMA, M., ONOGI, T. & HIRAGA, S. 2005. Localization of replication forks in wild-type and mukB mutant cells of *Escherichia coli*. *Mol Genet Genomics,* 274**,** 264-71.

ALBERTS, B. M., BARRY, J., BEDINGER, P., FORMOSA, T., JONGENEEL, C. V. & KREUZER, K. N. 1983. Studies on DNA replication in the bacteriophage T4 *in vitro* system. *Cold Spring Harb Symp Quant Biol,* 47 Pt 2**,** 655-68.

ANDERSON, S. G., WILLIAMS, C. R., O'DONNELL, M. & BLOOM, L. B. 2007. A function for the psi subunit in loading the *Escherichia coli* DNA polymerase sliding clamp. *J Biol Chem,* 282**,** 7035-45.

ANFINSEN, C. B. 1973. Principles that govern the folding of protein chains. *Science,* 181**,** 223-30.

ARAKAWA, K., SAITO, R. & TOMITA, M. 2007. Noise-reduction filtering for accurate detection of replication termini in bacterial genomes. *FEBS Lett,* 581**,** 253-8.

ASKIN, S. P., MORIN, I. & SCHAEFFER, P. M. 2011. Development of a protease activity assay using heat-sensitive Tus-GFP fusion protein substrates. *Anal Biochem,* 415**,** 126-33.

AZVOLINSKY, A., GIRESI, P. G., LIEB, J. D. & ZAKIAN, V. A. 2009. Highly transcribed RNA polymerase II genes are impediments to replication fork progression in *Saccharomyces cerevisiae. Mol Cell,* 34**,** 722-34.

BAICHOO, N. & HELMANN, J. D. 2002. Recognition of DNA by Fur: a reinterpretation of the Fur box consensus sequence. *J Bacteriol,* 184**,** 5826-32.

BASTIA, D. & SINGH, S. K. 2011. "Chromosome kissing" and modulation of replication termination. *Bioarchitecture,* 1**,** 24-28.

BASTIA, D., ZZAMAN, S., KRINGS, G., SAXENA, M., PENG, X. & GREENBERG, M. M. 2008. Replication termination mechanism as revealed by Tus-mediated polar arrest of a sliding helicase. *Proc Natl Acad Sci U S A,* 105**,** 12831-6.

BATES, D., EPSTEIN, J., BOYE, E., FAHRNER, K., BERG, H. & KLECKNER, N. 2005. The *Escherichia coli* baby cell column: a novel cell synchronization method provides new insight into the bacterial cell cycle. *Mol Microbiol,* 57**,** 380-91.

BATES, D. & KLECKNER, N. 2005. Chromosome and replisome dynamics in *E. coli*: loss of sister cohesion triggers global chromosome movement and mediates chromosome segregation. *Cell,* 121**,** 899-911.

BEDROSIAN, C. L. & BASTIA, D. 1991. *Escherichia coli* replication terminator protein impedes simian virus 40 (SV40) DNA replication fork movement and SV40 large tumor antigen helicase activity *in vitro* at a prokaryotic terminus sequence. *Proc Natl Acad Sci U S A,* 88**,** 2618-22.

BENKOVIC, S. J., VALENTINE, A. M. & SALINAS, F. 2001. Replisome-mediated DNA replication. *Annu Rev Biochem,* 70**,** 181-208.

BERDIS, A. J. 2008. DNA polymerases as therapeutic targets. *Biochemistry,* 47**,** 8253-60.

BIDNENKO, V., EHRLICH, S. D. & MICHEL, B. 2002. Replication fork collapse at replication terminator sequences. *EMBO J,* 21**,** 3898-907.

BIDNENKO, V., LESTINI, R. & MICHEL, B. 2006. The *Escherichia coli* UvrD helicase is essential for Tus removal during recombination-dependent replication restart from *Ter* sites. *Mol Microbiol,* 62**,** 382-96.

BIRD, R. E., LOUARN, J., MARTUSCELLI, J. & CARO, L. 1972. Origin and sequence of chromosome replication in *Escherichia coli*. *J Mol Biol,* 70**,** 549-66.

BLAKELY, G. W. & SHERRATT, D. J. 1994. Interactions of the site-specific recombinases XerC and XerD with the recombination site *dif*. *Nucleic Acids Res,* 22**,** 5613-20.

BLANCO, C. & MATA-GILSINGER, M. 1986. A DNA sequence containing the control sites for the *uxaB* gene of *Escherichia coli*. *J Gen Microbiol,* 132**,** 697-705.

BLINKOWA, A. L. & WALKER, J. R. 1990. Programmed ribosomal frameshifting generates the *Escherichia coli* DNA polymerase III gamma subunit from within the tau subunit reading frame. *Nucleic Acids Res,* 18**,** 1725-9.

BLOOM, L. B. 2006. Dynamics of loading the *Escherichia coli* DNA polymerase processivity clamp. *Crit Rev Biochem Mol Biol,* 41**,** 179-208.

BOLOGNA, F. P., ANDREO, C. S. & DRINCOVICH, M. F. 2007. *Escherichia coli* malic enzymes: two isoforms with substantial differences in kinetic properties, metabolic regulation, and structure. *J Bacteriol,* 189**,** 5937-46.

BRAMHILL, D. & KORNBERG, A. 1988. Duplex opening by DnaA protein at novel sequences in initiation of replication at the origin of the *E. coli* chromosome. *Cell,* 52**,** 743-55.

BRANDTS, J. F. & LIN, L. N. 1990. Study of strong to ultratight protein interactions using differential scanning calorimetry. *Biochemistry,* 29**,** 6927-40.

BREIER, A. M. & GROSSMAN, A. D. 2009. Dynamic association of the replication initiator and transcription factor DnaA with the *Bacillus subtilis* chromosome during replication stress. *J Bacteriol,* 191**,** 486-93.

BREIER, A. M., WEIER, H. U. & COZZARELLI, N. R. 2005. Independence of replisomes in *Escherichia coli* chromosomal replication. *Proc Natl Acad Sci U S A,* 102**,** 3942-7.

BROWN, T. A. & CLAYTON, D. A. 2002. Release of replication termination controls mitochondrial DNA copy number after depletion with 2',3'-dideoxycytidine. *Nucleic Acids Res,* 30**,** 2004-10.

BURNOUF, D. Y., OLIERIC, V., WAGNER, J., FUJII, S., REINBOLT, J., FUCHS, R. P. & DUMAS, P. 2004. Structural and biochemical analysis of sliding clamp/ligand interactions suggest a competition between replicative and translesion DNA polymerases. *J Mol Biol,* 335**,** 1187-97.

BUSSIERE, D. E., BASTIA, D. & WHITE, S. W. 1995. Crystal structure of the replication terminator protein from *B. subtilis* at 2.6 A. *Cell,* 80**,** 651-60.

CADMAN, C. J. & MCGLYNN, P. 2004. PriA helicase and SSB interact physically and functionally. *Nucleic Acids Res,* 32**,** 6378-87.

CARBALLES, F., BERTRAND, C., BOUCHE, J. P. & CAM, K. 1999. Regulation of *Escherichia coli* cell division genes ftsA and ftsZ by the two-component system rcsC-rcsB. *Mol Microbiol,* 34**,** 442-50.

CASSLER, M. R., GRIMWADE, J. E. & LEONARD, A. C. 1995. Cell cycle-specific changes in nucleoprotein complexes at a chromosomal replication origin. *EMBO J,* 14**,** 5833-41.

CAWLEY, S., BEKIRANOV, S., NG, H. H., KAPRANOV, P., SEKINGER, E. A., KAMPA, D., PICCOLBONI, A., SEMENTCHENKO, V., CHENG, J., WILLIAMS, A. J., WHEELER, R., WONG, B., DRENKOW, J., YAMANAKA, M., PATEL, S., BRUBAKER, S., TAMMANA, H., HELT, G., STRUHL, K. & GINGERAS, T. R. 2004. Unbiased mapping of transcription factor binding sites along human chromosomes 21 and 22 points to widespread regulation of noncoding RNAs. *Cell,* 116**,** 499-509.

CHATTERJEE, D. K., SITARAMAN, K., BAPTISTA, C., HARTLEY, J., HILL, T. M. & MUNROE, D. J. 2008. Protein microarray on-demand: a novel protein microarray system. *PLoS One,* 3**,** e3265.

CHI, E. Y., KRISHNAN, S., RANDOLPH, T. W. & CARPENTER, J. F. 2003. Physical stability of proteins in aqueous solution: mechanism and driving forces in nonnative protein aggregation. *Pharm Res,* 20**,** 1325-36.

CLARK, P. L. 2004. Protein folding in the cell: reshaping the folding funnel. *Trends Biochem Sci,* 29**,** 527-34.

CLAYTON, D. A. 1991. Replication and transcription of vertebrate mitochondrial DNA. *Annu Rev Cell Biol,* 7**,** 453-78.

CORN, J. E., PELTON, J. G. & BERGER, J. M. 2008. Identification of a DNA primase template tracking site redefines the geometry of primer synthesis. *Nat Struct Mol Biol,* 15**,** 163-9.

COSKUN-ARI, F. F. & HILL, T. M. 1997. Sequence-specific interactions in the Tus-*Ter* complex and the effect of base pair substitutions on arrest of DNA replication in *Escherichia coli. J Biol Chem,* 272**,** 26448-56.

CRAMERI, A., WHITEHORN, E. A., TATE, E. & STEMMER, W. P. 1996. Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nat Biotechnol,* 14**,** 315-9.

CROOKE, E., THRESHER, R., HWANG, D. S., GRIFFITH, J. & KORNBERG, A. 1993. Replicatively active complexes of DnaA protein and the *Escherichia coli* chromosomal origin observed in the electron microscope. *J Mol Biol,* 233**,** 16-24.

DAHDAH, D. B., MORIN, I., MOREAU, M. J., DIXON, N. E. & SCHAEFFER, P. M. 2009. Site-specific covalent attachment of DNA to proteins using a photoactivatable Tus-*Ter* complex. *Chem Commun (Camb)***,** 3050-2.

DALEY, D. O., RAPP, M., GRANSETH, E., MELEN, K., DREW, D. & VON HEIJNE, G. 2005. Global topology analysis of the *Escherichia coli* inner membrane proteome. *Science,* 308**,** 1321-3.

DALGAARD, J. Z. & KLAR, A. J. 2001. A DNA replication-arrest site RTS1 regulates imprinting by determining the direction of replication at mat1 in *S. pombe. Genes Dev,* 15**,** 2060-8.

DAME, R. T., KALMYKOWA, O. J. & GRAINGER, D. C. 2011. Chromosomal macrodomains and associated proteins: implications for DNA organization and replication in gram negative bacteria. *PLoS Genet,* 7**,** e1002123.

DAVEY, M. J., JERUZALMI, D., KURIYAN, J. & O'DONNELL, M. 2002. Motors and switches: AAA+ machines within the replisome. *Nat Rev Mol Cell Biol,* 3**,** 826-35.

DE MASSY, B., BEJAR, S., LOUARN, J., LOUARN, J. M. & BOUCHE, J. P. 1987. Inhibition of replication forks exiting the terminus region of the *Escherichia coli* chromosome occurs at two loci separated by 5 min. *Proc Natl Acad Sci U S A,* 84**,** 1759-63.

DELAGOUTTE, E. & VON HIPPEL, P. H. 2003. Helicase mechanisms and the coupling of helicases within macromolecular machines. Part II: Integration of helicases into cellular processes. *Q Rev Biophys,* 36**,** 1-69.

DELMAS, S. & MATIC, I. 2006. Interplay between replication and recombination in *Escherichia coli*: impact of the alternative DNA polymerases. *Proc Natl Acad Sci U S A,* 103**,** 4564-9.

DEN BLAAUWEN, T., AARSMAN, M. E., WHEELER, L. J. & NANNINGA, N. 2006. Pre-replication assembly of *E. coli* replisome components. *Mol Microbiol,* 62**,** 695-708.

DESHPANDE, A. M. & NEWLON, C. S. 1996. DNA replication fork pause sites dependent on transcription. *Science,* 272**,** 1030-3.

DOBSON, C. M., ŠALI, A. & KARPLUS, M. 1998. Protein Folding: A Perspective from Theory and Experiment. *Angewandte Chemie International Edition,* 37**,** 868-893.

DUDERSTADT, K. E., CHUANG, K. & BERGER, J. M. 2011. DNA stretching by bacterial initiators promotes replication origin opening. *Nature,* 478**,** 209-13.

DUGGIN, I. G. 2006. DNA replication fork arrest by the *Bacillus subtilis* RTP-DNA complex involves a mechanism that is independent of the affinity of RTP-DNA binding. *J Mol Biol,* 361**,** 1-6.

DUGGIN, I. G., ANDERSEN, P. A., SMITH, M. T., WILCE, J. A., KING, G. F. & WAKE, R. G. 1999. Site-directed mutants of RTP of *Bacillus subtilis* and the mechanism of replication fork arrest. *J Mol Biol,* 286**,** 1325-35.

DUGGIN, I. G. & BELL, S. D. 2009. Termination structures in the *Escherichia coli* chromosome replication fork trap. *J Mol Biol,* 387**,** 532-9.

DUGGIN, I. G., DUBARRY, N. & BELL, S. D. 2011. Replication termination and chromosome dimer resolution in the archaeon *Sulfolobus solfataricus*. *EMBO J,* 30**,** 145-53.

DUGGIN, I. G., MATTHEWS, J. M., DIXON, N. E., WAKE, R. G. & MACKAY, J. P. 2005. A complex mechanism determines polarity of DNA replication fork arrest by the replication terminator complex of *Bacillus subtilis*. *J Biol Chem,* 280**,** 13105-13.

DUGGIN, I. G., WAKE, R. G., BELL, S. D. & HILL, T. M. 2008. The replication fork trap and termination of chromosome replication. *Mol Microbiol,* 70**,** 1323-33.

DURFEE, T., NELSON, R., BALDWIN, S., PLUNKETT, G., 3RD, BURLAND, V., MAU, B., PETROSINO, J. F., QIN, X., MUZNY, D. M., AYELE, M., GIBBS, R. A., CSORGO, B., POSFAI, G., WEINSTOCK, G. M. & BLATTNER, F. R. 2008. The complete genome sequence of *Escherichia coli* DH10B: insights into the biology of a laboratory workhorse. *J Bacteriol,* 190**,** 2597-606.

EL HOURY MIGNAN, S., WITTE, G., NAUE, N. & CURTH, U. 2011. Characterization of the chipsi subcomplex of *Pseudomonas aeruginosa* DNA polymerase III. *BMC Mol Biol,* 12**,** 43.

ENGLER, L. E., WELCH, K. K. & JEN-JACOBSON, L. 1997. Specific binding by EcoRV endonuclease to its DNA recognition site GATATC. *J Mol Biol,* 269**,** 82-101.

EPPS, D. E. & TAYLOR, B. M. 2001. A competitive fluorescence assay to measure the reactivity of compounds. *Anal Biochem,* 295**,** 101-6.

ERICSSON, U. B., HALLBERG, B. M., DETITTA, G. T., DEKKER, N. & NORDLUND, P. 2006. Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Anal Biochem,* 357**,** 289-98.

ERZBERGER, J. P., MOTT, M. L. & BERGER, J. M. 2006. Structural basis for ATP-dependent DnaA assembly and replication-origin remodeling. *Nat Struct Mol Biol,* 13**,** 676-83.

ERZBERGER, J. P., PIRRUCCELLO, M. M. & BERGER, J. M. 2002. The structure of bacterial DnaA: implications for general mechanisms underlying DNA replication initiation. *EMBO J,* 21**,** 4763-73.

ESNAULT, E., VALENS, M., ESPELI, O. & BOCCARD, F. 2007. Chromosome structuring limits genome plasticity in *Escherichia coli*. *PLoS Genet,* 3**,** e226.

ESPELI, O., MERCIER, R. & BOCCARD, F. 2008. DNA dynamics vary according to macrodomain topography in the *E. coli* chromosome. *Mol Microbiol,* 68**,** 1418-27.

EYDMANN, T., SOMMARIVA, E., INAGAWA, T., MIAN, S., KLAR, A. J. & DALGAARD, J. Z. 2008. Rtf1-mediated eukaryotic site-specific replication termination. *Genetics,* 180**,** 27-39.

FACHINETTI, D., BERMEJO, R., COCITO, A., MINARDI, S., KATOU, Y., KANOH, Y., SHIRAHIGE, K., AZVOLINSKY, A., ZAKIAN, V. A. & FOIANI, M. 2010. Replication termination at eukaryotic chromosomes is mediated by Top2 and occurs at genomic loci containing pausing elements. *Mol Cell,* 39**,** 595-605.

FALKENBERG, M., LARSSON, N. G. & GUSTAFSSON, C. M. 2007. DNA replication and transcription in mammalian mitochondria. *Annu Rev Biochem,* 76**,** 679-99.

FANG, L., DAVEY, M. J. & O'DONNELL, M. 1999. Replisome assembly at *oriC*, the replication origin of *E. coli*, reveals an explanation for initiation sites outside an origin. *Mol Cell,* 4**,** 541-53.

FAVICCHIO, R., DRAGAN, A. I., KNEALE, G. G. & READ, C. M. 2009. Fluorescence spectroscopy and anisotropy in the analysis of DNA-protein interactions. *Methods Mol Biol,* 543**,** 589-611.

FELCZAK, M. M., SIMMONS, L. A. & KAGUNI, J. M. 2005. An essential tryptophan of *Escherichia coli* DnaA protein functions in oligomerization at the *E. coli* replication origin. *J Biol Chem,* 280**,** 24627-33.

FERULLO, D. J., COOPER, D. L., MOORE, H. R. & LOVETT, S. T. 2009. Cell cycle synchronization of *Escherichia coli* using the stringent response, with fluorescence labeling assays for DNA content and replication. *Methods,* 48**,** 8-13.

FLANAGIN, S., NELSON, J. D., CASTNER, D. G., DENISENKO, O. & BOMSZTYK, K. 2008. Microplate-based chromatin immunoprecipitation method, Matrix ChIP: a platform to study signaling of complex genomic events. *Nucleic Acids Res,* 36**,** e17.

FLOWER, A. M. & MCHENRY, C. S. 1990. The gamma subunit of DNA polymerase III holoenzyme of *Escherichia coli* is produced by ribosomal frameshifting. *Proc Natl Acad Sci U S A,* 87**,** 3713-7.

FORWOOD, J. K. & JANS, D. A. 2006. Quantitative analysis of DNA-protein interactions using double-labeled native gel electrophoresis and fluorescence-based imaging. *Electrophoresis,* 27**,** 3166-70.

FOSTER, B. A., COFFEY, H. A., MORIN, M. J. & RASTINEJAD, F. 1999. Pharmacological rescue of mutant p53 conformation and function. *Science,* 286**,** 2507-10.

FRANCOIS, V., LOUARN, J. & LOUARN, J. M. 1989. The terminus of the *Escherichia coli* chromosome is flanked by several polar replication pause sites. *Mol Microbiol,* 3**,** 995-1002.

FRANKE, S., GRASS, G., RENSING, C. & NIES, D. H. 2003. Molecular analysis of the copper-transporting efflux system CusCFBA of *Escherichia coli*. *J Bacteriol,* 185**,** 3804-12.

FRICK, D. N. & RICHARDSON, C. C. 2001. DNA primases. *Annu Rev Biochem,* 70**,** 39-80.

FRIEDBERG, E. C., WAGNER, R. & RADMAN, M. 2002. Specialized DNA polymerases, cellular survival, and the genesis of mutations. *Science,* 296**,** 1627-30.

FUJIMITSU, K., SENRIUCHI, T. & KATAYAMA, T. 2009. Specific genomic sequences of *E. coli* promote replicational initiation by directly reactivating ADP-DnaA. *Genes Dev,* 23**,** 1221-33.

FUKUSHIMA, S., ITAYA, M., KATO, H., OGASAWARA, N. & YOSHIKAWA, H. 2007. Reassessment of the *in vivo* functions of DNA polymerase I and RNase H in bacterial cell growth. *J Bacteriol,* 189**,** 8575-83.

FUNNELL, B. E., BAKER, T. A. & KORNBERG, A. 1987. *In vitro* assembly of a prepriming complex at the origin of the *Escherichia coli* chromosome. *J Biol Chem,* 262**,** 10327-34.

FURUKOHRI, A., GOODMAN, M. F. & MAKI, H. 2008. A dynamic polymerase exchange with *Escherichia coli* DNA polymerase IV replacing DNA polymerase III on the sliding clamp. *J Biol Chem,* 283**,** 11260-9.

GAO, D. & MCHENRY, C. S. 2001. Tau binds and organizes *Escherichia coli* replication proteins through distinct domains. Domain IV, located within the unique C terminus of tau, binds the replication fork, helicase, DnaB. *J Biol Chem,* 276**,** 4441-6.

GARNER, M. M. & RAU, D. C. 1995. Water release associated with specific binding of gal repressor. *EMBO J,* 14**,** 1257-63.

GEERTZ, M. & MAERKL, S. J. 2010. Experimental strategies for studying transcription factor-DNA binding specificities. *Brief Funct Genomics,* 9**,** 362-73.

GENSCHEL, J., CURTH, U. & URBANKE, C. 2000. Interaction of *E. coli* single-stranded DNA binding protein (SSB) with exonuclease I. The carboxy-terminus of SSB is the recognition site for the nuclease. *Biol Chem,* 381**,** 183-92.

GEORGESCU, R. E., KURTH, I. & O'DONNELL, M. E. 2012. Single-molecule studies reveal the function of a third polymerase in the replisome. *Nat Struct Mol Biol,* 19**,** 113-6.

GEORGESCU, R. E., KURTH, I., YAO, N. Y., STEWART, J., YURIEVA, O. & O'DONNELL, M. 2009. Mechanism of polymerase collision release from sliding clamps on the lagging strand. *EMBO J,* 28**,** 2981-91.

GERMINO, J. & BASTIA, D. 1981. Termination of DNA replication *in vitro* at a sequence-specific replication terminus. *Cell,* 23**,** 681-7.

GOODMAN, M. F. 2000. Coping with replication 'train wrecks' in *Escherichia coli* using Pol V, Pol II and RecA proteins. *Trends Biochem Sci,* 25**,** 189-95.

GOTTESMAN, S. 2005. Micros for microbes: non-coding regulatory RNAs in bacteria. *Trends Genet,* 21**,** 399-404.

GOTTLIEB, P. A., WU, S., ZHANG, X., TECKLENBURG, M., KUEMPEL, P. & HILL, T. M. 1992. Equilibrium, kinetic, and footprinting studies of the Tus-*Ter* protein-DNA interaction. *J Biol Chem,* 267**,** 7434-43.

GRAINGE, I., LESTERLIN, C. & SHERRATT, D. J. 2011. Activation of XerCD-dif recombination by the FtsK DNA translocase. *Nucleic Acids Res,* 39**,** 5140-8.

GRAINGER, D. C., AIBA, H., HURD, D., BROWNING, D. F. & BUSBY, S. J. 2007. Transcription factor distribution in *Escherichia coli*: studies with FNR protein. *Nucleic Acids Res,* 35**,** 269-78.

GRAINGER, D. C., HURD, D., HARRISON, M., HOLDSTOCK, J. & BUSBY, S. J. 2005. Studies of the distribution of *Escherichia coli* cAMP-receptor protein and RNA polymerase along the *E. coli* chromosome. *Proc Natl Acad Sci U S A,* 102**,** 17693-8.

GRAINGER, D. C., OVERTON, T. W., REPPAS, N., WADE, J. T., TAMAI, E., HOBMAN, J. L., CONSTANTINIDOU, C., STRUHL, K., CHURCH, G. & BUSBY, S. J. 2004. Genomic studies with *Escherichia coli* MelR protein: applications of chromatin immunoprecipitation and microarrays. *J Bacteriol,* 186**,** 6938-43.

GRAY, M. W. 2012. Mitochondrial evolution. *Cold Spring Harb Perspect Biol,* 4**,** a011403.

GRIFFITHS, A. A., ANDERSEN, P. A. & WAKE, R. G. 1998. Replication terminator protein-based replication fork-arrest systems in various *Bacillus* species. *J Bacteriol,* 180**,** 3360-7.

GRIGORIAN, A. V., LUSTIG, R. B., GUZMAN, E. C., MAHAFFY, J. M. & ZYSKIND, J. W. 2003. *Escherichia coli* cells with increased levels of DnaA and deficient in recombinational repair have decreased viability. *J Bacteriol,* 185**,** 630-44.

GRIGORIEV, A. 1998. Analyzing genomes with cumulative skew diagrams. *Nucleic Acids Res,* 26**,** 2286-90.

GUAJARDO, R. & SOUSA, R. 1999. Characterization of the effects of *Escherichia coli* replication terminator protein (Tus) on transcription reveals dynamic nature of the Tus block to transcription complex progression. *Nucleic Acids Res,* 27**,** 2814-24.

GUILLON, J. M., MECHULAM, Y., BLANQUET, S. & FAYAT, G. 1993. Importance of formylability and anticodon stem sequence to give a tRNA(Met) an initiator identity in *Escherichia coli*. *J Bacteriol,* 175**,** 4507-14.

HANCOCK, V., DAHL, M. & KLEMM, P. 2010. Abolition of biofilm formation in urinary tract *Escherichia coli* and Klebsiella isolates by metal interference through competition for fur. *Appl Environ Microbiol,* 76**,** 3836-41.

HARTIG, E., SCHIEK, U., VOLLACK, K. U. & ZUMFT, W. G. 1999. Nitrate and nitrite control of respiratory nitrate reduction in denitrifying *Pseudomonas stutzeri* by a two-component regulatory system homologous to NarXL of *Escherichia coli*. *J Bacteriol,* 181**,** 3658-65.

HASTINGS, A. F., OTTING, G., FOLMER, R. H., DUGGIN, I. G., WAKE, R. G., WILCE, M. C. & WILCE, J. A. 2005. Interaction of the replication terminator protein of *Bacillus subtilis* with DNA probed by NMR spectroscopy. *Biochem Biophys Res Commun,* 335**,** 361-6.

HENDERSON, T. A., NILLES, A. F., VALJAVEC-GRATIAN, M. & HILL, T. M. 2001. Site-directed mutagenesis and phylogenetic comparisons of the *Escherichia coli* Tus protein: DNA-protein interactions alone can not account for Tus activity. *Mol Genet Genomics,* 265**,** 941-53.

HENDRICKSON, H. & LAWRENCE, J. G. 2007. Mutational bias suggests that replication termination occurs near the *dif* site, not at *Ter* sites. *Mol Microbiol,* 64**,** 42-56.

HENDRICKSON, W. G., KUSANO, T., YAMAKI, H., BALAKRISHNAN, R., KING, M., MURCHIE, J. & SCHAECHTER, M. 1982. Binding of the origin of replication of *Escherichia coli* to the outer membrane. *Cell,* 30**,** 915-23.

HIASA, H. & MARIANS, K. J. 1992. Differential inhibition of the DNA translocation and DNA unwinding activities of DNA helicases by the *Escherichia coli* Tus protein. *J Biol Chem,* 267**,** 11379-85.

HIDAKA, M., AKIYAMA, M. & HORIUCHI, T. 1988. A consensus sequence of three DNA replication terminus sites on the *E. coli* chromosome is highly homologous to the *terR* sites of the R6K plasmid. *Cell,* 55**,** 467-75.

HIDAKA, M., KOBAYASHI, T. & HORIUCHI, T. 1991. A newly identified DNA replication terminus site, *TerE*, on the *Escherichia coli* chromosome. *J Bacteriol,* 173**,** 391-3.

HIDAKA, M., KOBAYASHI, T., ISHIMI, Y., SEKI, M., ENOMOTO, T., ABDEL-MONEM, M. & HORIUCHI, T. 1992. Termination complex in *Escherichia coli* inhibits SV40 DNA replication *in vitro* by impeding the action of T antigen helicase. *J Biol Chem,* 267**,** 5361-5.

HIDAKA, M., KOBAYASHI, T., TAKENAKA, S., TAKEYA, H. & HORIUCHI, T. 1989. Purification of a DNA replication terminus (*ter*) site-binding protein in *Escherichia coli* and identification of the structural gene. *J Biol Chem,* 264**,** 21031-7.

HILL, T. M. 1992. Arrest of bacterial DNA replication. *Annu Rev Microbiol,* 46**,** 603-33.

HILL, T. M., HENSON, J. M. & KUEMPEL, P. L. 1987. The terminus region of the *Escherichia coli* chromosome contains two separate loci that exhibit polar inhibition of replication. *Proc Natl Acad Sci U S A,* 84**,** 1754-8.

HILL, T. M., KOPP, B. J. & KUEMPEL, P. L. 1988a. Termination of DNA replication in *Escherichia coli* requires a trans-acting factor. *J Bacteriol,* 170**,** 662-8.

HILL, T. M., PELLETIER, A. J., TECKLENBURG, M. L. & KUEMPEL, P. L. 1988b. Identification of the DNA sequence from the *E. coli* terminus region that halts replication forks. *Cell,* 55**,** 459-66.

HILL, T. M., TECKLENBURG, M. L., PELLETIER, A. J. & KUEMPEL, P. L. 1989. Tus, the trans-acting gene required for termination of DNA replication in *Escherichia coli*, encodes a DNA-binding protein. *Proc Natl Acad Sci U S A,* 86**,** 1593-7.

HOLT, I. J., LORIMER, H. E. & JACOBS, H. T. 2000. Coupled leading- and lagging-strand synthesis of mammalian mitochondrial DNA. *Cell,* 100**,** 515-24.

HORIUCHI, T. & HIDAKA, M. 1988. Core sequence of two separable terminus sites of the R6K plasmid that exhibit polar inhibition of replication is a 20 bp inverted repeat. *Cell,* 54**,** 515-23.

HORIUCHI, T., NISHITANI, H. & KOBAYASHI, T. 1995. A new type of *E. coli* recombinational hotspot which requires for activity both DNA replication termination events and the Chi sequence. *Adv Biophys,* 31**,** 133-47.

HYVARINEN, A. K., POHJOISMAKI, J. L., REYES, A., WANROOIJ, S., YASUKAWA, T., KARHUNEN, P. J., SPELBRINK, J. N., HOLT, I. J. & JACOBS, H. T. 2007. The mitochondrial transcription termination factor mTERF modulates replication pausing in human mitochondrial DNA. *Nucleic Acids Res,* 35**,** 6458-74.

INOUE, T., SHINGAKI, R., HIROSE, S., WAKI, K., MORI, H. & FUKUI, K. 2007. Genome-wide screening of genes required for swarming motility in *Escherichia coli* K-12. *J Bacteriol,* 189**,** 950-7.

ISAACS, F. J., CARR, P. A., WANG, H. H., LAJOIE, M. J., STERLING, B., KRAAL, L., TOLONEN, A. C., GIANOULIS, T. A., GOODMAN, D. B., REPPAS, N. B., EMIG, C. J., BANG, D., HWANG, S. J., JEWETT, M. C., JACOBSON, J. M. & CHURCH, G. M. 2011. Precise manipulation of chromosomes *in vivo* enables genome-wide codon replacement. *Science,* 333**,** 348-53.

ISHII, M., KUNIMURA, J. S., JENG, H. T., PENNA, T. C. & CHOLEWA, O. 2007a. Evaluation of the pH- and thermal stability of the recombinant green fluorescent protein (GFP) in the presence of sodium chloride. *Appl Biochem Biotechnol,* 137-140**,** 555-71.

ISHII, M., KUNIMURA, J. S., PENNA, T. C. & CHOLEWA, O. 2007b. Study on the thermal stability of green fluorescent protein (GFP) in glucose parenteral formulations. *Int J Pharm,* 337**,** 109-17.

ISHIKAWA, S., OGURA, Y., YOSHIMURA, M., OKUMURA, H., CHO, E., KAWAI, Y., KUROKAWA, K., OSHIMA, T. & OGASAWARA, N. 2007. Distribution of stable DnaA-binding sites on the *Bacillus subtilis* genome detected using a modified ChIP-chip method. *DNA Res,* 14**,** 155-68.

ISHINO, Y., SHINAGAWA, H., MAKINO, K., AMEMURA, M. & NAKATA, A. 1987. Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J Bacteriol,* 169**,** 5429-33.

JELESAROV, I. & BOSSHARD, H. R. 1999. Isothermal titration calorimetry and differential scanning calorimetry as complementary tools to investigate the energetics of biomolecular recognition. *J Mol Recognit,* 12**,** 3-18.

JENSEN, K. F., LARSEN, J. N., SCHACK, L. & SIVERTSEN, A. 1984. Studies on the structure and expression of *Escherichia coli* pyrC, pyrD, and pyrF using the cloned genes. *Eur J Biochem,* 140**,** 343-52.

JERGIC, S., OZAWA, K., WILLIAMS, N. K., SU, X. C., SCOTT, D. D., HAMDAN, S. M., CROWTHER, J. A., OTTING, G. & DIXON, N. E. 2007. The unstructured C-terminus of the tau subunit of *Escherichia coli* DNA polymerase III holoenzyme is the site of interaction with the alpha subunit. *Nucleic Acids Res,* 35**,** 2813-24.

JERUZALMI, D., O'DONNELL, M. & KURIYAN, J. 2001. Crystal structure of the processivity clamp loader gamma (gamma) complex of *E. coli* DNA polymerase III. *Cell,* 106**,** 429-41.

JIANG, M., CHEN, X., GUO, Z. F., CAO, Y., CHEN, M. & GUO, Z. 2008. Identification and characterization of (1R,6R)-2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase in the menaquinone biosynthesis of *Escherichia coli*. *Biochemistry,* 47**,** 3426-34.

JOHNSON, A. & O'DONNELL, M. 2005. Cellular DNA replicases: components and dynamics at the replication fork. *Annu Rev Biochem,* 74**,** 283-315.

JOYCE, C. M. & GRINDLEY, N. D. 1984. Method for determining whether a gene of *Escherichia coli* is essential: application to the polA gene. *J Bacteriol,* 158**,** 636-43.

KACZMARCZYK, S. J., SITARAMAN, K., HILL, T., HARTLEY, J. L. & CHATTERJEE, D. K. 2010. Tus, an *E. coli* protein, contains mammalian nuclear targeting and exporting signals. *PLoS One,* 5**,** e8889.

KAGUNI, J. M. 2006. DnaA: controlling the initiation of bacterial DNA replication and more. *Annu Rev Microbiol,* 60**,** 351-75.

KAGUNI, J. M. 2011. Replication initiation at the *Escherichia coli* chromosomal origin. *Curr Opin Chem Biol,* 15**,** 606-13.

KAMADA, K., HORIUCHI, T., OHSUMI, K., SHIMAMOTO, N. & MORIKAWA, K. 1996. Structure of a replication-terminator protein complexed with DNA. *Nature,* 383**,** 598-603.

KAPLAN, D. L. 2006. Replication termination: mechanism of polar arrest revealed. *Curr Biol,* 16**,** R684-6.

KASIVISWANATHAN, R., COLLINS, T. R. & COPELAND, W. C. 2012. The interface of transcription and DNA replication in the mitochondria. *Biochim Biophys Acta,* 1819**,** 970-8.

KATAYAMA, T. & SEKIMIZU, K. 1999. Inactivation of *Escherichia coli* DnaA protein by DNA polymerase III and negative regulations for initiation of chromosomal replication. *Biochimie,* 81**,** 835-40.

KAWAKAMI, H., KEYAMURA, K. & 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**,** 27420-30.

KELMAN, Z. & O'DONNELL, M. 1995. DNA polymerase III holoenzyme: structure and function of a chromosomal replicating machine. *Annu Rev Biochem,* 64**,** 171-200.

KELMAN, Z., YUZHAKOV, A., ANDJELKOVIC, J. & O'DONNELL, M. 1998. Devoted to the lagging strand-the subunit of DNA polymerase III holoenzyme contacts SSB to promote processive elongation and sliding clamp assembly. *EMBO J,* 17**,** 2436-49.

KEYAMURA, K., FUJIKAWA, N., ISHIDA, T., OZAKI, S., SU'ETSUGU, M., FUJIMITSU, K., KAGAWA, W., YOKOYAMA, S., KURUMIZAKA, H. & 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**,** 2083-99.

KHATRI, G. S., MACALLISTER, T., SISTA, P. R. & BASTIA, D. 1989. The replication terminator protein of *E. coli* is a DNA sequence-specific contra-helicase. *Cell,* 59**,** 667-74.

KIM, D. R. & MCHENRY, C. S. 1996. Identification of the beta-binding domain of the alpha subunit of *Escherichia coli* polymerase III holoenzyme. *J Biol Chem,* 271**,** 20699-704.

KIM, W., KIM, Y., MIN, J., KIM, D. J., CHANG, Y. T. & HECHT, M. H. 2006. A high-throughput screen for compounds that inhibit aggregation of the Alzheimer's peptide. *ACS Chem Biol,* 1**,** 461-9.

KITANI, T., YODA, K., OGAWA, T. & OKAZAKI, T. 1985. Evidence that discontinuous DNA replication in *Escherichia coli* is primed by approximately 10 to 12 residues of RNA starting with a purine. *J Mol Biol,* 184**,** 45-52.

KOBAYASHI, T., HIDAKA, M. & HORIUCHI, T. 1989. Evidence of a *Ter* specific binding protein essential for the termination reaction of DNA replication in *Escherichia coli*. *EMBO J,* 8**,** 2435-41.

KOBORI, J. A. & KORNBERG, A. 1982. The *Escherichia coli* dnaC gene product. III. Properties of the DnaB-DnaC protein complex. *J Biol Chem,* 257**,** 13770-5.

KOGA, Y., KATSUMI, R., YOU, D. J., MATSUMURA, H., TAKANO, K. & KANAYA, S. 2008. Crystal structure of highly thermostable glycerol kinase from a hyperthermophilic archaeon in a dimeric form. *Febs J,* 275**,** 2632-43.

KONIECZNY, I. 2003. Strategies for helicase recruitment and loading in bacteria. *EMBO Rep,* 4**,** 37-41.

KONO, N., ARAKAWA, K. & TOMITA, M. 2012. Validation of bacterial replication termination models using simulation of genomic mutations. *PLoS One,* 7**,** e34526.

KRINGS, G. & BASTIA, D. 2005. Sap1p binds to Ter1 at the ribosomal DNA of *Schizosaccharomyces pombe* and causes polar replication fork arrest. *J Biol Chem,* 280**,** 39135-42.

KUCHERER, C., LOTHER, H., KOLLING, R., SCHAUZU, M. A. & MESSER, W. 1986. Regulation of transcription of the chromosomal DnaA gene of *Escherichia coli*. *Mol Gen Genet,* 205**,** 115-21.

KUEMPEL, P. L., DUERR, S. A. & SEELEY, N. R. 1977. Terminus region of the chromosome in *Escherichia coli* inhibits replication forks. *Proc Natl Acad Sci U S A,* 74**,** 3927-31.

KUNZELMANN, S., MORRIS, C., CHAVDA, A. P., ECCLESTON, J. F. & WEBB, M. R. 2010. Mechanism of interaction between single-stranded DNA binding protein and DNA. *Biochemistry,* 49**,** 843-52.

KUZNETSOV, S. V., KOZLOV, A. G., LOHMAN, T. M. & ANSARI, A. 2006. Microsecond dynamics of protein-DNA interactions: direct observation of the wrapping/unwrapping kinetics of single-stranded DNA around the *E. coli* SSB tetramer. *J Mol Biol,* 359**,** 55-65.

LANGE, R. P., LOCHER, H. H., WYSS, P. C. & THEN, R. L. 2007. The targets of currently used antibacterial agents: lessons for drug discovery. *Curr Pharm Des,* 13**,** 3140-54.

LANGLEY, D. B., SMITH, M. T., LEWIS, P. J. & WAKE, R. G. 1993. Protein-nucleoside contacts in the interaction between the replication terminator protein of *Bacillus subtilis* and the DNA terminator. *Mol Microbiol,* 10**,** 771-9.

LANKA, E. & SCHUSTER, H. 1983. The DnaC protein of *Escherichia coli*. Purification, physical properties and interaction with dnaB protein. *Nucleic Acids Res,* 11**,** 987-97.

LAVINDER, J. J., HARI, S. B., SULLIVAN, B. J. & MAGLIERY, T. J. 2009. High-throughput thermal scanning: a general, rapid dye-binding thermal shift screen for protein engineering. *J Am Chem Soc,* 131**,** 3794-5.

LEE, E. H., KORNBERG, A., HIDAKA, M., KOBAYASHI, T. & HORIUCHI, T. 1989. *Escherichia coli* replication termination protein impedes the action of helicases. *Proc Natl Acad Sci U S A,* 86**,** 9104-8.

LEE, J., CHASTAIN, P. D., KUSAKABE, T., GRIFFITH, J. D. & RICHARDSON, C. C. 1998. Coordinated leading and lagging strand DNA synthesis on a minicircular template. *Mol Cell,* 1**,** 1001-10.

LEE, J. B., HITE, R. K., HAMDAN, S. M., XIE, X. S., RICHARDSON, C. C. & VAN OIJEN, A. M. 2006. DNA primase acts as a molecular brake in DNA replication. *Nature,* 439**,** 621-4.

LEE, P. S. & LEE, K. H. 2003. *Escherichia coli*-a model system that benefits from and contributes to the evolution of proteomics. *Biotechnol Bioeng,* 84**,** 801-14.

LEHMAN, I. R. 1974. DNA ligase: structure, mechanism, and function. *Science,* 186**,** 790-7.

LEIBOWITZ, P. J. & SCHAECHTER, M. 1975. The attachment of the bacterial chromosome to the cell membrane. *Int Rev Cytol,* 41**,** 1-28.

LEMON, K. P. & GROSSMAN, A. D. 1998. Localization of bacterial DNA polymerase: evidence for a factory model of replication. *Science,* 282**,** 1516-9.

LEMON, K. P. & GROSSMAN, A. D. 2000. Movement of replicating DNA through a stationary replisome. *Mol Cell,* 6**,** 1321-30.

LEONARD, A. C. & GRIMWADE, J. E. 2005. Building a bacterial orisome: emergence of new regulatory features for replication origin unwinding. *Mol Microbiol,* 55**,** 978-85.

LEPOCK, J. R., RITCHIE, K. P., KOLIOS, M. C., RODAHL, A. M., HEINZ, K. A. & KRUUV, J. 1992. Influence of transition rates and scan rate on kinetic simulations of differential scanning calorimetry profiles of reversible and irreversible protein denaturation. *Biochemistry,* 31**,** 12706-12.

LEUNG, S. M., SENISTERRA, G., RITCHIE, K. P., SADIS, S. E., LEPOCK, J. R. & HIGHTOWER, L. E. 1996. Thermal activation of the bovine Hsc70 molecular chaperone at physiological temperatures: physical evidence of a molecular thermometer. *Cell Stress Chaperones,* 1**,** 78-89.

LIA, G., MICHEL, B. & ALLEMAND, J. F. 2012. Polymerase exchange during Okazaki fragment synthesis observed in living cells. *Science,* 335**,** 328-31.

LO, M. C., AULABAUGH, A., JIN, G., COWLING, R., BARD, J., MALAMAS, M. & ELLESTAD, G. 2004. Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Anal Biochem,* 332**,** 153-9.

LOBRY, J. R. 1996. Asymmetric substitution patterns in the two DNA strands of bacteria. *Mol Biol Evol,* 13**,** 660-5.

LOHMAN, T. M. & FERRARI, M. E. 1994. *Escherichia coli* single-stranded DNA-binding protein: multiple DNA-binding modes and cooperativities. *Annu Rev Biochem,* 63**,** 527-70.

LOMAKIN, I. B., SHIROKIKH, N. E., YUSUPOV, M. M., HELLEN, C. U. & PESTOVA, T. V. 2006. The fidelity of translation initiation: reciprocal activities of eIF1, IF3 and YciH. *EMBO J,* 25**,** 196-210.

LOPEZ DE SARO, F. J., GEORGESCU, R. E. & O'DONNELL, M. 2003. A peptide switch regulates DNA polymerase processivity. *Proc Natl Acad Sci U S A,* 100**,** 14689-94.

LOPEZ DE SARO, F. J. & O'DONNELL, M. 2001. Interaction of the beta sliding clamp with MutS, ligase, and DNA polymerase I. *Proc Natl Acad Sci U S A,* 98**,** 8376-80.

LOUARN, J., PATTE, J. & LOUARN, J. M. 1977. Evidence for a fixed termination site of chromosome replication in *Escherichia coli* K12. *J Mol Biol,* 115**,** 295-314.

LOUARN, J., PATTE, J. & LOUARN, J. M. 1979. Map position of the replication terminus on the *Escherichia coli* chromosome. *Mol Gen Genet,* 172**,** 7-11.

MADSEN, C. S., GHIVIZZANI, S. C. & HAUSWIRTH, W. W. 1993. Protein binding to a single termination-associated sequence in the mitochondrial DNA D-loop region. *Mol Cell Biol,* 13**,** 2162-71.

MAERKL, S. J. & QUAKE, S. R. 2007. A systems approach to measuring the binding energy landscapes of transcription factors. *Science,* 315**,** 233-7.

MAGLIERY, T. J., LAVINDER, J. J. & SULLIVAN, B. J. 2011. Protein stability by number: high-throughput and statistical approaches to one of protein science's most difficult problems. *Curr Opin Chem Biol,* 15**,** 443-51.

MAKISE, M., TSUCHIYA, T. & MIZUSHIMA, T. 2002. Sequence-independent DNA binding activity of DnaA protein, the initiator of chromosomal DNA replication in *Escherichia coli*. *J Biochem,* 131**,** 419-25.

MANNING, G. S. 1978. The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. *Q Rev Biophys,* 11**,** 179-246.

MARGULIES, C. & KAGUNI, J. M. 1996. Ordered and sequential binding of DnaA protein to *oriC*, the chromosomal origin of *Escherichia coli*. *J Biol Chem,* 271**,** 17035-40.

MARSZALEK, J. & KAGUNI, J. M. 1994. DnaA protein directs the binding of DnaB protein in initiation of DNA replication in *Escherichia coli*. *J Biol Chem,* 269**,** 4883-90.

MARSZALEK, J., ZHANG, W., HUPP, T. R., MARGULIES, C., CARR, K. M., CHERRY, S. & KAGUNI, J. M. 1996. Domains of DnaA protein involved in interaction with DnaB protein, and in unwinding the *Escherichia coli* chromosomal origin. *J Biol Chem,* 271**,** 18535-42.

MATULIS, D., KRANZ, J. K., SALEMME, F. R. & TODD, M. J. 2005. Thermodynamic stability of carbonic anhydrase: measurements of binding affinity and stoichiometry using ThermoFluor. *Biochemistry,* 44**,** 5258-66.

MCGARRY, K. C., RYAN, V. T., GRIMWADE, J. E. & LEONARD, A. C. 2004. Two discriminatory binding sites in the *Escherichia coli* replication origin are required for DNA strand opening by initiator DnaA-ATP. *Proc Natl Acad Sci U S A,* 101**,** 2811-6.

MCHENRY, C. S. 2003. Chromosomal replicases as asymmetric dimers: studies of subunit arrangement and functional consequences. *Mol Microbiol,* 49**,** 1157-65.

MCHUGH, J. P., RODRIGUEZ-QUINONES, F., ABDUL-TEHRANI, H., SVISTUNENKO, D. A., POOLE, R. K., COOPER, C. E. & ANDREWS, S. C. 2003. Global iron-dependent gene regulation in *Escherichia coli*. A new mechanism for iron homeostasis. *J Biol Chem,* 278**,** 29478-86.

MENG, F. G., HONG, Y. K., HE, H. W., LYUBAREV, A. E., KURGANOV, B. I., YAN, Y. B. & ZHOU, H. M. 2004. Osmophobic effect of glycerol on irreversible thermal denaturation of rabbit creatine kinase. *Biophys J,* 87**,** 2247-54.

MESSER, W. 2002. The bacterial replication initiator DnaA. DnaA and *oriC*, the bacterial mode to initiate DNA replication. *FEMS Microbiol Rev,* 26**,** 355-74.

MEYER, R. R. & LAINE, P. S. 1990. The single-stranded DNA-binding protein of *Escherichia coli*. *Microbiol Rev,* 54**,** 342-80.

MILLER, D. T., GRIMWADE, J. E., BETTERIDGE, T., ROZGAJA, T., TORGUE, J. J. & LEONARD, A. C. 2009. Bacterial origin recognition complexes direct assembly of higher-order DnaA oligomeric structures. *Proc Natl Acad Sci U S A,* 106**,** 18479-84.

MINTON, A. P. 2001. The influence of macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media. *J Biol Chem,* 276**,** 10577-80.

MITKOVA, A. V., KHOPDE, S. M. & BISWAS, S. B. 2003. Mechanism and stoichiometry of interaction of DnaG primase with DnaB helicase of *Escherichia coli* in RNA primer synthesis. *J Biol Chem,* 278**,** 52253-61.

MOHANTY, B. K., BAIRWA, N. K. & BASTIA, D. 2009. Contrasting roles of checkpoint proteins as recombination modulators at Fob1-*Ter* complexes with or without fork arrest. *Eukaryot Cell,* 8**,** 487-95.

MOHANTY, B. K. & BASTIA, D. 2004. Binding of the replication terminator protein Fob1p to the *Ter* sites of yeast causes polar fork arrest. *J Biol Chem,* 279**,** 1932-41.

MOHANTY, B. K., SAHOO, T. & BASTIA, D. 1996. The relationship between sequence-specific termination of DNA replication and transcription. *EMBO J,* 15**,** 2530-9.

MOHANTY, B. K., SAHOO, T. & BASTIA, D. 1998. Mechanistic studies on the impact of transcription on sequence-specific termination of DNA replication and vice versa. *J Biol Chem,* 273**,** 3051-9.

MOLINA, F. & SKARSTAD, K. 2004. Replication fork and SeqA focus distributions in *Escherichia coli* suggest a replication hyperstructure dependent on nucleotide metabolism. *Mol Microbiol,* 52**,** 1597-612.

MOREAU, M. J., MORIN, I. & SCHAEFFER, P. M. 2010. Quantitative determination of protein stability and ligand binding using a green fluorescent protein reporter system. *Mol Biosyst,* 6**,** 1285-92.

MOREAU, M. J. & SCHAEFFER, P. M. 2012a. Differential Tus-*Ter* binding and lock formation: implications for DNA replication termination in *Escherichia coli*. *Mol Biosyst,* 8**,** 2783-91.

MOREAU, M. J. & SCHAEFFER, P. M. 2012b. A polyplex qPCR-based binding assay for protein-DNA interactions. *Analyst,* 137**,** 4111-3.

MOREAU, M. J. J., MORIN, I., ASKIN, S. P., COOPER, A., MORELAND, N. J., VASUDEVAN, S. G. & SCHAEFFER, P. M. 2012. Rapid determination of protein stability and ligand binding by differential scanning fluorimetry of GFP-tagged proteins. *RSC Advances,* 2**,** 11892-11900.

MORIN, I., ASKIN, S. P. & SCHAEFFER, P. M. 2011. IgG-detection devices for the Tus-*Ter*-lock immuno-PCR diagnostic platform. *Analyst,* 136**,** 4815-21.

MORIN, I., DIXON, N. E. & SCHAEFFER, P. M. 2010. Ultrasensitive detection of antibodies using a new Tus-*Ter*-lock immunoPCR system. *Mol Biosyst,* 6**,** 1173-5.

MORIN, I., SCHAEFFER, P. M., ASKIN, S. P. & DIXON, N. E. 2012. Combining RNA-DNA swapping and quantitative polymerase chain reaction for the detection of influenza A nucleoprotein

Development of a protease activity assay using heat-sensitive Tus-GFP fusion protein substrates

Ultrasensitive detection of antibodies using a new Tus-*Ter*-lock immunoPCR system. *Anal Biochem,* 420**,** 121-6.

MOTT, M. L. & BERGER, J. M. 2007. DNA replication initiation: mechanisms and regulation in bacteria. *Nat Rev Microbiol,* 5**,** 343-54.

MOTT, M. L., ERZBERGER, J. P., COONS, M. M. & BERGER, J. M. 2008. Structural Synergy and Molecular Crosstalk between Bacterial Helicase Loaders and Replication Initiators. *Cell,* 135**,** 623-634.

MULCAIR, M. D., SCHAEFFER, P. M., OAKLEY, A. J., CROSS, H. F., NEYLON, C., HILL, T. M. & DIXON, N. E. 2006. A molecular mousetrap determines polarity of termination of DNA replication in *E. coli*. *Cell,* 125**,** 1309-19.

MULUGU, S., POTNIS, A., SHAMSUZZAMAN, TAYLOR, J., ALEXANDER, K. & BASTIA, D. 2001. Mechanism of termination of DNA replication of *Escherichia coli* involves helicase-contrahelicase interaction. *Proc Natl Acad Sci U S A,* 98**,** 9569-74.

NAKTINIS, V., TURNER, J. & O'DONNELL, M. 1996. A molecular switch in a replication machine defined by an internal competition for protein rings. *Cell,* 84**,** 137-45.

NATARAJAN, S., KAUL, S., MIRON, A. & BASTIA, D. 1993. A 27 kd protein of *E. coli* promotes antitermination of replication *in vitro* at a sequence-specific replication terminus. *Cell,* 72**,** 113-20.

NATARAJAN, S., KELLEY, W. L. & BASTIA, D. 1991. Replication terminator protein of *Escherichia coli* is a transcriptional repressor of its own synthesis. *Proc Natl Acad Sci U S A,* 88**,** 3867-71.

NEIDHARDT, F. C. 1987. *Escherichia coli* and *Salmonella typhimurium:* cellular and molecular biology, American Society for Microbiology.

NEUWALD, A. F., ARAVIND, L., SPOUGE, J. L. & KOONIN, E. V. 1999. AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res,* 9**,** 27-43.

NEYLON, C., BROWN, S. E., KRALICEK, A. V., MILES, C. S., LOVE, C. A. & DIXON, N. E. 2000. Interaction of the *Escherichia coli* replication terminator protein (Tus)

with DNA: a model derived from DNA-binding studies of mutant proteins by surface plasmon resonance. *Biochemistry,* 39**,** 11989-99.

NEYLON, C., KRALICEK, A. V., HILL, T. M. & DIXON, N. E. 2005. Replication termination in *Escherichia coli*: structure and antihelicase activity of the Tus-*Ter* complex. *Microbiol Mol Biol Rev,* 69**,** 501-26.

NIBA, E. T., NAKA, Y., NAGASE, M., MORI, H. & KITAKAWA, M. 2007. A genome-wide approach to identify the genes involved in biofilm formation in *E. coli. DNA Res,* 14**,** 237-46.

NIELSEN, H. J., LI, Y., YOUNGREN, B., HANSEN, F. G. & AUSTIN, S. 2006. Progressive segregation of the *Escherichia coli* chromosome. *Mol Microbiol,* 61**,** 383-93.

NIELSEN, H. J., YOUNGREN, B., HANSEN, F. G. & AUSTIN, S. 2007. Dynamics of *Escherichia coli* chromosome segregation during multifork replication. *J Bacteriol,* 189**,** 8660-6.

NIESEN, F. H., BERGLUND, H. & VEDADI, M. 2007. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc,* 2**,** 2212-21.

NIEVERA, C., TORGUE, J. J., GRIMWADE, J. E. & LEONARD, A. C. 2006. SeqA blocking of DnaA-*oriC* interactions ensures staged assembly of the *E. coli* pre-RC. *Mol Cell,* 24**,** 581-92.

NOVIKOVA, M., METLITSKAYA, A., DATSENKO, K., KAZAKOV, T., KAZAKOV, A., WANNER, B. & SEVERINOV, K. 2007. The *Escherichia coli* Yej transporter is required for the uptake of translation inhibitor microcin C. *J Bacteriol,* 189**,** 8361-5.

O'DONNELL, M. 2006. Replisome architecture and dynamics in *Escherichia coli*. *J Biol Chem,* 281**,** 10653-6.

O'DONNELL, M., JERUZALMI, D. & KURIYAN, J. 2001. Clamp loader structure predicts the architecture of DNA polymerase III holoenzyme and RFC. *Curr Biol,* 11**,** R935-46.

OAKLEY, A. J., LOSCHA, K. V., SCHAEFFER, P. M., LIEPINSH, E., PINTACUDA, G., WILCE, M. C., OTTING, G. & DIXON, N. E. 2005. Crystal and solution structures of the helicase-binding domain of *Escherichia coli* primase. *J Biol Chem,* 280**,** 11495-504.

OEHLER, S., ALEX, R. & BARKER, A. 1999. Is nitrocellulose filter binding really a universal assay for protein-DNA interactions? *Anal Biochem,* 268**,** 330-6.

OKAZAKI, R., ARISAWA, M. & SUGINO, A. 1971. Slow joining of newly replicated DNA chains in DNA polymerase I-deficient *Escherichia coli* mutants. *Proc Natl Acad Sci U S A,* 68**,** 2954-7.

OLSON, M. W., DALLMANN, H. G. & MCHENRY, C. S. 1995. DnaX complex of *Escherichia coli* DNA polymerase III holoenzyme. The chi psi complex functions by increasing the affinity of tau and gamma for delta.delta' to a physiologically relevant range. *J Biol Chem,* 270**,** 29570-7.

OSHIMA, T., ISHIKAWA, S., KUROKAWA, K., AIBA, H. & OGASAWARA, N. 2006. *Escherichia coli* histone-like protein H-NS preferentially binds to horizontally acquired DNA in association with RNA polymerase. *DNA Res,* 13**,** 141-53.

OZAKI, S. & KATAYAMA, T. 2009. DnaA structure, function, and dynamics in the initiation at the chromosomal origin. *Plasmid.*

OZAKI, S. & KATAYAMA, T. 2012. Highly organized DnaA-*oriC* complexes recruit the single-stranded DNA for replication initiation. *Nucleic Acids Res,* 40**,** 1648-65.

OZAKI, S., KAWAKAMI, H., NAKAMURA, K., FUJIKAWA, N., KAGAWA, W., PARK, S. Y., YOKOYAMA, S., KURUMIZAKA, H. & KATAYAMA, T. 2008. A common

mechanism for the ATP-DnaA-dependent formation of open complexes at the replication origin. *J Biol Chem,* 283**,** 8351-62.

OZAWA, K., HEADLAM, M. J., MOURADOV, D., WATT, S. J., BECK, J. L., RODGERS, K. J., DEAN, R. T., HUBER, T., OTTING, G. & DIXON, N. E. 2005. Translational incorporation of L-3,4-dihydroxyphenylalanine into proteins. *Febs J,* 272**,** 3162-71.

PAGES, V., MAZON, G., NAIMAN, K., PHILIPPIN, G. & FUCHS, R. P. 2012. Monitoring bypass of single replication-blocking lesions by damage avoidance in the *Escherichia coli* chromosome. *Nucleic Acids Res,* 40**,** 9036-43.

PANCHENKO, T., ZHU, W. W. & MONTCLARE, J. K. 2006. Influence of global fluorination on chloramphenicol acetyltransferase activity and stability. *Biotechnol Bioeng,* 94**,** 921-30.

PANDEY, M., SYED, S., DONMEZ, I., PATEL, G., HA, T. & PATEL, S. S. 2009. Coordinating DNA replication by means of priming loop and differential synthesis rate. *Nature,* 462**,** 940-3.

PANTOLIANO, M. W., PETRELLA, E. C., KWASNOSKI, J. D., LOBANOV, V. S., MYSLIK, J., GRAF, E., CARVER, T., ASEL, E., SPRINGER, B. A., LANE, P. & SALEMME, F. R. 2001. High-density miniaturized thermal shift assays as a general strategy for drug discovery. *J Biomol Screen,* 6**,** 429-40.

PATEL, S. S. & PICHA, K. M. 2000. Structure and function of hexameric helicases. *Annu Rev Biochem,* 69**,** 651-97.

PELLEGRINI, M., ASIN-CAYUELA, J., ERDJUMENT-BROMAGE, H., TEMPST, P., LARSSON, N. G. & GUSTAFSSON, C. M. 2009. MTERF2 is a nucleoid component in mammalian mitochondria. *Biochim Biophys Acta,* 1787**,** 296-302.

PENNA, T. C., ISHII, M., KUNIMURA, J. S. & CHOLEWA, O. 2005. Stability of recombinant green fluorescent protein (GFPuv) in glucose solutions at different concentrations and pH values. *Appl Biochem Biotechnol,* 121-124**,** 501-27.

POLOSA, P. L., DECEGLIE, S., ROBERTI, M., GADALETA, M. N. & CANTATORE, P. 2005. Contrahelicase activity of the mitochondrial transcription termination factor mtDBP. *Nucleic Acids Res,* 33**,** 3812-20.

POMERANTZ, R. T. & O'DONNELL, M. 2007. Replisome mechanics: insights into a twin DNA polymerase machine. *Trends Microbiol,* 15**,** 156-64.

POTEETE, A. R. 2001. What makes the bacteriophage lambda Red system useful for genetic engineering: molecular mechanism and biological function. *FEMS Microbiol Lett,* 201**,** 9-14.

RAGHUNATHAN, S., KOZLOV, A. G., LOHMAN, T. M. & WAKSMAN, G. 2000. Structure of the DNA binding domain of *E. coli* SSB bound to ssDNA. *Nat Struct Biol,* 7**,** 648-52.

RAGHUNATHAN, S., RICARD, C. S., LOHMAN, T. M. & WAKSMAN, G. 1997. Crystal structure of the homo-tetrameric DNA binding domain of *Escherichia coli* single-stranded DNA-binding protein determined by multiwavelength x-ray diffraction on the selenomethionyl protein at 2.9-A resolution. *Proc Natl Acad Sci U S A,* 94**,** 6652-7.

RECORD, M. T., JR., ANDERSON, C. F. & LOHMAN, T. M. 1978. Thermodynamic analysis of ion effects on the binding and conformational equilibria of proteins and nucleic acids: the roles of ion association or release, screening, and ion effects on water activity. *Q Rev Biophys,* 11**,** 103-78.

RECORD, M. T., JR., HA, J. H. & FISHER, M. A. 1991. Analysis of equilibrium and kinetic measurements to determine thermodynamic origins of stability and specificity and mechanism of formation of site-specific complexes between proteins and helical DNA. *Methods Enzymol,* 208**,** 291-343.

RECORD, M. T., JR., LOHMAN, M. L. & DE HASETH, P. 1976. Ion effects on ligand-nucleic acid interactions. *J Mol Biol,* 107**,** 145-58.

REGEV, T., MYERS, N., ZARIVACH, R. & FISHOV, I. 2012. Association of the chromosome replication initiator DnaA with the *Escherichia coli* inner membrane *in vivo*: quantity and mode of binding. *PLoS One,* 7**,** e36441.

REPOILA, F. & DARFEUILLE, F. 2009. Small regulatory non-coding RNAs in bacteria: physiology and mechanistic aspects. *Biol Cell,* 101**,** 117-31.

REYES-LAMOTHE, R., POSSOZ, C., DANILOVA, O. & SHERRATT, D. J. 2008a. Independent positioning and action of *Escherichia coli* replisomes in live cells. *Cell,* 133**,** 90-102.

REYES-LAMOTHE, R., SHERRATT, D. J. & LEAKE, M. C. 2010. Stoichiometry and architecture of active DNA replication machinery in *Escherichia coli*. *Science,* 328**,** 498-501.

REYES-LAMOTHE, R., WANG, X. & SHERRATT, D. 2008b. *Escherichia coli* and its chromosome. *Trends Microbiol,* 16**,** 238-45.

ROCHA, E. P. 2004. The replication-related organization of bacterial genomes. *Microbiology,* 150**,** 1609-27.

RODINA, A. & GODSON, G. N. 2006. Role of Conserved Amino Acids in the Catalytic Activity of *Escherichia coli* Primase. *J Bacteriol,* 188**,** 3614-21.

ROECKLEIN, B., PELLETIER, A. & 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**,** 169-75.

ROECKLEIN, B. A. & KUEMPEL, P. L. 1992. *In vivo* characterization of *tus* gene expression in *Escherichia coli*. *Mol Microbiol,* 6**,** 1655-61.

ROTHSTEIN, R., MICHEL, B. & GANGLOFF, S. 2000. Replication fork pausing and recombination or "gimme a break". *Genes Dev,* 14**,** 1-10.

ROWEN, L. & KORNBERG, A. 1978. Primase, the DnaG protein of *Escherichia coli*. An enzyme which starts DNA chains. *J Biol Chem,* 253**,** 758-64.

ROY, R., KOZLOV, A. G., LOHMAN, T. M. & HA, T. 2007. Dynamic structural rearrangements between DNA binding modes of *E. coli* SSB protein. *J Mol Biol,* 369**,** 1244-57.

RUPP, W. D. & HOWARD-FLANDERS, P. 1968. Discontinuities in the DNA synthesized in an excision-defective strain of *Escherichia coli* following ultraviolet irradiation. *J Mol Biol,* 31**,** 291-304.

RYTER, A., HIROTA, Y. & JACOB, F. 1968. DNA-membrane complex and nuclear segregation in bacteria. *Cold Spring Harb Symp Quant Biol,* 33**,** 669-76.

SAECKER, R. M. 2001. Protein–DNA Interactions: Polyelectrolyte Effects. *eLS.* John Wiley & Sons, Ltd.

SALINAS, F. & BENKOVIC, S. J. 2000. Characterization of bacteriophage T4-coordinated leading- and lagging-strand synthesis on a minicircle substrate. *Proc Natl Acad Sci U S A,* 97**,** 7196-201.

SAMITT, C. E., HANSEN, F. G., MILLER, J. F. & SCHAECHTER, M. 1989. *In vivo* studies of DnaA binding to the origin of replication of *Escherichia coli*. *EMBO J,* 8**,** 989-93.

SCHAEFFER, P. M. & DIXON, N. E. 2009. Synthesis and applications of covalent protein-DNA conjugates. *Aust J Chem,* 62**,** 1328-1332.

SCHAEFFER, P. M., HEADLAM, M. J. & DIXON, N. E. 2005. Protein-protein interactions in the eubacterial replisome. *IUBMB Life,* 57**,** 5-12.

SCHAPER, S. & MESSER, W. 1995. Interaction of the initiator protein DnaA of *Escherichia coli* with its DNA target. *J Biol Chem,* 270**,** 17622-6.

SCHELLMAN, J. A. 1975. Macromolecular binding. *Biopolymers,* 14**,** 999-1018.

SCHEUERMANN, R., TAM, S., BURGERS, P. M., LU, C. & ECHOLS, H. 1983. Identification of the epsilon-subunit of *Escherichia coli* DNA polymerase III holoenzyme as the DnaQ gene product: a fidelity subunit for DNA replication. *Proc Natl Acad Sci U S A,* 80**,** 7085-9.

SCHLACHER, K., PHAM, P., COX, M. M. & GOODMAN, M. F. 2006. Roles of DNA polymerase V and RecA protein in SOS damage-induced mutation. *Chem Rev,* 106**,** 406-19.

SCOLARI, V. F., BASSETTI, B., SCLAVI, B. & LAGOMARSINO, M. C. 2011. Gene clusters reflecting macrodomain structure respond to nucleoid perturbations. *Mol Biosyst,* 7**,** 878-88.

SEITZ, H., WEIGEL, C. & MESSER, W. 2000. The interaction domains of the DnaA and DnaB replication proteins of *Escherichia coli. Mol Microbiol,* 37**,** 1270-9.

SEKIMIZU, K., BRAMHILL, D. & KORNBERG, A. 1987. ATP activates DnaA protein in initiating replication of plasmids bearing the origin of the *E. coli* chromosome. *Cell,* 50**,** 259-65.

SEKIMIZU, K., YUNG, B. Y. & KORNBERG, A. 1988. The DnaA protein of *Escherichia coli*. Abundance, improved purification, and membrane binding. *J Biol Chem,* 263**,** 7136-40.

SENISTERRA, G. A. & FINERTY, P. J., JR. 2009. High throughput methods of assessing protein stability and aggregation. *Mol Biosyst,* 5**,** 217-23.

SENISTERRA, G. A., MARKIN, E., YAMAZAKI, K., HUI, R., VEDADI, M. & AWREY, D. E. 2006. Screening for ligands using a generic and high-throughput light-scattering-based assay. *J Biomol Screen,* 11**,** 940-8.

SENISTERRA, G. A., SOO HONG, B., PARK, H. W. & VEDADI, M. 2008. Application of high-throughput isothermal denaturation to assess protein stability and screen for ligands. *J Biomol Screen,* 13**,** 337-42.

SEUFERT, W. & MESSER, W. 1987. DnaA protein binding to the plasmid origin region can substitute for primosome assembly during replication of pBR322 *in vitro*. *Cell,* 48**,** 73-8.

SHARMA, B. & HILL, T. M. 1992. *TerF*, the sixth identified replication arrest site in *Escherichia coli*, is located within the *rcsC* gene. *J Bacteriol,* 174**,** 7854-8.

SHARMA, B. & HILL, T. M. 1995. Insertion of inverted *Ter* sites into the terminus region of the *Escherichia coli* chromosome delays completion of DNA replication and disrupts the cell cycle. *Mol Microbiol,* 18**,** 45-61.

SIDOROVA, N. Y. & RAU, D. C. 1996. Differences in water release for the binding of EcoRI to specific and nonspecific DNA sequences. *Proc Natl Acad Sci U S A,* 93**,** 12272-7.

SINGH, S. K., SABATINOS, S., FORSBURG, S. & BASTIA, D. 2010. Regulation of replication termination by Reb1 protein-mediated action at a distance. *Cell,* 142**,** 868-78.

SISTA, P. R., MUKHERJEE, S., PATEL, P., KHATRI, G. S. & BASTIA, D. 1989. A host-encoded DNA-binding protein promotes termination of plasmid replication at a sequence-specific replication terminus. *Proc Natl Acad Sci U S A,* 86**,** 3026-30.

SITARAMAN, K. & CHATTERJEE, D. K. 2011. Protein-protein interactions: an application of Tus-*Ter* mediated protein microarray system. *Methods Mol Biol,* 723**,** 185-200.

SKOKOTAS, A., HIASA, H., MARIANS, K. J., O'DONNELL, L. & HILL, T. M. 1995. Mutations in the *Escherichia coli* Tus protein define a domain positioned close to the DNA in the Tus-*Ter* complex. *J Biol Chem,* 270**,** 30941-8.

SKOKOTAS, A., WROBLESKI, M. & HILL, T. M. 1994. Isolation and characterization of mutants of Tus, the replication arrest protein of *Escherichia coli*. *J Biol Chem,* 269**,** 20446-55.

SMITH, M. T., LANGLEY, D. B., YOUNG, P. A. & WAKE, R. G. 1994. The minimal sequence needed to define a functional DNA terminator in *Bacillus subtilis*. *J Mol Biol,* 241**,** 335-40.

SMITH, M. T. & WAKE, R. G. 1992. Definition and polarity of action of DNA replication terminators in *Bacillus subtilis*. *J Mol Biol,* 227**,** 648-57.

SMITS, W. K., MERRIKH, H., BONILLA, C. Y. & GROSSMAN, A. D. 2011. Primosomal proteins DnaD and DnaB are recruited to chromosomal regions bound by DnaA in *Bacillus subtilis*. *J Bacteriol,* 193**,** 640-8.

SOO, V. W., HANSON-MANFUL, P. & PATRICK, W. M. 2011. Artificial gene amplification reveals an abundance of promiscuous resistance determinants in *Escherichia coli*. *Proc Natl Acad Sci U S A,* 108**,** 1484-9.

SPECK, C., WEIGEL, C. & MESSER, W. 1999. ATP- and ADP-DnaA protein, a molecular switch in gene regulation. *EMBO J,* 18**,** 6169-76.

STANO, N. M., JEONG, Y. J., DONMEZ, I., TUMMALAPALLI, P., LEVIN, M. K. & PATEL, S. S. 2005. DNA synthesis provides the driving force to accelerate DNA unwinding by a helicase. *Nature,* 435**,** 370-3.

STOUT, V. & GOTTESMAN, S. 1990. RcsB and RcsC: a two-component regulator of capsule synthesis in *Escherichia coli*. *J Bacteriol,* 172**,** 659-69.

STRAUME, M. & FREIRE, E. 1992. Two-dimensional differential scanning calorimetry: simultaneous resolution of intrinsic protein structural energetics and ligand binding interactions by global linkage analysis. *Anal Biochem,* 203**,** 259-68.

STUDWELL-VAUGHAN, P. S. & O'DONNELL, M. 1991. Constitution of the twin polymerase of DNA polymerase III holoenzyme. *J Biol Chem,* 266**,** 19833-41.

STUKENBERG, P. T., STUDWELLVAUGHAN, P. S. & ODONNELL, M. 1991. Mechanism of the sliding beta-clamp of DNA polymerase-III holoenzyme. *Journal of Biological Chemistry,* 266**,** 11328-11334.

SUN, W. & GODSON, G. N. 1998. Structure of the *Escherichia coli* primase/single-strand DNA-binding protein/phage G4*oric* complex required for primer RNA synthesis. *J Mol Biol,* 276**,** 689-703.

SUTTON, M. D. & KAGUNI, J. M. 1997. The *Escherichia coli* DnaA gene: four functional domains. *J Mol Biol,* 274**,** 546-61.

TANNER, N. A., HAMDAN, S. M., JERGIC, S., SCHAEFFER, P. M., DIXON, N. E. & VAN OIJEN, A. M. 2008. Single-molecule studies of fork dynamics in *Escherichia coli* DNA replication. *Nature Structural & Molecular Biology,* 15**,** 170-176.

TANNER, N. A., LOPARO, J. J., HAMDAN, S. M., JERGIC, S., DIXON, N. E. & VAN OIJEN, A. M. 2009. Real-time single-molecule observation of rolling-circle DNA replication. *Nucleic Acids Research,* 37.

TANNER, N. A., TOLUN, G., LOPARO, J. J., JERGIC, S., GRIFFITH, J. D., DIXON, N. E. & VAN OIJEN, A. M. 2011. *E. coli* DNA replication in the absence of free beta clamps. *EMBO J,* 30**,** 1830-40.

THORNER, J. W. & PAULUS, H. 1973. Catalytic and allosteric properties of glycerol kinase from *Escherichia coli*. *J Biol Chem,* 248**,** 3922-32.

TORO, E. & SHAPIRO, L. 2010. Bacterial chromosome organization and segregation. *Cold Spring Harb Perspect Biol,* 2**,** a000349.

TOSA, T. & PIZER, L. I. 1971. Biochemical bases for the antimetabolite action of L-serine hydroxamate. *J Bacteriol,* 106**,** 972-82.

TSIEN, R. Y. 1998. The green fluorescent protein. *Annu Rev Biochem,* 67**,** 509-44.

TSUCHIHASHI, Z. & KORNBERG, A. 1990. Translational frameshifting generates the gamma subunit of DNA polymerase III holoenzyme. *Proc Natl Acad Sci U S A,* 87**,** 2516-20.

TURNER, J., HINGORANI, M. M., KELMAN, Z. & O'DONNELL, M. 1999. The internal workings of a DNA polymerase clamp-loading machine. *EMBO J,* 18**,** 771-83.

VALENS, M., PENAUD, S., ROSSIGNOL, M., CORNET, F. & BOCCARD, F. 2004. Macrodomain organization of the *Escherichia coli* chromosome. *EMBO J,* 23**,** 4330-41.

VALJAVEC-GRATIAN, M., HENDERSON, T. A. & HILL, T. M. 2005. Tus-mediated arrest of DNA replication in *Escherichia coli* is modulated by DNA supercoiling. *Mol Microbiol,* 58**,** 758-73.

VEDADI, M., NIESEN, F. H., ALLALI-HASSANI, A., FEDOROV, O. Y., FINERTY, P. J., JR., WASNEY, G. A., YEUNG, R., ARROWSMITH, C., BALL, L. J., BERGLUND, H., HUI, R., MARSDEN, B. D., NORDLUND, P., SUNDSTROM, M., WEIGELT, J. & EDWARDS, A. M. 2006. Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proc Natl Acad Sci U S A,* 103**,** 15835-40.

VESSONI PENNA, T. C., ISHII, M., CHOLEWA, O. & DE SOUZA, L. C. 2004. Thermal characteristics of recombinant green fluorescent protein (GFPuv) extracted from *Escherichia coli*. *Lett Appl Microbiol,* 38**,** 135-9.

VIVIAN, J. P., PORTER, C. J., WILCE, J. A. & WILCE, M. C. 2007. An asymmetric structure of the *Bacillus subtilis* replication terminator protein in complex with DNA. *J Mol Biol,* 370**,** 481-91.

VOLKMER, B. & HEINEMANN, M. 2011. Condition-dependent cell volume and concentration of *Escherichia coli* to facilitate data conversion for systems biology modeling. *PLoS One,* 6**,** e23126.

WAHLE, E., LASKEN, R. S. & KORNBERG, A. 1989. The DnaB-DnaC replication protein complex of *Escherichia coli*. I. Formation and properties. *J Biol Chem,* 264**,** 2463-8.

WAKE, R. & KING, G. 1997. A tale of two terminators: crystal structures sharpen the debate on DNA replication fork arrest mechanisms. *Structure,* 5**,** 1-5.

WALDO, G. S. 2003. Genetic screens and directed evolution for protein solubility. *Curr Opin Chem Biol,* 7**,** 33-8.

WALDO, G. S., STANDISH, B. M., BERENDZEN, J. & TERWILLIGER, T. C. 1999. Rapid protein-folding assay using green fluorescent protein. *Nat Biotechnol,* 17**,** 691-5.

WALDRON, T. T., SCHRIFT, G. L. & MURPHY, K. P. 2005. The salt-dependence of a protein-ligand interaction: ion-protein binding energetics. *J Mol Biol,* 346**,** 895-905.

WANG, X., LESTERLIN, C., REYES-LAMOTHE, R., BALL, G. & SHERRATT, D. J. 2011. Replication and segregation of an *Escherichia coli* chromosome with two replication origins. *Proc Natl Acad Sci U S A,* 108**,** E243-50.

WARD, W. W. & BOKMAN, S. H. 1982. Reversible denaturation of *Aequorea* green-fluorescent protein: physical separation and characterization of the renatured protein. *Biochemistry,* 21**,** 4535-40.

WEIGEL, C., SCHMIDT, A., RUCKERT, B., LURZ, R. & MESSER, W. 1997. DnaA protein binding to individual DnaA boxes in the *Escherichia coli* replication origin, oriC. *EMBO J,* 16**,** 6574-83.

WEIGEL, C., SCHMIDT, A., SEITZ, H., TUNGLER, D., WELZECK, M. & MESSER, W. 1999. The N-terminus promotes oligomerization of the *Escherichia coli* initiator protein DnaA. *Mol Microbiol,* 34**,** 53-66.

WEISS, A. S. & WAKE, R. G. 1984. Impediment to replication fork movement in the terminus region of the *Bacillus subtilis* chromosome. *J Mol Biol,* 179**,** 745-50.

WHITE, M. A., EYKELENBOOM, J. K., LOPEZ-VERNAZA, M. A., WILSON, E. & LEACH, D. R. 2008. Non-random segregation of sister chromosomes in *Escherichia coli*. *Nature,* 455**,** 1248-50.

WICKNER, S. & HURWITZ, J. 1975. Interaction of *Escherichia coli dnaB* and *dnaC(D)* gene products *in vitro*. *Proc Natl Acad Sci U S A,* 72**,** 921-5.

WING, R. A., BAILEY, S. & STEITZ, T. A. 2008. Insights into the replisome from the structure of a ternary complex of the DNA polymerase III alpha-subunit. *J Mol Biol,* 382**,** 859-69.

WITTE, G., URBANKE, C. & CURTH, U. 2003. DNA polymerase III chi subunit ties single-stranded DNA binding protein to the bacterial replication machinery. *Nucleic Acids Res,* 31**,** 4434-40.

WU, C. A., ZECHNER, E. L. & MARIANS, K. J. 1992. Coordinated leading- and lagging-strand synthesis at the *Escherichia coli* DNA replication fork. I. Multiple effectors act to modulate Okazaki fragment size. *J Biol Chem,* 267**,** 4030-44.

XIAO, H., DONG, Z. & O'DONNELL, M. 1993. DNA polymerase III accessory proteins. IV. Characterization of chi and psi. *J Biol Chem,* 268**,** 11779-84.

YAKUBOVSKAYA, E., MEJIA, E., BYRNES, J., HAMBARDJIEVA, E. & GARCIA-DIAZ, M. 2010. Helix unwinding and base flipping enable human MTERF1 to terminate mitochondrial transcription. *Cell,* 141**,** 982-93.

YAMAMOTO, K., HIRAO, K., OSHIMA, T., AIBA, H., UTSUMI, R. & ISHIHAMA, A. 2005. Functional characterization *in vitro* of all two-component signal transduction systems from *Escherichia coli*. *J Biol Chem,* 280**,** 1448-56.

YANG, M. Y., BOWMAKER, M., REYES, A., VERGANI, L., ANGELI, P., GRINGERI, E., JACOBS, H. T. & HOLT, I. J. 2002. Biased incorporation of ribonucleotides on the mitochondrial L-strand accounts for apparent strand-asymmetric DNA replication. *Cell,* 111**,** 495-505.

YAO, N. Y. & O'DONNELL, M. 2008. Replisome dynamics and use of DNA trombone loops to bypass replication blocks. *Mol Biosyst,* 4**,** 1075-84.

YEELES, J. T. & MARIANS, K. J. 2011. The *Escherichia coli* replisome is inherently DNA damage tolerant. *Science,* 334**,** 235-8.

YODA, K., YASUDA, H., JIANG, X. W. & OKAZAKI, T. 1988. RNA-primed initiation sites of DNA replication in the origin region of bacteriophage lambda genome. *Nucleic Acids Res,* 16**,** 6531-46.

YUZHAKOV, A., KELMAN, Z. & O'DONNELL, M. 1999. Trading places on DNA-a three-point switch underlies primer handoff from primase to the replicative DNA polymerase. *Cell,* 96**,** 153-63.

ZAKRZEWSKA-CZERWINSKA, J., JAKIMOWICZ, D., ZAWILAK-PAWLIK, A. & MESSER, W. 2007. Regulation of the initiation of chromosomal replication in bacteria. *FEMS Microbiol Rev,* 31**,** 378-87.

ZHOU, P., BOGAN, J. A., WELCH, K., PICKETT, S. R., WANG, H. J., ZARITSKY, A. & HELMSTETTER, C. E. 1997. Gene transcription and chromosome replication in *Escherichia coli*. *J Bacteriol,* 179**,** 163-9.

ZUBRIENE, A., MATULIENE, J., BARANAUSKIENE, L., JACHNO, J., TORRESAN, J., MICHAILOVIENE, V., CIMMPERMAN, P. & MATULIS, D. 2009. Measurement of nanomolar dissociation constants by titration calorimetry and thermal shift assay - radicicol binding to Hsp90 and ethoxzolamide binding to CAII. *Int J Mol Sci,* 10**,** 2662-80.

Appendix

Appendix A
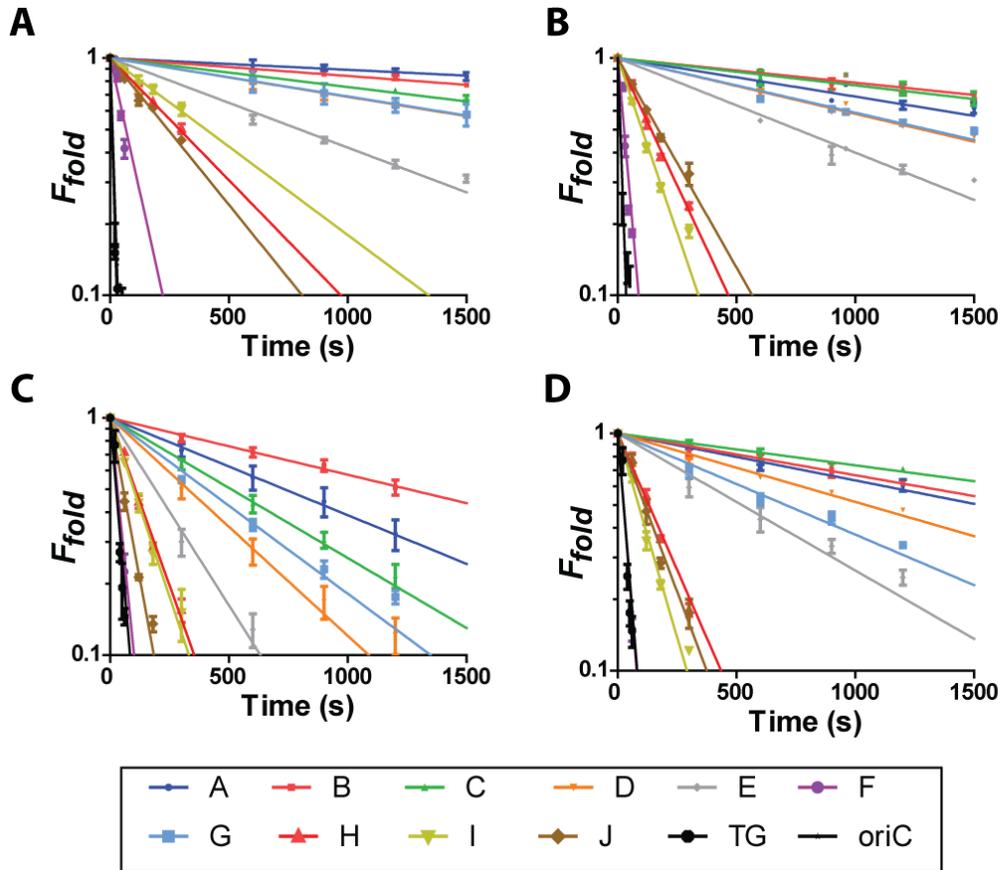
Figure S1: Tus-GFP aggregation rates in presence of *Ter*, *Ter-lock* or *oriC* (non-specific control) sequences obtained with GFP-Basta. (A) Aggregation rates of Tus-GFP-*Ter* complexes in 150 mM KCl at 58°C. (B) Aggregation rates of Tus-GFP-*Ter-lock* complexes in 150 mM KCl at 58°C. (C) Aggregation rates of Tus-GFP-*Ter* complexes in 250 mM KCl at 52°C. (D) Aggregation rates of Tus-GFP-*Ter-lock* complexes in 250 mM KCl at 52 °C.
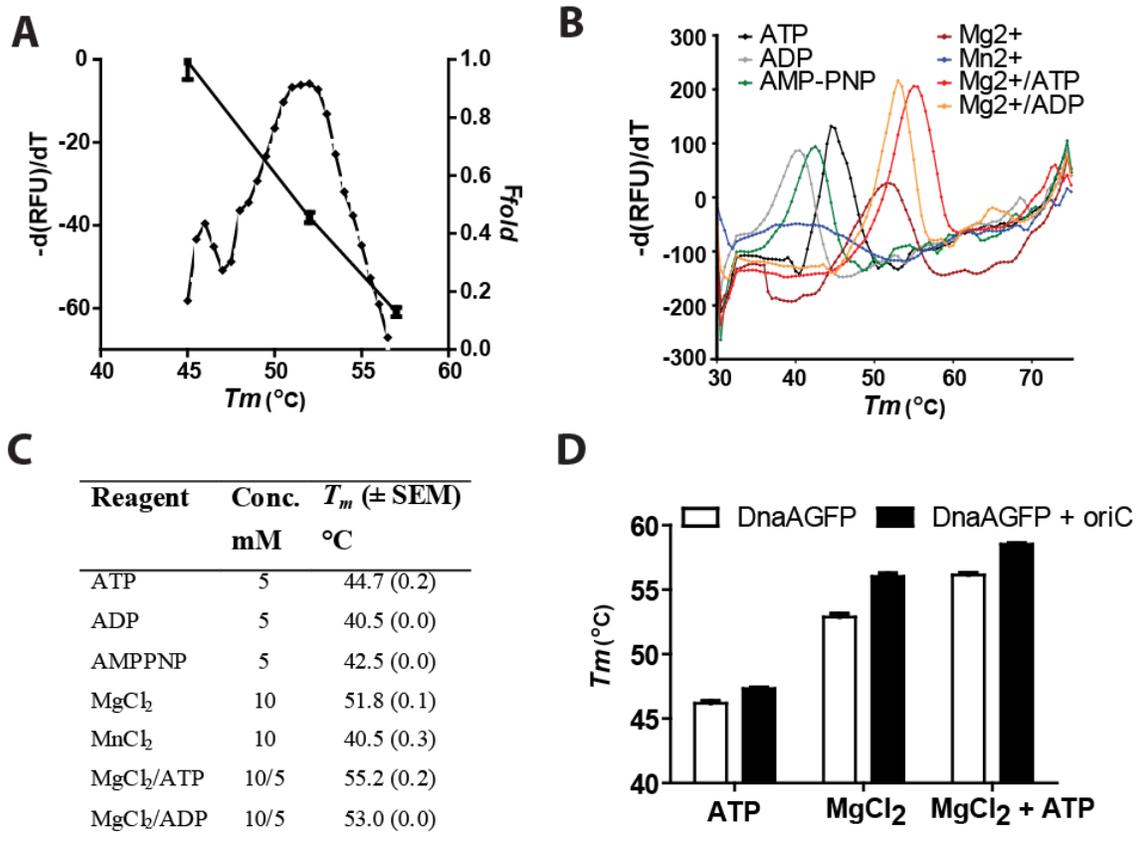
Figure S2: Thermal stability of DnaA-GFP. (A) Correlation between the T_m peak and the aggregation state of DnaA-GFP in the presence of 10 mM $MgCl_2$. Melting curves were stopped immediately before, at the midpoint, and at the end of the T_m peak and centrifuged to discard aggregates (F_{fold}; squares). The melting curve of DnaA-GFP expressed as -dRFU/dT is represented for comparison (diamonds). **(B)** Effects of nucleotides and divalent cations on DnaA-GFP. **(C)** T_m values of DnaA-GFP in the presence of nucleotides and divalent cations. **(D)** Effect of DNA binding in the presence of ATP, $MgCl_2$ or ATP/$MgCl_2$. Reactions contained 60 µl of DnaA-GFP at 2.5 mM. Initial RFUs were 3800–4700. Melting curves were recorded from 30–75°C with 0.5°C/cycle and 30 s dwell time.

Appendix C

List of current publications

DAHDAH, D. B., MORIN, I., **MOREAU, M. J.**, DIXON, N. E. & SCHAEFFER, P. M. 2009. Site-specific covalent attachment of DNA to proteins using a photoactivatable Tus-*Ter* complex. *Chem Commun (Camb)*, 3050-2.

MOREAU, M. J., MORIN, I. & SCHAEFFER, P. M. 2010. Quantitative determination of protein stability and ligand binding using a green fluorescent protein reporter system. *Mol Biosyst,* 6**,** 1285-92.

MOREAU, M. J. & SCHAEFFER, P. M. 2012a. Differential Tus-*Ter* binding and lock formation: implications for DNA replication termination in *Escherichia coli*. *Mol Biosyst,* 8**,** 2783-91.

MOREAU, M. J. & SCHAEFFER, P. M. 2012b. A polyplex qPCR-based binding assay for protein-DNA interactions. *Analyst,* 137**,** 4111-3.

MOREAU, M. J.., MORIN, I., ASKIN, S. P., COOPER, A., MORELAND, N. J., VASUDEVAN, S. G. & SCHAEFFER, P. M. 2012. Rapid determination of protein stability and ligand binding by differential scanning fluorimetry of GFP-tagged proteins. *RSC Advances,* 2**,** 11892-11900.