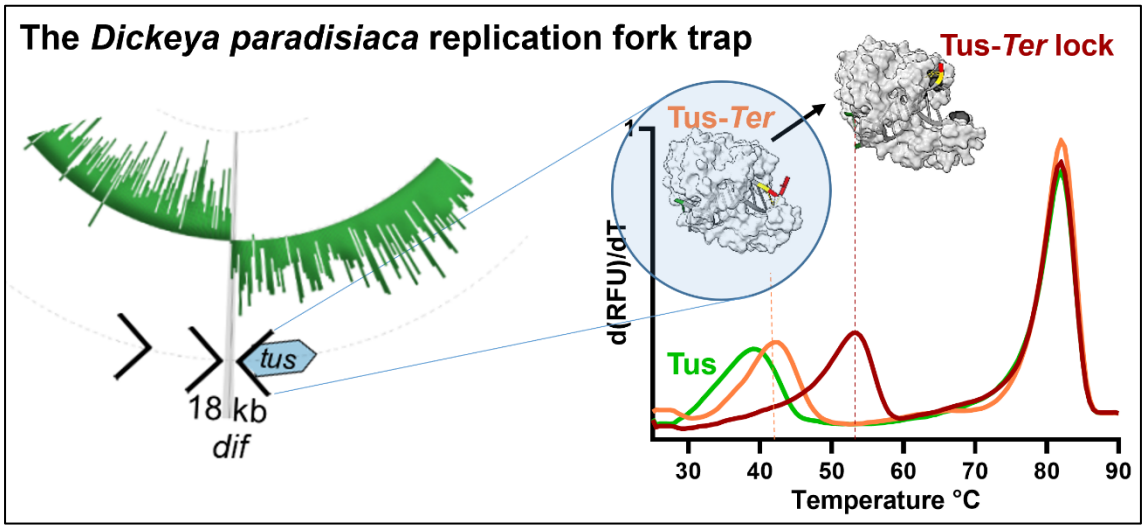
A soft Tus-*Ter* interaction is hiding a fail-safe lock in the replication fork trap of *Dickeya paradisiaca*

Casey J. Toft, Alanna E. Sorenson and Patrick M. Schaeffer*

Molecular and Cell Biology, College of Public Health, Medical and Veterinary Sciences, James Cook University, Douglas, QLD, 4811, Australia

* To whom correspondence should be addressed. Tel: +61 (0) 7 4781 4448; Fax: +61 (0) 7 4781 6078; Email: patrick.schaeffer@jcu.edu.au



Abstract

A variety of replication fork traps have recently been characterised in Enterobacterales, unveiling two different types of architecture. Of these, the degenerate type II fork traps are commonly found in *Enterobacteriaceae* such as *Escherichia coli*. The newly characterised type I fork traps are found almost exclusively outside *Enterobacteriaceae* within Enterobacterales and include several archetypes of possible ancestral architectures. *Dickeya paradisiaca* harbours a somewhat degenerate type I fork trap with a unique *Ter1* adjacent to *tus* gene on one side of the circular chromosome and three putative *Ter2-4* sites on the other side of the fork trap. The two innermost *Ter1* and *Ter2* sites are only separated by 18 kb, which is the shortest distance between two innermost *Ter* sites of any chromosomal fork trap identified so far. Of note, the *dif* site is located between these two sites, coinciding with a sharp GC-skew flip. Here we examined and compared the binding modalities of *E. coli* and *D. paradisiaca* Tus proteins for these *Ter* sites. Surprisingly, while *Ter1-3* were functional, no significant Tus binding was observed for *Ter4* even in low salt conditions, which is in stark contrast with the significant non-specific protein-DNA interactions that occur with *E. coli* Tus. Even more surprising was the finding that *D. paradisiaca* Tus has a relatively moderate binding affinity to double-stranded *Ter* while retaining an extremely high affinity to *Ter*-lock sequences. Our data revealed major differences in the salt resistance and stability between the *D. paradisiaca* and *E. coli* Tus protein complexes, suggesting that while Tus protein evolution can be quite flexible regarding the initial *Ter* binding step, it requires a highly stringent purifying selection for its final locked complex formation.

Keywords

DNA replication termination, Enterobacterales, Tus-*Ter*, Fork trap, DSF-GTP, EMSA

1. Introduction

Most of our understanding of DNA replication termination originates from the biochemical characterisation of the replisomal proteins of *Bacillus subtilis* and *Escherichia coli* (Neylon et al., 2005; Schaeffer et al., 2005). Both model bacteria carry a replication fork trap system situated in the terminus region ensuring that replication forks can enter but not leave, thus preventing over-replication (Midgley-Smith et al., 2018). Initially, the *E. coli* replication fork trap system was assumed to be conserved across all bacteria (Louarn et al., 1977; Hill et al., 1988). It was later found that the *E. coli* Tus-*Ter* system did not share any sequence or structural homology nor a similar binding mechanism with the *B. subtilis* RTP-*Ter* system (Bussiere et al., 1995; Wilce et al., 2001; Neylon et al., 2005). The RTP protein is only conserved in a subset of *Bacillus* species (Griffiths et al., 1998) and evolved independently of Tus. The notion that a replication fork trap is essential, was later dismissed by the observation that deletion of chromosomal *tus* in *E. coli* has little effect on overall cell viability (Roecklein et al., 1991; Gottlieb et al., 1992; Sharma and Hill, 1995). The rapid growth of genomic data supported this notion, *e.g. Vibrio cholerae*, a bacterium closely related to *E. coli,* is devoid of any replication fork trap system (Galli et al., 2019). Nevertheless, the *tus* gene is commonly found across Enterobacterales (Galli et al., 2019; Toft et al., 2021).

Since its discovery (Hill et al., 1989; Kuempel et al., 1989), the *E. coli* Tus-*Ter* interaction has been intensely scrutinized, culminating into the elucidation of the Tus-*Ter* lock (TT-lock) mechanism of polar fork arrest (Mulcair et al., 2006), and more recently fuelling the development of new technological applications (Dahdah et al., 2009; Morin et al., 2010; Askin et al., 2011; Askin and Schaeffer, 2012; Moreau and Schaeffer, 2012b; Cooper et al., 2013; Johnston et al., 2014; Kamath et al., 2014; Willis et al., 2014; Willis et al., 2018; Jorgensen et al., 2019; Norouzi et al., 2021). The complex replication fork trap architecture in terms of

spatial distribution and number of *Ter* sites within the *E. coli* chromosome has recently been revised down to only six functional fork-arresting sites (Figure 1)(Toft et al., 2021). An examination of the fork traps of other Tus-dependent bacteria, revealed that the large number of *Ter* sites in *E. coli* are more of an exception (Figure 1). The study also revealed two distinct replication fork trap architectures (Toft et al., 2021): type I where the innermost *Ter,* acting as a repressor of the *tus* gene, is the primary fork arrest site (*e.g. Edwardsiella tarda*); and type II where the *Ter* vicinal to *tus* is less-frequently used as it is in the second position from the terminus (*e.g. E. coli*). Of importance, the newly identified type I fork trap architecture is almost exclusively found in bacteria outside of the *Enterobacteriaceae* family and for many of these species, such as *E. tarda,* it contains only two closely positioned *Ter* sequences in opposite orientation, near the *dif* site (Figure 1). Interestingly, the amino acid residues of *E. coli* Tus ($_{Ec}$Tus) that make critical interactions with the C(6) base of *Ter* and are part of the cytosine binding pocket, are mostly conserved in *E. tarda* Tus ($_{Et}$Tus) despite having a moderate sequence identity with the *E. coli* protein (48%) (Figure 2A and B). Toft et al. (2021) hypothesised that in bacteria containing only two *Ter* sites to sequester replication forks, the Tus protein must be able to form a fail-safe TT-lock complex (Toft et al., 2021). The sharp terminus GC-skew switch observed for these species strongly suggests that replication forks never break through these *Ter* sites (Figure 1) whether due to fork arrest or collisions. Indeed, Kono et al. (2012) showed that for 65 Tus-harbouring bacteria, the sharp GC-skew pattern was best recreated by a model in which replication forks were arrested by random collision (10% of the time), the Tus-*Ter* fork trap (70% of the time) and at *dif* (20% of the time) (Kono et al., 2012). Moreover, Rudolph et al. (2013) found that a sharp focusing of termination that is evident in wild-type *E. coli* is identical in a *tus* deletion mutant, suggesting that additional factors may work in concert with Tus and *Ter* (Rudolph et al., 2013). One of the most intriguing type I fork traps, with respect to its wide span and *Ter* sites as well as its *dif* and GC-skew

84 switch loci, was identified in *Cedecea* species (Figure 1) and suggests that replication stalling
85 activity may not be a primary function of Tus here (Toft et al., 2021). These recent findings
86 raise questions about the *Ter*-binding and lock forming ability of orthologous Tus proteins in
87 type I replication fork traps.
88

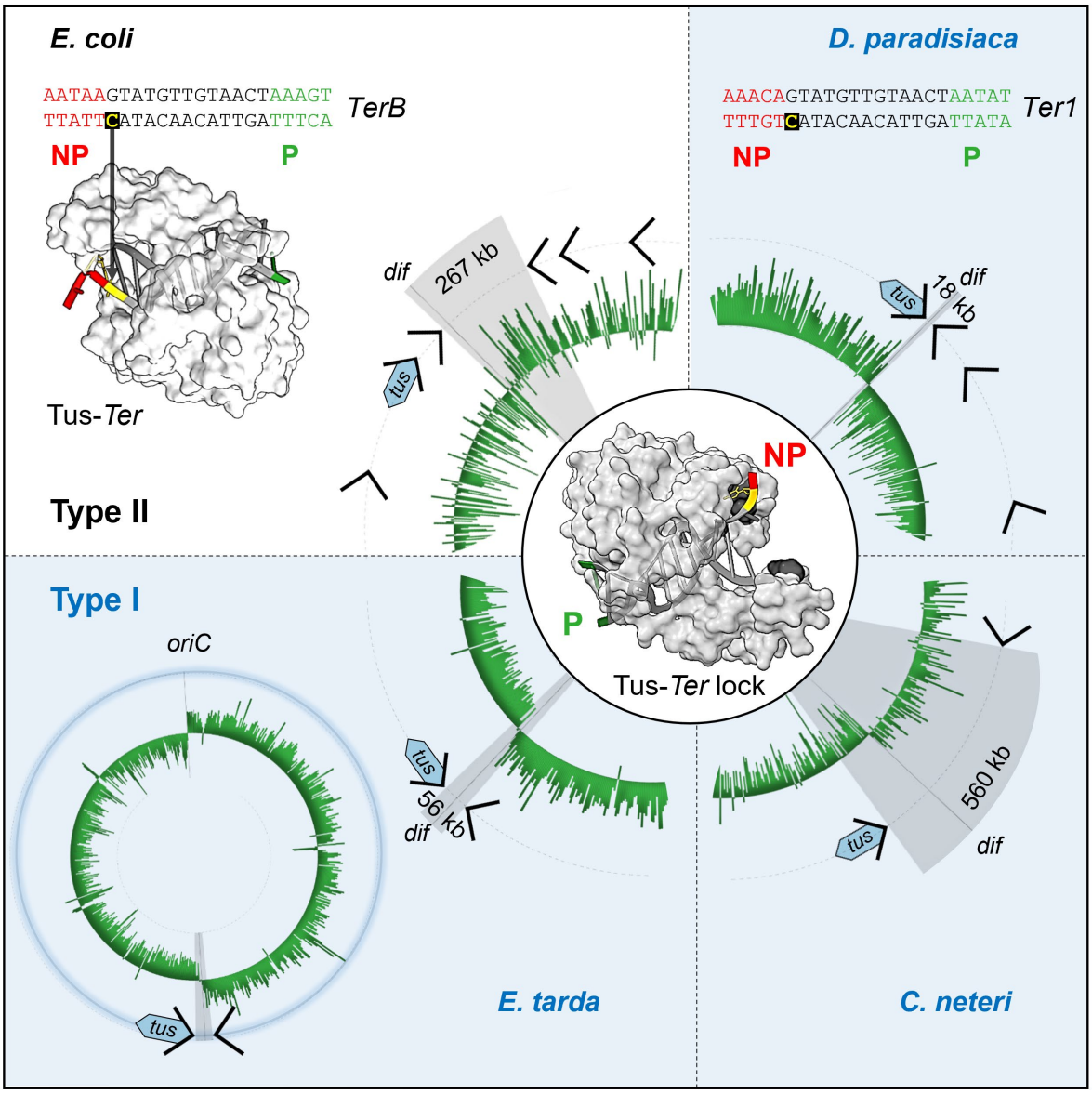
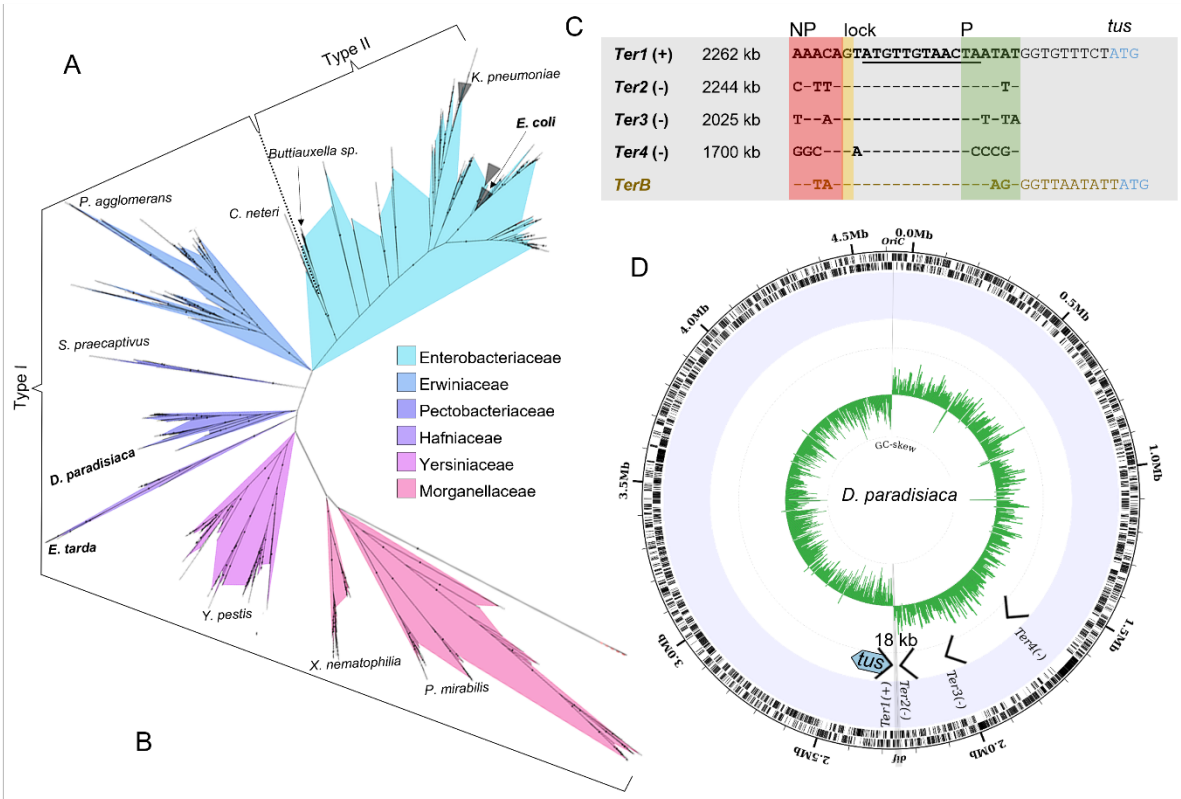


Figure 1: Replication fork trap architectures and *Ter* site loci in *Escherichia coli*, *Dickeya paradisiaca*, *Cedecea neteri* and *Edwardsiella tarda*. The type I replication fork traps are shaded in light blue. The area shaded in grey represents the distance between the innermost *Ter* sites. The *tus* gene position is indicated (blue). GC-skew (green) and *dif* site are indicated for reference. Structures of Tus-*Ter* (PDB 2I05) and locked complexes (PDB 2I06) with the C(6) of *Ter* and Ter-lock DNA highlighted in yellow and their permissive (P) and non-permissive (NP) ends in green and red respectively.

Some orthologous Tus proteins of *Salmonella typhimurium* and *enteritidis* (type II), *Klebsiella ozaenae* (type II), *Yersinia pestis* and *enterocolitica* (type I), *Proteus mirabilis* (type I) and *Serratia marcescens* (type I) have only been examined briefly (Henderson et al., 2001). Two main conclusions emanated from the study: firstly, a selection of these Tus proteins was shown to bind to the *E. coli TerB* sequence using bacterial cell extracts in a band shift assay; secondly, four of these Tus proteins (*S. typhimurium, S. enteritidis, K. ozaenae and Y. pestis*) could be expressed in *E. coli* and were capable of arresting replication forks within a mini-chromosome carrying two oppositely-oriented *E. coli TerB* sites with their non-permissive end pointing toward the *oriC*. A somewhat lower fork arrest efficiency was observed for *Y. pestis* Tus (53% sequence identity between *E. coli* and *Y. pestis* Tus) at *TerB* (Henderson et al., 2001). All other Tus proteins were found to be slightly more efficient than the *E. coli* protein. Of note, only two *Ter* sites are present within a narrow type I replication fork trap in *Y. pestis* (Toft et al., 2021), suggesting that *Y. pestis* Tus may have a higher binding affinity and form a more efficient lock with its specific *Ter* sequences. Indeed, the vicinal *Ter1* sequence in *Y. pestis* is one of the most distantly related to *E. coli TerB*, with only 17 conserved bases out of 23 (Toft et al., 2021). Unfortunately, none of these orthologous Tus proteins have been examined in any further detail for their ability to form a TT-lock complex.

D. paradisiaca (*Pectobacteriaceae*) is a plant pathogen, recently proposed for reclassification into a new genus on the basis of genomic, phylogenetic and phenotypic analysis (Hugouvieux-Cotte-Pattat et al., 2021). Only two genomes have been sequenced to date, with available literature limited to taxonomic studies and biochemical analysis for phenotyping purposes (Samson et al., 2005; Salplachta et al., 2015; Hugouvieux-Cotte-Pattat et al., 2021). *D. paradisiaca* harbours a somewhat degenerate type I fork trap with a unique *Ter1* adjacent to the *tus* gene on one side of the circular chromosome, and three putative *Ter2-4* sites on the other side of the fork trap (Figure 1 and Figure 2). The two innermost *Ter1* and *Ter2* sites are

only separated by 18 kb. This is the shortest distance between two innermost *Ter* sites of any chromosomal fork trap identified so far. Of note, the *dif* site is located right between these two sites also coinciding with a sharp GC-skew flip (Toft et al., 2021). Here we examined the binding properties of *D. paradisiaca* Tus ($_{Dp}$Tus) for these *Ter* sites to gain some insight into the possible precursor and mechanism leading to such a fork trap architecture, *i.e.* whether the bacterium initially adopted a wide or a narrow fork trap. A systematic approach including a green fluorescent protein-based electrophoretic mobility shift assay (GFP-EMSA) (Morin et al., 2012; Askin et al., 2016; Sorenson and Schaeffer, 2020a) and differential scanning fluorimetry of GFP-tagged Proteins (DSF-GTP) (Moreau et al., 2012; Moreau and Schaeffer, 2013; Sorenson and Schaeffer, 2020b) was used to characterise and compare the binding of $_{Dp}$Tus and $_{Ec}$Tus to *Ter* and *Ter*-lock DNA species. Unexpectedly, no significant $_{Dp}$Tus binding occurred with the outermost *Ter4*. Most surprising was the finding that $_{Dp}$Tus exhibits moderate affinity for double-stranded *Ter* while retaining extremely high affinity for *Ter*-lock species even in low salt conditions, which is completely at odds with the *E. coli* data. Overall, our data may suggest that *D. paradisiaca* initially adopted a wide fork trap that subsequently narrowed through acquisition of additional *Ter* sites.



Species	*Ter* sites	Fork trap size (kb)	Vicinal Ter sequence	Identity to *TerB*	% Identity to $_{Ec}$Tus	Fork-trap classification
E. coli	6*	267	AATAAGTATGTTGTAACTAAAGT	-	-	Type II
D. paradisiaca	**4**	**18**	**AAACAGTATGTTGTAACTAATAT**	**19/23**	**53.9**	**Type I**
E. tarda	2	58	CAAAAGTATGTTGTAACTAATAA	18/23	48.3	Type I

52

Figure 2: Phylogenetic analysis of Tus in Enterobacterales and the chromosomal fork trap architecture in *D. paradisiaca*. (A) Phylogenetic relationship of ~2500 Tus protein sequences using InterPro entries (IPR008865) highlighting the evolutionary distance of *D. paradisiaca* from *E. coli* and *E. tarda* (B) Chromosomal fork trap characteristics and classification of *D. paradisiaca* compared to *E. coli* and *E. tarda*. Fork trap size (kb) corresponds to the distance between the two innermost *Ter* sites of opposite polarity. Underlined bases represent a continuous identical sequence shared between *Ter* sequences vicinal to *Tus*. (C) Sequence alignment of the *D. paradisiaca Ter* and *E. coli TerB* sequences. *Ter1* and *TerB* are located upstream the start codon of *tus.* The conserved 12 bp core *Ter* sequence is underlined and the G(6) base complementary to C(6) is highlighted in yellow. NP: non-permissive end (red), P: permissive end (green). (D) Circular representation of *D.*

paradisiaca chromosome. Illustrated from the outside to the centre of the circle: forward and reverse genes, labelled genomic locations of putative *Ter* sites, simplified annotation of the termination fork trap utilised, GC-skew over a 5000 bp moving window. The sharp GC-skew switches polarity at the replication origin and between the *Ter1* (locus: 2262 kb) and *Ter2* (locus: 2244 kb) sites near *dif* (locus: 2251 kb). Adapted from Toft et al. (Toft et al., 2021)

2. Materials and methods

2.1. Expression and purification of TusGFP proteins

$_{Ec}$TusGFP was expressed and Ni-affinity purified as previously described (Moreau et al., 2012). The $_{Dp}$Tus coding sequence (Dd703_1900) was synthesised in a codon optimized form (Bioneer) and ligated into the GFP vector pIM013 (Moreau et al., 2010) to create pCT300 (pET-N-6His-$_{Dp}$TusGFP-C). The pCT300 expression vector carrying the coding sequence of $_{Dp}$TusGFP was expressed in *E. coli* BL21(DE3)RIPL using Overnight Express TB Medium, containing 100 µg/ml ampicillin and 50 µg/ml chloramphenicol. In a 1 L flask, 100 mL of TB expression medium was inoculated with a bacterial loop sourced from a fresh overnight master plate culture and incubated at 37°C and 200 RPM until the optical density reached 0.7. Proteins were expressed over 72 hours at 16°C. Lysis and purification procedures were performed as for $_{Ec}$TusGFP (Moreau et al., 2012). Protein concentrations were determined by Bradford Assay and purity assessed by SDS-PAGE. Purified TusGFP samples were stored in buffer A (50 mM sodium phosphate (pH 7.8), 10% glycerol (v/v) and 2 mM β-mercaptoethanol).

2.2. Size exclusion chromatography

Purified $_{Dp}$TusGFP, $_{Ec}$TusGFP and GFP in buffer A were loaded (*i.e.* 200 µL of 100 µM stocks) individually on a Superdex 200 10/300 GL column connected to a BioLogic Duo-Flow system and BioLogic QuadTec UV-Vis detector (Bio-Rad). Size exclusion chromatography (SEC) was performed at 4°C, in buffer B (buffer A + 125 mM NaCl and without β-mercaptoethanol)

and a flow rate of 0.5 mL/min. Absorbance was recorded at 280 nm using BioLogic QuadTec UV-Vis detector (Bio-Rad) and plotted as a function of the retention time. GFP Fluorescence in each fraction (0.5 mL) of $_{Dp}$TusGFP was measured in the Bio-Rad CFX96 C1000 Touch Thermal Cycler (FAM channel) and superimposed onto the SEC elution profile obtained at 280 nm absorbance. Purity of fractions containing purified proteins were evaluated by SDS-PAGE and Coomassie Blue staining.

2.3. GFP-EMSA

The binding of *Ter* species to TusGFP proteins was examined using a modified electrophoretic mobility shift assay (GFP-EMSA) (Askin et al., 2016; Sorenson and Schaeffer, 2020a). Briefly, 3 µL of a working stock of $_{Dp}$TusGFP or $_{Ec}$TusGFP (4 µM) in buffer A were mixed with 3 µL of *Ter* or *Ter*-lock DNA (4.5 µM) in buffer C (20 mM Tris-HCl (pH 8), 150 mM NaCl) and 6 µL of either $_{dd}H_2O$ (yielding 37.5 mM final NaCl) or a 425 mM NaCl solution (yielding 250 mM final NaCl). Reactions were left at RT for 10 min and loaded onto a 1% agarose gel in TBE and run at 80 V for 40 min. Protein and DNA bands were visualised using a G:BOX Chemi XRQ. GFP fluorescence was first captured (Blue LED module, Filt525), followed by Gelred staining (in 3XGelred solution for 30 min) and DNA fluorescence (TLUM mid-wave, UV06).

2.4. DSF-GTP

All DSF-GTP (Moreau et al., 2012) reactions were run in Hard-Shell 96-Well PCR Plates (Bio-Rad) sealed with Microseal B Adhesive sealer (Bio-Rad). Briefly, 50 µl reactions were equilibrated for 10 min at RT prior to melt curve analysis in a real-time thermal cycler (Bio-Rad CFX96 C1000 Touch Thermal Cycler). Temperature range was set from 25 to 90°C, increasing in 0.5°C increments every 30 s and a stabilization phase of 30 s between increments.

GFP fluorescence was recorded using the FAM channel. Transition midpoint temperature (T_m) data were analysed in GraphPad Prism (version 8.3.1).

2.5. Thermal stability of TusGFP-*Ter* and *Ter-lock* complexes

The T_m of free TusGFP and TusGFP-*Ter* complexes were analysed at 37.5 mM, 144 mM, 250 mM and 400 mM of NaCl by DSF-GTP. For this, 12.5 µL of $_{Dp}$TusGFP or $_{Ec}$TusGFP (4 µM in buffer A) were mixed with 12.5 µL of *Ter* or *Ter*-lock DNA (4.5 µM in buffer C) and 25 µL of water or NaCl solutions (212.5, 425 or 725 mM). Reactions were left at RT for 10 min and subject to the DSF-GTP protocol described above. All reactions were performed at least in triplicate.

Additionally, the T_m of the TusGFP proteins in the presence of increasing *Ter1* or *Ter1*-lock species (ranging from 1 - 10 µM final concentration) was analysed at 37.5 mM and 144 mM of NaCl by DSF-GTP to estimate their apparent dissociation constant (K_{obs}). Here, reactions were performed in duplicate.

2.6. Mathematical and statistical analyses

Statistics and number of biological and technical repeats are indicated in the relevant figure legends and methods. Statistical analyses were performed using GraphPad Prism 7. Data are expressed as mean values ± SD. To determine if TusGFP proteins are affected equally by the ionic strength of the buffer, a simple linear regression analysis was performed for the $_{Dp}$ and $_{Ec}$ TusGFP slopes.

3. Results and discussion

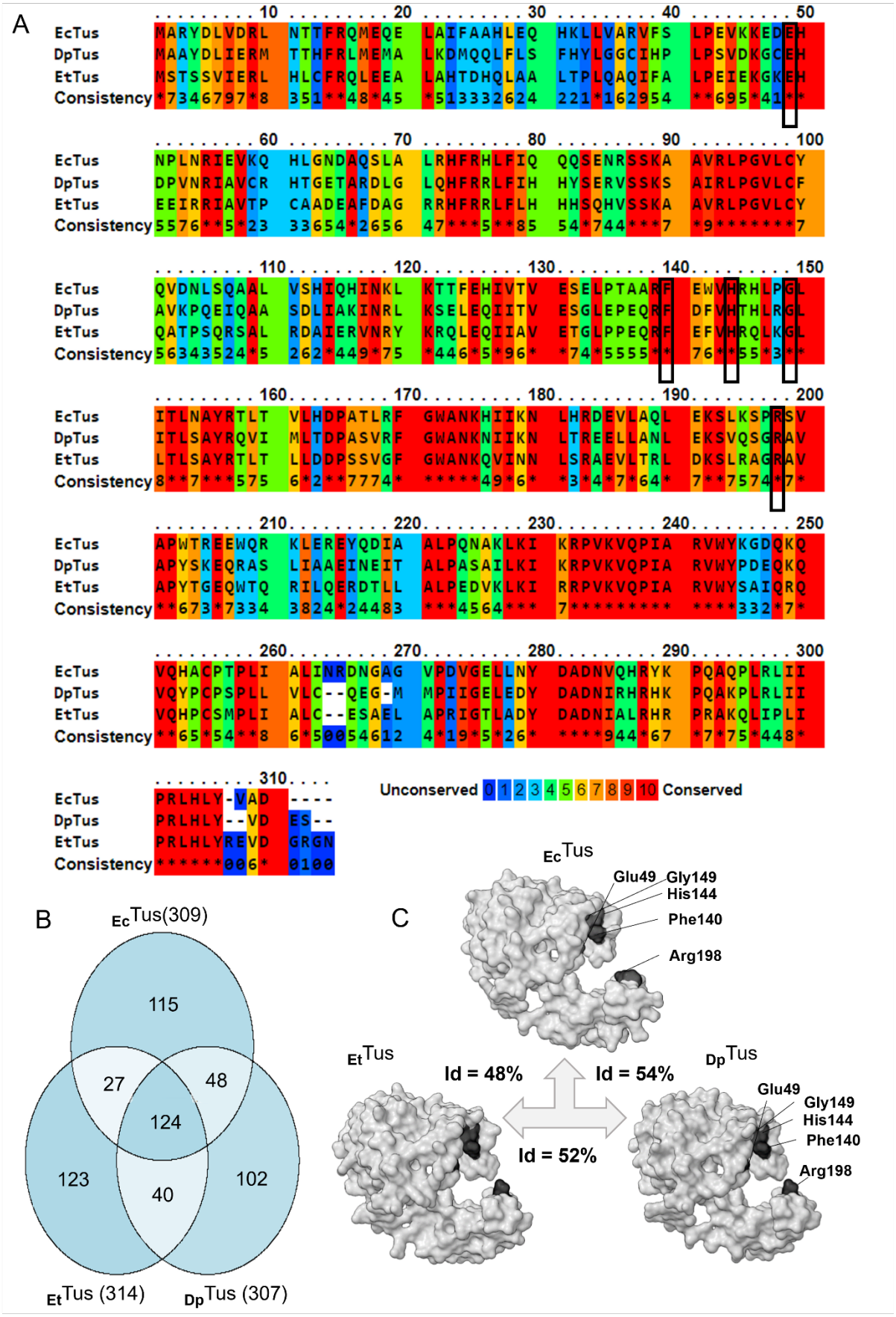
3.1. *In silico* analyses and Rationale

Recently, four putative *Ter* sites were identified in the *D. paradisiaca* replication fork trap (type I, 'slightly degenerate', Figure 2C) (Toft et al., 2021). Here, the left chromosomal arm contains a single *Ter* site vicinal to *tus*, while the right chromosomal arm contains three putative *Ter* sites in opposite orientation (Figure 2D). Of note, the distance between the innermost *Ter* sites is just ~18 kb (Figure 2D) which is in extreme contrast with the wide *E. coli* fork trap. Moreover, an extremely sharp terminal GC-skew switch coincides with the *dif* site midway between the vicinal *Ter1* and *Ter2* sites (Figure 2D), and is located halfway from the origin. Here, delineation of the leading mechanism of termination would be challenging. Yet, it is tempting to hypothesize that the innermost *Ter* sites must be extremely efficient at arresting replication forks, which can be examined by comparing their locking strength. Additionally, scrutinizing the binding of $_{Dp}$Tus to the three putative *Ter* sequences of the right chromosomal arm may provide valuable insight into a possible need for backup *Ter* sites (*e.g.* are additional *Ter* sites required due to a lower Tus-*Ter* complex stability?), and the mechanisms of adoption and evolution of the replication fork trap architecture in *D. paradisiaca*.

The formation of a TT-lock complex is crucial to arrest DNA replication in *E. coli*. Several residues (E49, I79, F140, H144, G149 and R198) found at the non-permissive face of $_{Ec}$Tus have been shown to be important for *Ter* binding (Kamada et al., 1996) and efficient TT-lock formation to arrest incoming DNA replication forks (Mulcair et al., 2006; Elshenawy et al., 2015). Most of these are fully conserved amongst Tus-harbouring bacteria (Toft et al., 2021). $_{Dp}$Tus shares ~54% sequence identity with $_{Ec}$Tus and 52% with $_{Et}$Tus (Figure 3A). A triangular comparison reveals that only 124 amino acid residues are identical between these proteins which are mostly clustered in highly conserved uninterrupted sequences (Figure 3A-B).

Among these, the critical E49, I79, F140, H144, G149 and R198 residues are fully conserved in $_{Dp}$Tus (Figure 3A). However, in $_{Et}$Tus the corresponding $_{Ec}$Tus I79 residue is conservatively substituted to L79. I79 is a key residue within the cytosine binding pocket of $_{Ec}$Tus that makes a critical interaction with the C(6) base in the TT-lock complex (Mulcair et al., 2006; Elshenawy et al., 2015). Of note, the replication fork traps of *E. coli* (type II, 'degenerate') and *E. tarda* (type I, 'ancestral') are extremely different, *i.e.* six and two *Ter* sites respectively (*cf* Figure 1), and their respective Tus proteins are most divergent, sharing just 27 distinct residues in addition to the 124 residues in common between the three Tus proteins (Figure 3B). $_{Dp}$Tus shares an additional 40 residues exclusively with $_{Et}$Tus, and another 48 distinct residues with $_{Ec}$Tus (Figure 1 & Figure 3B).

While there is no structural data for free $_{Ec}$Tus, the structural coordinates of $_{Ec}$Tus in complex with *Ter* and *Ter*-lock species are available (Kamada et al., 1996; Mulcair et al., 2006; Elshenawy et al., 2015). Protein structure homology modelling was performed on $_{Dp}$Tus using the *E. coli* structure (PDB 2I06) as a template. The $_{Et}$Tus model structure is also shown for reference (Figure 3C). Despite the high similarities between these Tus structures, some differences can be seen in the 'bottom jaw' structural domain bearing the R198 residue. Structural changes that increase or decrease the flexibility of the 'bottom jaw' could significantly affect the properties of the TT-lock (Figure 3C). Here, we decided to focus on $_{Dp}$Tus as it seemed to be the perfect choice to help us gain a better understanding of the evolution of the TT-lock mechanism within a simpler replication fork trap.


A
EcTus
DpTus
EtTus
Consistency
Unconserved 0 1 2 3 4 5 6 7 8 9 10 Conserved
B
EcTus(309)
115
27
48
124
123
40
102
EtTus (314)
DpTus (307)
C
EcTus
Glu49
Gly149
His144
Phe140
Arg198
EtTus
DpTus
Id = 48%
Id = 54%
Id = 52%
Glu49
Gly149
His144
Phe140
Arg198

262

Figure 3: *In silico* analyses of $_{Dp}$Tus. (A) Multiple sequence alignment (PRALINE) of Tus protein sequences from *E. tarda* (NC_013508), *E. coli* (U00096) and *D. paradisiaca* (NC_012880). (B) Shared and distinct amino acid residues between selected Tus proteins. Total number of Tus amino acid residues are in brackets. $_{Dp}$Tus shares a similar number of residues with both $_{Et}$Tus and $_{Ec}$Tus species. (C) Comparison of the 3D structure of $_{Ec}$Tus (PDB 2I06) with the modelled structures of $_{Dp}$Tus (QMEAN = -0.79) and $_{Et}$Tus (QMEAN = -1.89) using SWISS-MODEL. Essential cytosine binding pocket residues are boxed (A) and dark grey (C).

3.2. Thermal stability and quaternary structure analyses of $_{Dp}$TusGFP

A pET vector system was used to express similar levels of GFP-tagged $_{Ec}$Tus ($_{Ec}$TusGFP) and $_{Dp}$Tus ($_{Dp}$TusGFP) and the proteins were purified in identical fashion using a streamlined lysis and Ni-affinity chromatography process (Moreau and Schaeffer, 2012a). The two fusion proteins include a slightly different linker sequence (Figure 4A) but show no significant difference in their electrophoretic mobility (Figure 4B). DSF-GTP has been successfully applied to examine the thermal stability of $_{Ec}$TusGFP and its binding to *Ter* sites (Moreau and Schaeffer, 2013). The DSF-GTP assay can accurately measure the transition midpoint (T_m) of the temperature-dependent denaturation of TusGFP in the presence or absence of DNA (Moreau et al., 2012). Consequently, it was important to ascertain that $_{Dp}$TusGFP was compatible with the DSF-GTP assay in the same buffer conditions as for $_{Ec}$TusGFP. A T_m value of 39.4±0.7°C was determined for $_{Dp}$TusGFP in low salt buffer conditions (37.5 mM NaCl). The thermal stability of $_{Ec}$TusGFP is slightly higher in the same buffer with a T_m of 45.4±0.3°C (Figure 4C). Neither $_{Dp}$TusGFP nor $_{Ec}$TusGFP generated a T_m during a second full DSF-GTP run, immediately after an initial short one from 25–50°C, demonstrating that these Tus proteins cannot refold after heat denaturation.

287 $_{Ec}$Tus has previously been reported as a monomeric protein (Coskun-Ari et al., 1994). We
288 examined the quaternary structure of $_{Ec}$ and $_{Dp}$TusGFP by size exclusion chromatography
289 (SEC). $_{Dp}$TusGFP (63.8 kDa) and $_{Ec}$TusGFP (64.6 kDa) samples were run on a Superdex 200
290 10/300 GL column and their elution profile confirmed a monomeric quaternary structure for
291 both proteins (Figure 4D). The retention time (Rt) values of $_{Ec}$ and $_{Dp}$TusGFP peaks were 34.0
292 and 34.5 min respectively. GFP was run as a proteolysis control yielding a single peak with a
293 Rt of 35.6 min. The elution curves indicated that neither $_{Dp}$ nor $_{Ec}$TusGFP were proteolyzed,
294 and this was confirmed by SDS-PAGE (Figure 4E). SDS-PAGE analysis (Figure 4E) clearly
295 shows the removal of a protein band contaminant at 80 kDa that could be seen prior to SEC
296 (Figure 4B). The difference in Rt observed for $_{Ec}$ and $_{Dp}$TusGFP may be due to a difference in
297 their respective Stokes radius (Horiike et al., 1983), and would suggest that $_{Dp}$TusGFP could
298 be more loosely packed, consistent with its lower thermal stability. Taken together, the SEC
299 data show that $_{Dp}$Tus just like $_{Ec}$Tus is a monomer and does not significantly aggregate in the
300 buffer conditions used, despite its lower thermal stability.
301
302

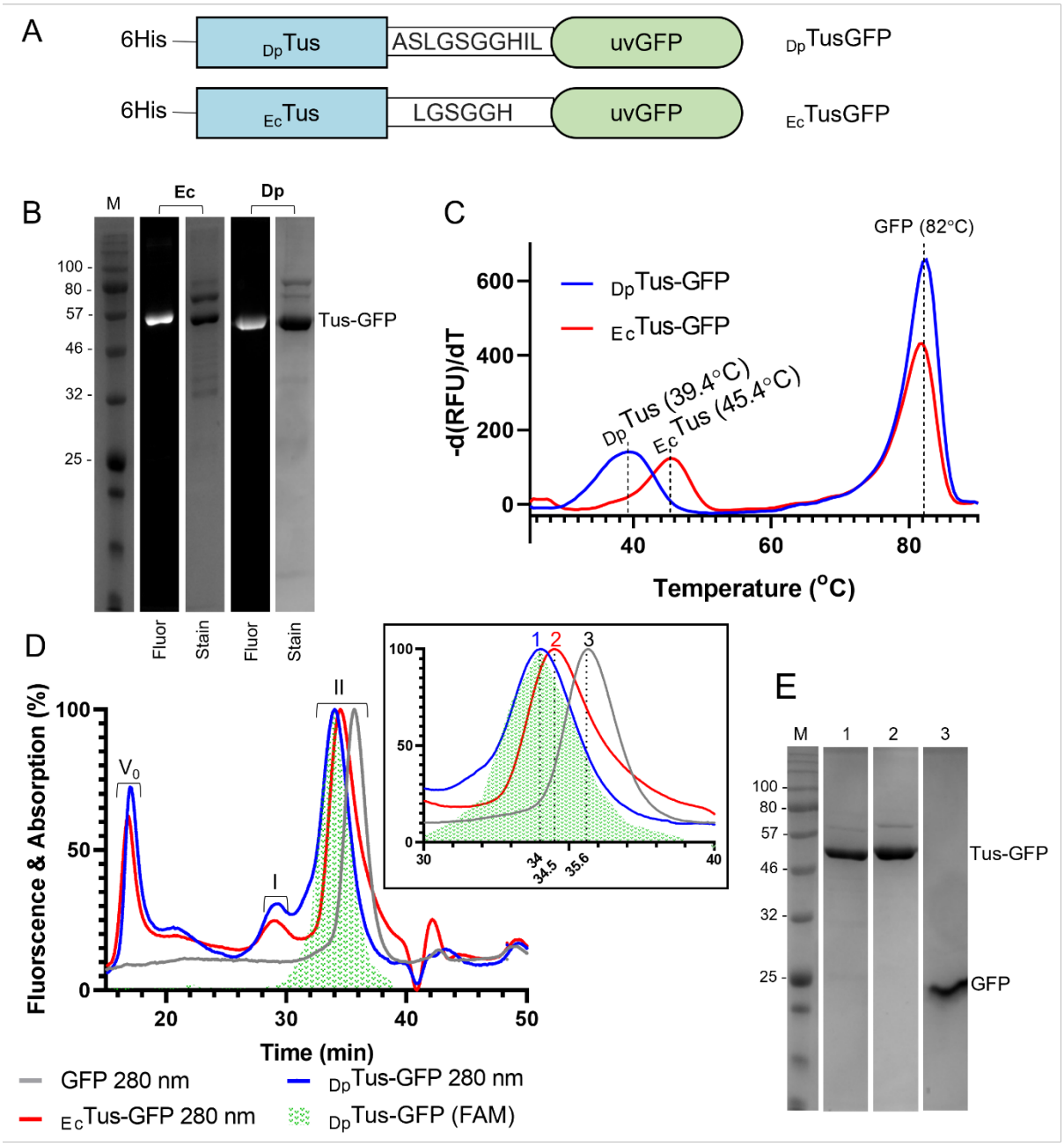


Figure 4: Comparison of $_{Dp}$TusGFP and $_{Ec}$TusGFP. (A) TusGFP constructs. Linker sequence differences are indicated. (B) SDS-PAGE of purified TusGFP. Gels were analysed with a G:BOX Chemi XRQ. From left to right: protein marker (M), GFP Fluorescence (Blue LED module, Filt 525) of $_{Ec}$TusGFP gel; same gel stained with Coomasie blue R-250; $_{Dp}$TusGFP fluorescence; and Coomassie-stained $_{Dp}$TusGFP gel. (C) DSF-GTP with $_{Dp}$TusGFP and $_{Ec}$TusGFP. (D) SEC of GFP (grey line, Rt = 35.6 min), $_{Ec}$TusGFP (red line, Rt = 34.5 min) and $_{Dp}$TusGFP (blue line, Rt = 34 min) measured at 280 nm (flowrate: 0.5 mL/min). The GFP

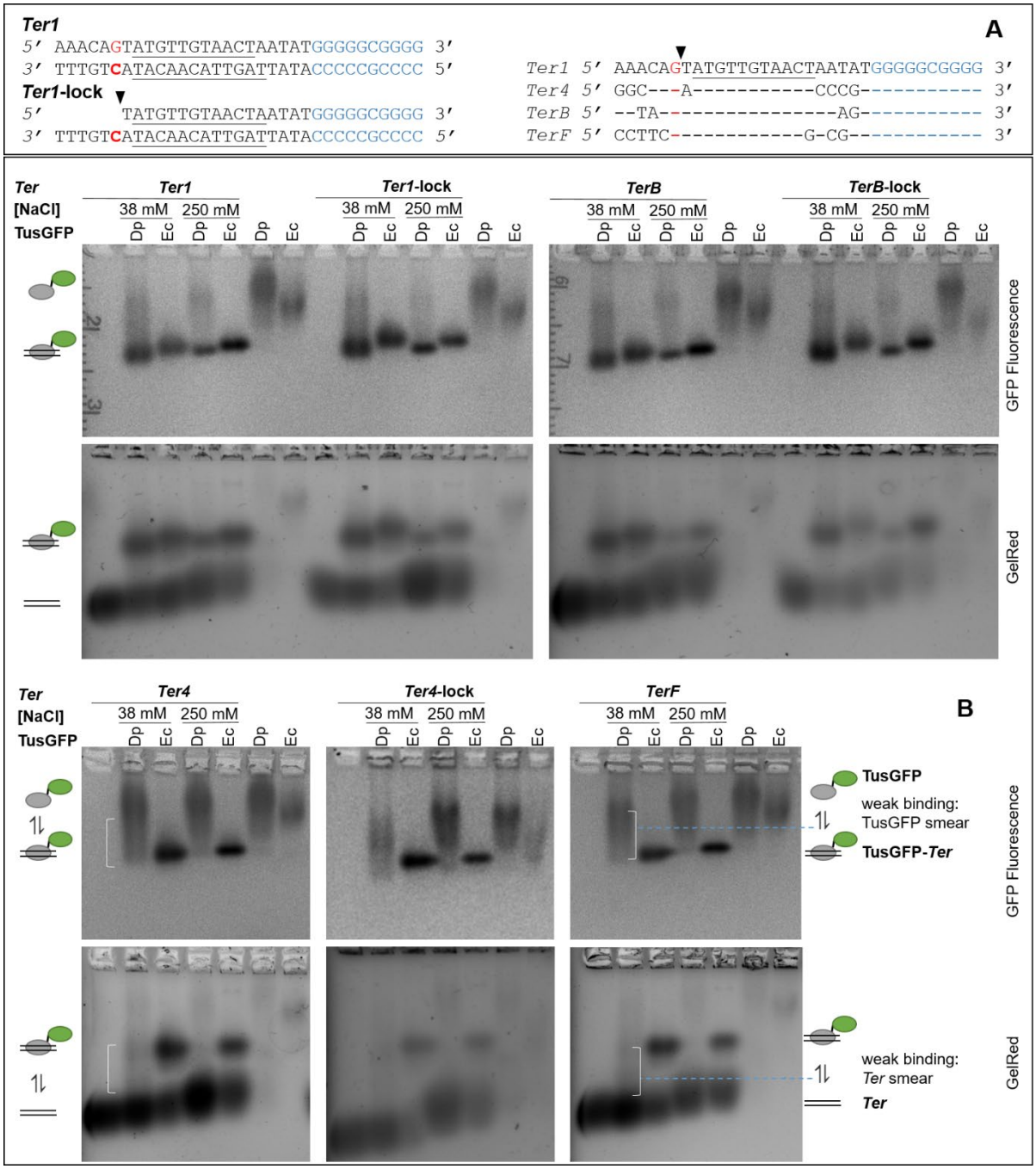
fluorescence of $_{Dp}$TusGFP fractions was measured using CFX96 Touch Real-Time PCR (FAM channel). (E) SDS-PAGE analysis of SEC peak fractions: $_{Dp}$TusGFP (1); $_{Ec}$TusGFP (2) and GFP (3). Note: SDS-PAGE samples were not heat denatured before loading to allow in-gel measurement of GFP fluorescence. As a result the TusGFP bands appear lower than expected for $_{Dp}$TusGFP (63.8 kDa) and $_{Ec}$TusGFP (64.6 kDa).

3.3. Comparison of $_{Dp}$ and $_{Ec}$Tus binding to *Ter* and *Ter*-lock by GFP-EMSA

A GFP-EMSA (Bond et al., 2017; Sorenson and Schaeffer, 2020a) was used to evaluate and compare the binding of $_{Dp}$ and $_{Ec}$TusGFP to different *Ter* and *Ter*-lock DNA species. The *Ter*-lock oligonucleotides used in this study are partially single-stranded at their non-permissive end to free the C(6), allowing it to interact with the cytosine binding pocket of Tus (Figure 5A) (Moreau and Schaeffer, 2013). The *D. paradisiaca Ter1* and *Ter4* (Toft et al., 2021) as well as the strong TT-lock forming *E. coli TerB* and pseudo-*TerF* (Moreau and Schaeffer, 2012a) were chosen for cross comparison (Figure 5A). *Ter1* and *Ter4* were selected as they were hypothesised to be the strongest and weakest $_{Dp}$Tus binders respectively, based on their loci in the type I fork trap architecture of *D. paradisiaca* (Figure 2D) as well as their sequence identities to *TerB* (Figure 5A). All protein-DNA complexes were assembled in low-salt conditions to increase their stability and maximise the chances of observing a band shift for weak *Ter* species, as well as in high-salt conditions to weaken non-specific protein-DNA interactions and make apparent any potential binding preference of $_{Dp}$TusGFP for the selected *Ter* or *Ter*-lock species.

$_{Ec}$TusGFP produced similar discrete shifted bands with all *Ter* and *Ter*-lock DNA species in both low and high salt conditions demonstrating that its binding to weak or strong *Ter* species leads to formation of stable protein-DNA complexes that cannot be distinguished using GFP-

335 EMSA. This was in stark contrast with $_{Dp}$TusGFP, which only produced discrete shifted bands
336 with *Ter1* and *Ter1*-lock as well as *TerB* and *TerB*-lock in both salt conditions (Figure 5B).
337 Interestingly, $_{Dp}$TusGFP in the presence of *Ter4, Ter4*-lock and pseudo-*TerF* only produced
338 faint smeared protein and DNA band shifts, that are almost completely lost in high salt (Figure
339 5B), indicating highly unstable protein-DNA complexes with these *Ter* species. The very weak
340 binding observed for the *Ter4* and *Ter4*-lock species at 1 μM $_{Dp}$TusGFP in low-salt conditions
341 suggests that this site is a pseudo-*Ter* in *D. paradisiaca*.
342



343

344 **Figure 5: DNA binding of TusGFP to selected *Ter* and *Ter*-lock sequences.** (A) Double-
345 stranded *Ter1* compared to partially single-stranded *Ter1*-lock sequences with free C(6) base
346 (red). The core *Ter* sequence is underlined. Base-changes between *Ter* species are indicated
347 relative to the top strand of *D. paradisiaca Ter1*. The GC rich sequence (blue) increases the T_m
348 and stability of the *Ter* species. (B) GFP-EMSA (Sorenson and Schaeffer, 2020a) was used to
349 compare the binding of $_{Dp}$TusGFP to *E. coli TerB* (strong), *TerB*-lock (strong) and pseudo-

TerF (weak), with its innermost *Ter1*, outermost *Ter4* and *Ter*-lock analogues. The DNA-binding of $_{Ec}$TusGFP with the same species was examined for reference. TusGFP (1 μM) were combined with *Ter* DNA (1.1 μM) in low and high-salt conditions and analysed by agarose gel electrophoresis. Top gels represent GFP fluorescence of TusGFP bands (Blue LED module, Filt525 on a G:BOX Chemi XRQ). Bottom gels represent GelRed fluorescence of DNA bands (TLUM mid-wave, UV06). GFP-EMSA was repeated twice per reaction.

3.4. Ranking of *D. paradisiaca Ter* sites and lock formation using DSF-GTP

DSF-GTP has previously been applied to examine the salt resistance of $_{Ec}$TusGFP in complex with various *Ter* and *Ter*-lock sequences (Moreau and Schaeffer, 2013). The authors found that the major factor that increased the salt resistance of $_{Ec}$Tus is the formation of the TT-lock complex due to the increased number of base-specific interactions over non-specific interactions (Moreau and Schaeffer, 2013). Here, DSF-GTP was performed to directly compare the binding of $_{Dp}$ and $_{Ec}$TusGFP to various *Ter* and *Ter*-lock DNA (Figure 6A) in low, moderate and high salt conditions (*i.e.* 37.5-250 mM NaCl).

The thermal stability of $_{Dp}$TusGFP increased at a linear rate of ~6°C/[NaCl (M)] which is twice the rate observed for $_{Ec}$TusGFP (~3°C/[NaCl (M)]), indicating that more stabilising interactions occur with salt ions for $_{Dp}$ than $_{Ec}$TusGFP (Figure 6B). No significant increase in T_m was observed for $_{Dp}$TusGFP in the presence of pseudo-*TerF*, *Ter4* or *Ter4*-lock, at all NaCl concentrations (Figure 6B) confirming our band shift data (Figure 5B). Yet, the thermal stability profile of $_{Ec}$TusGFP in the presence of *Ter4* was nearly identical to the one observed with pseudo-*TerF* (Figure 6B). Thus, the protein-DNA interactions in these complexes are mostly non-specific as previously reported for pseudo-*TerF* (Moreau and Schaeffer, 2012b,a,2013; Toft et al., 2021). Furthermore, the salt resistance profiles of $_{Ec}$TusGFP in

complex with either *Ter4* or *Ter4*-lock were essentially the same (*i.e.* curves do not cross), suggesting that *Ter4* cannot form a TT-lock with $_{Ec}$Tus. Taken together, our comparative GFP-EMSA and DSF-GTP data suggest that *Ter4* is a pseudo-*Ter* site that is not part of the *D. paradisiaca* replication fork trap.

Unexpectedly, all other *Ter*-lock species stabilised $_{Dp}$TusGFP significantly more, at any NaCl concentration, than their corresponding *Ter* species (*e.g.* $\Delta T_{m(37.5\ mM\ NaCl)}$ = +20.6±0.3°C for *TerB*-lock and +15.3±0.5°C for *TerB*). This is in complete contrast to what is observed with $_{Ec}$TusGFP, where the strong *Ter* species are more stabilising than their corresponding *Ter*-lock species in low salt (*e.g.* $\Delta T_{m(37.5\ mM\ NaCl)}$ = +20.4±0.4°C for *TerB*-lock and +23.4±0.4°C for *TerB*) (Figure 6B). For $_{Ec}$TusGFP, this trend inverts at 250 mM NaCl where more base-specific protein-DNA interactions subside with *TerB*-lock than with *TerB*. In fact, *TerB*, *Ter1* and their respective *Ter*-lock species are nearly identical in their respective protein stabilizing effects. At 144 mM NaCl, the ΔT_m of $_{Dp}$TusGFP in complex with *Ter1* and *Ter1*-lock ($\Delta T_{m(144\ mM\ NaCl)}$ = +1.8±1.2°C and +11.4±1.2°C respectively) are indistinguishable with *TerB* and *TerB*-lock ($\Delta T_{m(144\ mM\ NaCl)}$ = +2.1±1.0°C and +11.1±1.1°C respectively). The salt-resistance trends obtained with *Ter2* and *Ter2*-lock also suggest a strong *Ter* capable of forming a TT-lock. Indeed, $_{Dp}$TusGFP in complex with *Ter2*-lock was just slightly more sensitive to salt ($\Delta T_{m(144\ mM\ NaCl)}$ = +10.8±0.6°C) than with *Ter1*-lock.

In *E. coli*, it has previously been shown that an A20T base substitution in *TerB* reduces 3-fold the fork arrest activity (33% fork arrest activity compared to *TerB*) while it increases ~3-fold the K_{obs} (90 pM for *TerB* to 290 pM for A20T *TerB*) (Coskun-Ari and Hill, 1997). The same base substitution is present in *Ter3* and *TerH* (Figure 6A), calling into question as to whether *Ter3* would be capable of arresting a replication fork in either bacterial species. *In vitro, TerH* has previously been shown to be a moderate Tus binder that cannot form a TT-lock with

398 $_{Ec}$TusGFP (Moreau and Schaeffer, 2012a). More recently, *TerH* has been reclassified as a
399 pseudo-*Ter* site due to its inability to arrest replication forks *in vivo* in addition to its low protein
400 occupancy (*i.e.* only 13% of the Tus binding observed at *TerB* in overexpressed Tus conditions)
401 (Toft et al., 2021). As such, we hypothesised that *TerH* and its pseudo-lock form could help
402 identify if *D. paradiasiaca Ter3* is a true or pseudo-*Ter* site. We found that *Ter3* and *TerH* are
403 quite similar with respect to their $_{Ec}$TusGFP stabilization profiles and trends, however *Ter3* is
404 slightly more stabilising (Figure 6B). At 144 mM NaCl, the T_m values obtained for $_{Ec}$TusGFP
405 in complex with *Ter3* were higher than with *TerH* ($\Delta T_{m(144\ mM\ NaCl)}$ = +12.0±0.6°C and
406 +9.8±0.4°C respectively). The same trend was observed for $_{Dp}$TusGFP with *Ter3* and *TerH*
407 ($\Delta T_{m(144\ mM\ NaCl)}$ = +1.5±0.4°C and +1.2±0.6°C respectively).

408 The *TerH*-lock dataset (Figure 6B) obtained with $_{Dp}$TusGFP was rather unexpected. $_{Ec}$TusGFP
409 in complex with *TerH*-lock has been shown to be less stable than with *TerH* at 150 mM KCl
410 using either SPR (Moreau and Schaeffer, 2012a) or DSF-GTP (Moreau and Schaeffer, 2013),
411 *i.e. TerH*-lock $t_{1/2}$ = 412 s and *TerH* $t_{1/2}$ = 1244 s. A G(5) base next to C(6) in *TerH* had been
412 suggested to impair TT-lock formation due to possible bulky steric hindrances as the same base
413 substitution is also found in pseudo-*TerF* (Moreau and Schaeffer, 2012a). As such, we expected
414 to see a similar stabilization of $_{Dp}$TusGFP with either *TerH* or *TerH*-lock at 144 mM NaCl.
415 However, *TerH*-lock was far more stabilizing than *TerH*. It seems that the much lower affinity
416 of $_{Dp}$TusGFP for double-stranded *Ter* DNA makes it possible to detect the specific stabilizing
417 effects of even very weak *Ter*-lock species. However, we cannot dismiss the alternative that
418 *TerH*-lock is simply forming a better TT-lock with $_{Dp}$TusGFP despite having no biological
419 relevance here.

420 In summary, we found that *Ter1* and *Ter3* are similar, and slightly less stabilizing than *Ter2*
421 with $_{Dp}$TusGFP (Figure 6C). However, the $_{Dp}$TusGFP-*Ter3*-lock data suggest formation of a

significantly weaker protein-DNA complex than with *Ter1*-lock and *Ter2*-lock (Figure 6C). Taken together, a clear ranking is obvious for *D. paradisiaca Ter* sites with respect to their binding and ability to lock onto $_{Dp}$Tus, *i.e. Ter1* ≈ *Ter2* > *Ter3*.

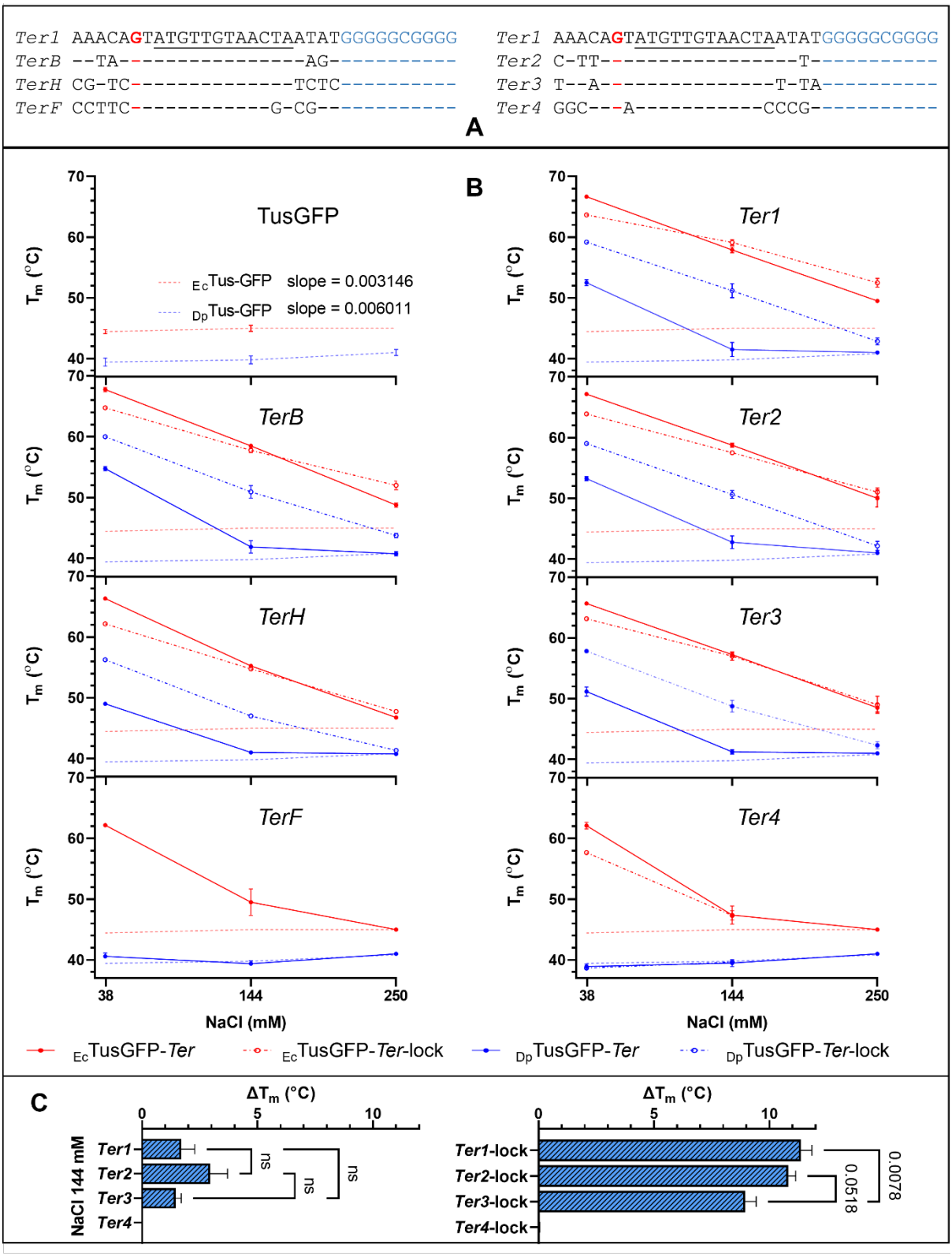


Figure 6: Thermal stability profiles of $_{Dp}$ and $_{Ec}$TusGFP bound to *Ter* and *Ter*-lock sequences in low, moderate and high ionic strength. (A) Sequence alignment of the *D. paradisiaca* and *E. coli Ter* sequences. The conserved 12 bp core sequence is underlined and the G(6) base complementary to C(6) is red. The GC rich sequence (blue) increases the T_m of the *Ter* species. (B) Thermal stability curves of $_{Dp}$TusGFP (blue) and $_{Ec}$TusGFP (red) in complex with *Ter* (continuous lines) and *Ter*-lock (dotted lines) species. Reactions were performed at 1 µM protein and 1.1 µM DNA in the presence of 37.5, 144 and 250 mM NaCl ($n \geq 3$). A linear regression model was used to determine the significant difference for the salt dependence slopes for free TusGFP ($p=0.0393$, DFd=27). (C) Ranking of *Ter* and *Ter*-lock-dependent thermal stabilization of $_{Dp}$TusGFP. ΔT_m values were calculated as $T_{m(TusGFP\text{-}Ter)}$ - $T_{m(TusGFP)}$ at 144 mM NaCl. P-values are indicated.

3.6. Major differences in DNA binding between $_{Dp}$ and $_{Ec}$Tus

Our DSF-GTP data revealed striking differences between $_{Dp}$ and $_{Ec}$TusGFP with respect to their *Ter* and *Ter*-lock binding modalities. The significantly lower thermal stability and salt-resistance of the $_{Dp}$TusGFP complex formed with *Ter* compared with *Ter*-lock species is puzzling. Here, the pseudo-*Ter4* was used to differentiate the non-specific and base-specific DNA binding contributions that are driving Tus-*Ter* and TT-lock complex formation for each Tus species. As such, the thermal stability curves obtained with *Ter4* were used as a proxy to highlight the thermal stabilization contributions of non-specific DNA binding to the proteins which are negligible in the case of $_{Dp}$TusGFP. For this, the ΔT_m values obtained with *Ter4* at NaCl concentrations ranging from 37.5-400 mM were simply overlaid on the ΔT_m curves obtained for each TusGFP with *TerB* and *Ter1* as well as their respective *Ter*-lock species (Figure 7). These *Ter* sites were chosen as they are the strongest *Ter* sites (vicinal to *tus*) in their respective bacterial replication fork traps. Our data presentation allows for easy

451 comparison of the individual and cumulative contributions of base-specific, non-specific (*i.e.*
452 DNA backbone) as well as concentration-dependent NaCl effects on the thermal stability of $_{Ec}$
453 and $_{Dp}$TusGFP.

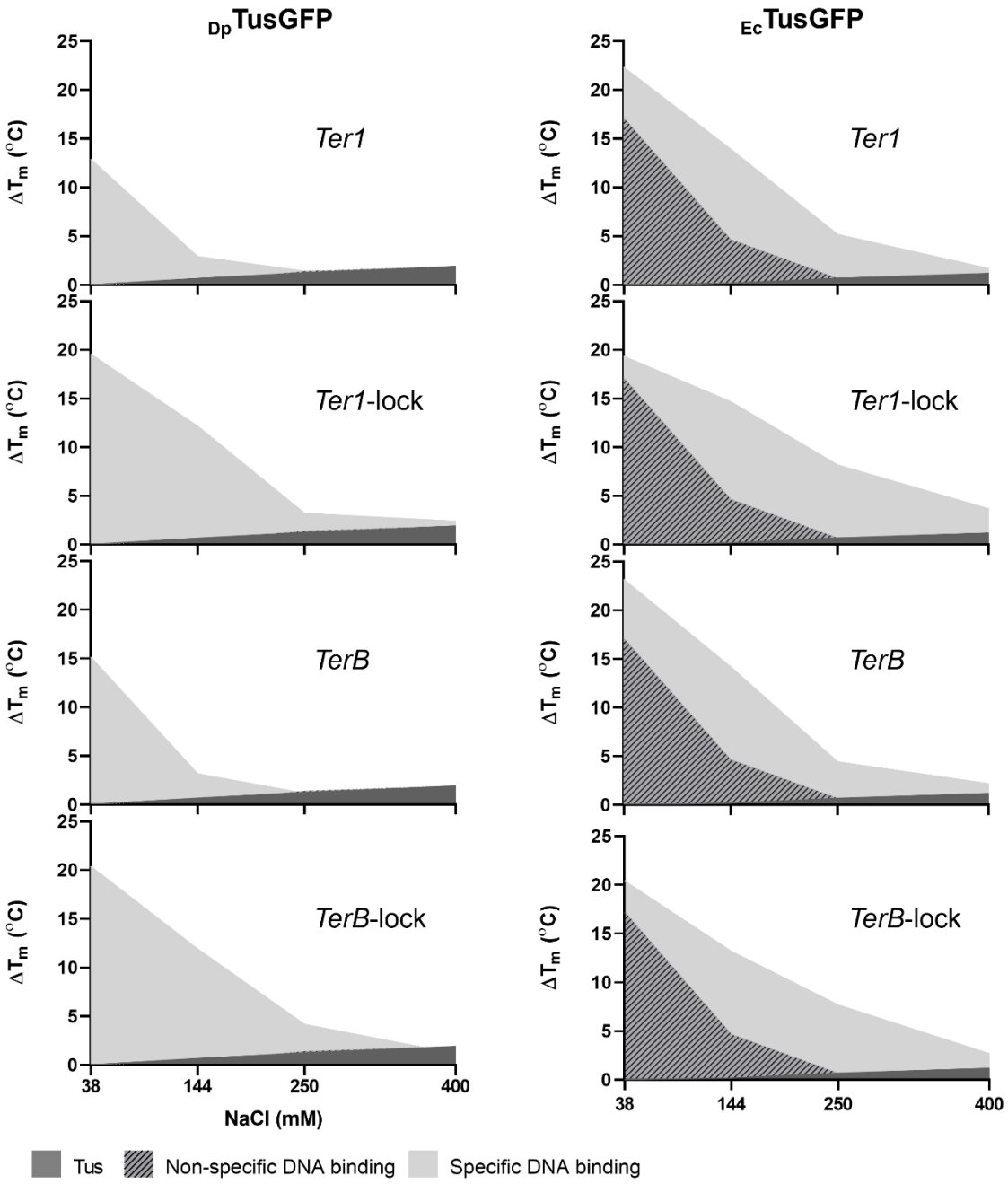


Figure 7: Differences in the specific and non-specific DNA binding properties of $_{Dp}$ and $_{Ec}$TusGFP in complex with strong *Ter* and *Ter*-lock species. The thermal stability curves of pseudo-*Ter4* obtained with $_{Dp}$ or $_{Ec}$TusGFP were overlaid to highlight the thermal stabilization contributions from non-specific DNA binding to these proteins. ΔT_m values were calculated as $T_{m(\text{free TusGFP or TusGFP-}Ter)}$ - $T_{m(\text{TusGFP at 37.5 mM NaCl})}$ to indicate the concentration-dependent stabilization effects of NaCl on free TusGFP.

The ΔT_m values obtained for $_{Dp}$TusGFP-*Ter1*-lock were systematically larger than with *Ter1* at all NaCl concentrations. The same pattern was observed with *TerB*-lock when compared to *TerB*. Hence, a greater number of protein-DNA interactions must occur with *Ter*-lock species in $_{Dp}$Tus. If these interactions are base-specific, the ionic strength of the buffer should have a comparatively lesser effect on the ΔT_m of $_{Dp}$TusGFP-*Ter*-lock complexes as base-specific protein interactions should be impacted less than electrostatic (non-specific) interactions, which is indeed what we do observe. As such, we can argue that $_{Dp}$Tus makes considerably more specific-contacts with *Ter*-lock than *Ter*, highlighting the importance of TT-lock formation to arrest replication forks in *D. paradisiaca*.

The $_{Ec}$TusGFP data with the same DNA species were in complete contrast, especially for *Ter* binding (*e.g.* $_{Dp}$TusGFP-*Ter1* $\Delta T_{m(144\ mM\ NaCl)}$ = + 1.7°C and $_{Ec}$TusGFP-*Ter1* $\Delta T_{m(144\ mM\ NaCl)}$ = + 15°C). It is clear that $_{Dp}$Tus has substantially weaker affinity for double stranded *Ter* than $_{Ec}$Tus which can reasonably be attributed to a comparatively lesser number of protein-DNA contacts. The salt-resistance profiles of these protein complexes are also dramatically different. Overall, it is evident that $_{Ec}$Tus-DNA complexes are systematically more resistant to salt (*cf* $_{Dp}$TusGFP-*Ter1* $\Delta T_{m(250\ mM\ NaCl)}$ = 0°C and $_{Ec}$TusGFP-*Ter1* $\Delta T_{m(144\ mM\ NaCl)}$ = + 4.5°C). Despite the major differences noted above, at 37.5 mM NaCl, the binding of *Ter*-lock species to $_{Dp}$TusGFP or $_{Ec}$TusGFP results in a near identical thermal stabilization of these proteins. Yet the non-specific protein-DNA binding contributions highlighted by the hashed area in the $_{Ec}$TusGFP data is quasi inexistent in $_{Dp}$TusGFP (Figure 7). This is probably the most striking difference between these proteins with respect to their DNA binding profile. If we tentatively subtract the non-specific contributions toward protein-DNA complex formation we could argue that different environmental growth conditions could have been the driving forces behind the different binding modes of $_{Ec}$ and $_{Dp}$Tus, where $_{Ec}$Tus evolved to bind and lock onto *Ter* in higher ionic strength conditions. One simple way to increase the salt resistance of a protein-

DNA complex is via increase of the number of basic amino acid residues and non-specific electrostatic interactions. The significant differences in net charge and theoretical isoelectric point (pI) between $_{Ec}$Tus (UniProt ID: P16525, (Asp + Glu) = 33, (Arg + Lys) = 43, net charge = +10, pI = 9.57) and $_{Dp}$Tus (UniProt ID: C6C580, (Asp + Glu) = 35, (Arg + Lys) = 39, net charge = +4, pI = 8.67) support this proposition.

With respect to the differences in DNA-binding modes between $_{Dp}$ and $_{Ec}$Tus, we hypothesise that $_{Ec}$Tus makes more interactions with the non-permissive end of *Ter* (*i.e.* with both DNA strands) than $_{Dp}$Tus. Furthermore, the larger increase in thermal stabilization of $_{Dp}$Tus with the *Ter*-lock species likely indicates that a greater number of protein-DNA contacts occur within its cytosine binding pocket compared to $_{Ec}$Tus.

3.7. Affinity of $_{Dp}$TusGFP for *Ter1* and *Ter1*-lock

DSF-GTP has previously been applied to compare the affinity of $_{Ec}$TusGFP for various *Ter* and *Ter*-lock sequences. It is the only method that is compatible with a wide range of ionic strength conditions (Moreau and Schaeffer, 2013). The differences in observed dissociation constant (K_{obs}) values for the various *Ter* and *Ter*-lock species in complex with $_{Ec}$TusGFP obtained by DSF-GTP (Moreau and Schaeffer, 2013), mirrored previous surface plasmon resonance (Moreau and Schaeffer, 2012a) and qPCR binding assay data (Moreau and Schaeffer, 2012b). The T_m values of $_{Dp}$TusGFP were determined in the presence of *Ter1* and *Ter1*-lock at concentrations ranging from 1 to 10 µM and compared to $_{Ec}$TusGFP in both low and moderate salt conditions.

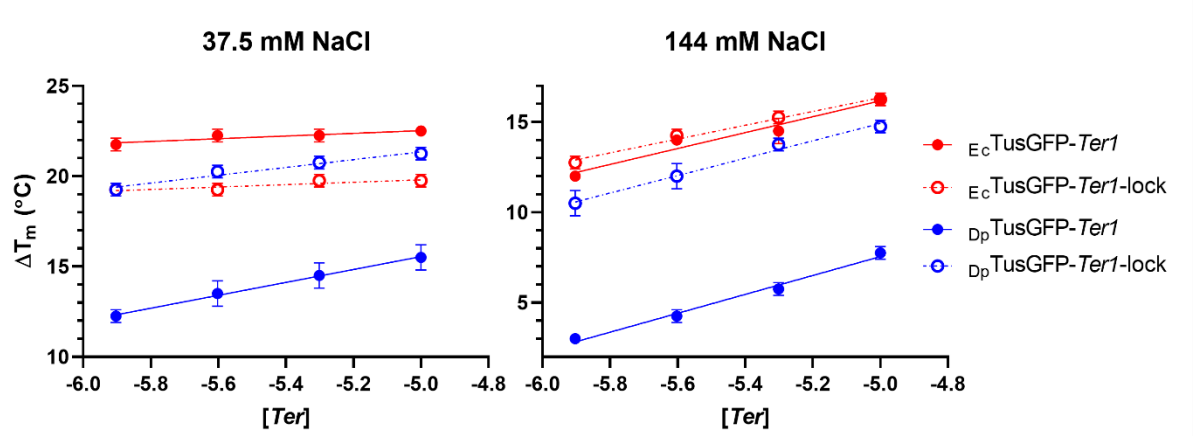


Figure 8: Affinity of $_{Dp}$ and $_{Ec}$TusGFP for *Ter1* and *Ter1*-lock in low and moderate salt conditions. DSF-GTP data were obtained with 1 µM constant TusGFP in the presence of increasing concentrations of *Ter* species (1-10 µM) and are presented with ΔT_m expressed as a function of the logarithm of *Ter* concentration. Error bars represent the SD (N = 2). The K_{obs}, slope and R-square values are shown in Table 1.

In low salt conditions (37.5 mM NaCl), the K_{obs} for $_{Ec}$TusGFP-*Ter1* and *Ter1*-lock complexes could not be accurately determined (<10^{-15} M) yet *Ter1* is clearly the most stabilizing. Previous

$_{Ec}$Tus low-salt data obtained with *TerB* and *TerB*-lock mirror this pattern (Moreau and Schaeffer, 2013). $_{Dp}$TusGFP-*Ter1* was the least stable complex ($K_{obs(37.5\ mM\ NaCl)} = 10^{-9.4}$). Interestingly, the relative thermal stability of the $_{Dp}$TusGFP-*Ter1*-lock complex ($K_{obs(37.5\ mM\ NaCl)} = 10^{-14.9}$ M) increases significantly more than the $_{Ec}$TusGFP-*Ter1*-lock with rising *Ter1*-lock concentration. An even steeper rise in thermal stabilization can be seen with $_{Dp}$TusGFP-*Ter1*. We are not sure what the steeper thermal stabilization slope means for $_{Dp}$Tus but suspect that it is the result of different kinetics in complex formation between these species. Astoundingly, the binding of $_{Dp}$TusGFP to *Ter1*-lock yields an impressive 100,000-fold increase in complex stability (*i.e.* when compared over *Ter1*) in these low-salt conditions. Indeed, we are not aware of any DNA-binding protein where the removal of 6 nucleotides from one strand of its DNA target leads to such a massive increase in complex stability.

Table 1: *Ter1* and *Ter1*-lock-induced stabilization of $_{Dp}$ and $_{Ec}$TusGFP

Complex	NaCl (mM)	X-intercept	K_{obs} (M)	R squared	Slope
$_{Dp}$TusGFP-*Ter1*	37.5	-9.35	$10^{-9.4}$	0.9968	3.57
	144	-6.44	$10^{-6.4}$	0.9885	5.23
$_{Dp}$TusGFP-*Ter1*-lock	37.5	-14.89	$10^{-14.9}$	0.9657	2.16
	144	-8.10	$10^{-8.1}$	0.9894	4.40
$_{Ec}$TusGFP-*Ter1*	37.5	-35.14	$<10^{-15*}$	0.8526*	0.75
	144	-8.68	$10^{-8.7}$	0.9571	4.82
$_{Ec}$TusGFP-*Ter1*-lock	37.5	-34.80	$<10^{-15*}$	0.8000*	0.66
	144	-9.28	$10^{-9.3}$	0.9888	3.82

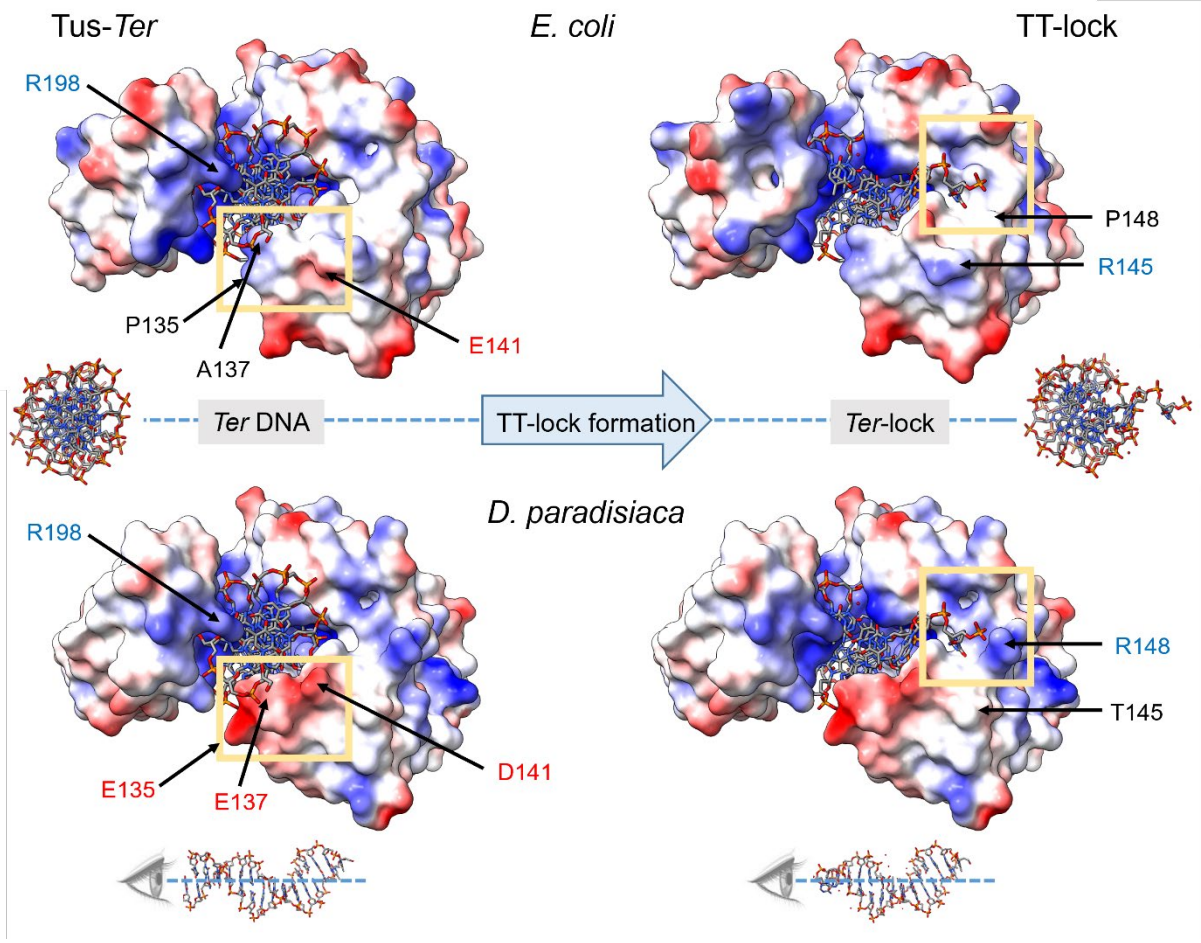
* K_{obs} cannot be estimated

At 144 NaCl, K_{obs} values could be determined for all complexes allowing a direct comparison between $_{Ec}$ and $_{Dp}$TusGFP. Overall, in these moderate salt conditions, $_{Dp}$TusGFP-*Ter1* is ~ 200-fold less stable than $_{Ec}$TusGFP-*Ter1* yet $_{Dp}$TusGFP-*Ter1*-lock is only 15-fold less stable than $_{Ec}$TusGFP-*Ter1*-lock. $_{Ec}$TusGFP has a similar affinity for *Ter1* and *Ter1-lock* obscuring the specific contributions of TT-lock formation in these conditions. Interestingly, the large

differences that were observed for the slopes between $_{Ec}$ and $_{Dp}$TusGFP complexes in low salt were not apparent at 144 mM NaCl. However, the binding of $_{Dp}$TusGFP to *Ter1*-lock ($K_{obs(144\ mM\ NaCl)} = 10^{-8.1}$ M) is still 50-fold stronger than to *Ter1* ($K_{obs(144\ mM\ NaCl)} = 10^{-6.4}$ M).

The relatively low pI of $_{Dp}$Tus (compared to $_{Ec}$Tus) already hinted toward the possibility that less electrostatic interactions might occur with the DNA backbone. Detailed examination of the electrostatic surface potential of the modelled structures of the $_{Dp}$Tus in complex with *Ter* and *Ter*-lock yielded simple yet elegant explanations for its unique DNA-binding profile (Figure 9). Indeed, major differences in electrostatic surface potential could be observed at the non-permissive face of $_{Ec}$ and $_{Dp}$Tus. Of these, the P135E and A137E switches in $_{Dp}$Tus (compared to $_{Ec}$Tus) immediately stand out as differences that would significantly reduce its non-specific binding to DNA and overall affinity to *Ter*. On the other hand, it seems that the cytosine-binding pocket is strengthened by the P148R switch in $_{Dp}$Tus, via formation of additional electrostatic interactions. The relatively moderate salt resistance of the $_{Dp}$TusGFP-*Ter1*-lock further supports this proposition.

The significant increase in stability for the *Ter*-lock complex is obviously fundamental to fork arrest activity in $_{Dp}$Tus. However, Tus must first find and bind a double stranded *Ter* sequence and wait until the GC(6) base pair is unzipped by a translocating DNA helicase to form a locked complex. The significantly lower binding affinity of $_{Dp}$Tus for *Ter1* raises a fundamental question as to how $_{Dp}$Tus can find a *Ter* site and stay on it? This is particularly puzzling, if we consider that no more than 100 $_{Ec}$Tus proteins are present in an *E. coli* bacteria at any given time (Gottlieb et al., 1992).



553

554 **Figure 9: Major differences in the electrostatic surface potential at the non-permissive**
555 **faces of $_{Ec}$Tus and $_{Dp}$Tus complexes.** Crystal structures of $_{Ec}$Tus in complex with *Ter* (top left,
556 PDB 2I05) and *Ter*-lock DNA (top right, PDB 2I06). The $_{Dp}$Tus structure was modelled using
557 SWISS-MODEL (https://swissmodel.expasy.org/) with $_{Ec}$Tus-*Ter* (PDB 2I05) as template and
558 aligned in ChimeraX (available at: https://www.cgl.ucsf.edu/chimerax/) to the $_{Ec}$Tus-*Ter*
559 (bottom left, PDB 2I05) and $_{Ec}$TT-lock (bottom right, PDB 2I06). Boxed regions display the
560 difference in charge distributions and residues of the proteins. Details of *Ter* and *Ter*-lock
561 structures without Tus are also shown for reference. Note: The portion of *Ter* sequence in the
562 $_{Ec}$Tus-*Ter* structure (PDB 2I05) is identical to *Ter1*. See Figure S1 for further sequence details
563 and structural viewpoints, and Figure S2 for reference *E. coli* NUCPLOT protein-DNA
564 interaction maps.

4. Conclusions and perspective

A recent *in silico* analysis of the fork trap architecture of *D. paradisiaca* identified four putative *Ter* sites, organised as a single unbreachable *Ter* site vicinal to *tus* on the left chromosomal arm to arrest clockwise travelling replication forks, and three *Ter* sites in opposite orientation on the right chromosomal arm with the outermost sites potentially acting as backups (Toft et al., 2021). Our DSF-GTP data unambiguously showed that $_{Dp}$Tus can bind to *Ter* and form a TT-lock complex with *Ter1-3* but not with *Ter4*, which we reclassified as a pseudo-*Ter* site. The innermost *Ter1* and *Ter2* are the strongest and most stabilising $_{Dp}$Tus binding sites followed by a slightly weaker *Ter3*. The ranking of these *Ter* sites correlated well with their individual chromosomal locations and sequence identities to *TerB.* The strongest *D. paradisiaca* TT-lock complexes are formed with the two innermost *Ter* sites and the GC-skew switch coincides with an equal use of these for fork arrest. These findings negate a requirement for backup *Ter* sites to arrest replication forks in *D. paradisiaca*. The obvious trend of weakening and increasing mutations from innermost to outermost *Ter* sites suggests that there are minimal evolutionary driving forces at play for outer *Ter* sites to remain functional. As such, it seems probable that *D. paradisiaca* adopted a relatively wide plasmid-borne fork trap (Galli et al., 2019) that was perfectly positioned to stop a clockwise moving replication fork yet required further adjustments to halt the anticlockwise moving fork.

The 50-fold lower $_{Dp}$Tus binding affinity for double-stranded *Ter1* compared with the *Ter1*-lock in moderate salt conditions (144 mM NaCl) was completely unexpected and most intriguing. This is in sharp contrast with the $_{Ec}$Tus protein, which binds with comparable high affinity to both *Ter* and *Ter*-lock species, and where the differences in dissociation rates and affinity between these species are mostly masked at 150 mM KCl (Mulcair et al., 2006; Moreau

590 and Schaeffer, 2012a). In fact, formation of the *E. coli* TT-lock can only be observed at 250
591 mM KCl as a ~40-fold decrease in dissociation rate for the *Ter*-lock species (Mulcair et al.,
592 2006; Moreau and Schaeffer, 2012a). It would be reasonable to assume that $_{Dp}$Tus would not
593 significantly bind to its *Ter* sites if expressed in the low numbers reported for $_{Ec}$Tus (Gottlieb
594 et al., 1992). Thus, given its moderate affinity to *Ter*, it can be argued that $_{Dp}$Tus must be
595 expressed at much higher levels, enabling it to act as a repressor and bind its other *Ter* sites in
596 *D. paradisiaca*. We propose that the integrated transcriptional negative feedback loop of the
597 Tus-*Ter* system essentially allows the affinity of Tus for its vicinal *Ter* site to evolve relatively
598 freely. However, the different salt resistance and binding affinity profiles of $_{Ec}$ and $_{Dp}$Tus
599 proteins for the highly conserved *Ter* sequence, also highlight the importance of the TT-lock
600 mechanism in fork arrest. This dynamic process must be central to the selection of functional
601 Tus proteins while they evolve and adapt to different conditions (*e.g.* ionic strength, pH and
602 temperature).

603 The pseudo-*Ter4* data highlighted the large contribution of non-specific and electrostatic
604 interactions towards $_{Ec}$Tus-DNA complex formation, and the extremely low affinity of $_{Dp}$Tus
605 for non-specific DNA even in low salt conditions. Initially, we hypothesised that the lower
606 affinity of $_{Dp}$Tus for its vicinal *Ter1* was due to a reduced number of basic residues, compared
607 to $_{Ec}$Tus. The $_{Dp}$Tus modelled structure showed obvious differences in the number and
608 distribution of basic and acidic residues, offering a better justification for its moderate *Ter*
609 binding. The pseudo-*TerH*-lock data indicated that significant stabilizing interactions are
610 occurring within the cytosine binding pocket of $_{Dp}$Tus, yet this species cannot form a lock with
611 $_{Ec}$Tus. Examination of the cytosine binding pocket revealed the presence of R148 that is
612 perfectly placed to increase the lock strength in $_{Dp}$Tus by promoting additional electrostatic
613 interactions with the phosphate backbone. This finding can explain the $_{Dp}$TusGFP stability data
614 with pseudo-*TerH*-lock as well as the lower salt resistance observed with *Ter*-lock species. We

cannot be sure how the 6 nucleotides that are missing in the *Ter1*-lock species (compared to *Ter1,* Figure 5A) are affecting the stability of $_{Dp}$TusGFP, however based on the $_{Ec}$TusGFP data we would expect a detrimental effect mirroring previous $_{Ec}$Tus-*TerB*-lock data (Mulcair et al., 2006; Moreau and Schaeffer, 2012a,2013). Thus, it is possible that *E. coli* and *D. paradisiaca* TT-lock complexes may be equally stable despite their differences in K_{obs}, especially if we consider the different environmental conditions of these bacteria. Indeed, gut contents have variable salinities that can reach 4% with a high sodium over potassium ratio (Seck et al., 2019), while potato tubers that are commonly infected by *Dickeya* species (Helmann et al., 2022) have much lower salinity with a high potassium to sodium ratio (Beals, 2019). Overall, our data support the notion that a loss of protein-DNA contacts in the Tus-*Ter* complex can be compensated for by increasing the interactions (*e.g.* electrostatic) in the locked species.

Further studies are clearly warranted to examine the kinetics, dynamics and structures of $_{Dp}$Tus in complex with its *Ter* and *Ter*-lock species. Mutational analyses, especially of the cytosine binding pocket residues would be valuable to reveal their specific contributions to the TT-lock. While the presence of six high-affinity and lock forming *Ter* sites (*TerA-E* and *G*) in *E. coli* will likely remain obscure, we are confident that future comparative studies of simpler fork traps (*e.g. E. tarda* and *C. neteri*) and their associated Tus proteins will shed new light on this unique system. Finally, with its surprisingly negligible binding to non-target DNA, $_{Dp}$Tus is poised to become an attractive alternative to $_{Ec}$Tus, which can be problematic in biotechnological applications (Toft et al., 2021; Toft et al., 2022) where specificity is key (Morin et al., 2010; Johnston et al., 2014; Larsen et al., 2014).

CRediT Authorship Contribution Statement

Casey J. Toft: Investigation, Formal analysis, Writing - Original Draft, Writing - Review & Editing, Visualization, Data Curation. Alanna E. Sorenson: Supervision, Resources, Writing -

Review & Editing. Patrick M. Schaeffer: Conceptualization, Methodology, Supervision, Writing - Original Draft, Writing - Review & Editing, Visualization, Project administration.

Funding

Casey J. Toft was supported by a merit-based Research Training Program Scholarship (James Cook University).

Declarations of interest

None

References

Askin SP, Schaeffer PM. A universal immuno-PCR platform for comparative and ultrasensitive quantification of dual affinity-tagged proteins in complex matrices. Analyst 2012;137(22):5193-5196.

Askin SP, Morin I, Schaeffer PM. Development of a protease activity assay using heat-sensitive Tus-GFP fusion protein substrates. Analytical biochemistry 2011;415(2):126-133.

Askin SP, Bond TEH, Schaeffer PM. Green fluorescent protein-based assays for high-throughput functional characterization and ligand-binding studies of biotin protein ligase. Anal Methods-Uk 2016;8(2):418-424.

Beals KA. Potatoes, Nutrition and Health. Am J Potato Res 2019;96(2):102-110.

Bond TEH, Sorenson AE, Schaeffer PM. Functional characterisation of Burkholderia pseudomallei biotin protein ligase: A toolkit for anti-melioidosis drug development. Microbiol Res 2017;199:40-48.

Bussiere DE, Bastia D, White SW. Crystal structure of the replication terminator protein from B. subtilis at 2.6 A. Cell 1995;80(4):651-60.

Cooper A, Williams NL, Morris JL, Norton RE, Ketheesan N, Schaeffer PM. ELISA and immuno-polymerase chain reaction assays for the sensitive detection of melioidosis. Diagn Micr Infec Dis 2013;75(2):135-138.

Coskun-Ari FF, Hill TM. 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 1997;272(42):26448-56.

Coskun-Ari FF, Skokotas A, Moe GR, Hill TM. Biophysical characteristics of Tus, the replication arrest protein of Escherichia coli. The Journal of biological chemistry 1994;269(6):4027-34.

Dahdah DB, Morin I, Moreau MJ, Dixon NE, Schaeffer PM. Site-specific covalent attachment of DNA to proteins using a photoactivatable Tus-Ter complex. Chem Commun (Camb) 2009(21):3050-2.

Elshenawy MM, Jergic S, Xu ZQ, Sobhy MA, Takahashi M, Oakley AJ, Dixon NE, Hamdan SM. Replisome speed determines the efficiency of the Tus-Ter replication termination barrier. Nature 2015;525(7569):394-8.

Galli E, Ferat JL, Desfontaines JM, Val ME, Skovgaard O, Barre FX, Possoz C. Replication termination without a replication fork trap. Sci Rep 2019;9(1):8315.

Gottlieb PA, Wu S, Zhang X, Tecklenburg M, Kuempel P, Hill TM. Equilibrium, kinetic, and footprinting studies of the Tus-Ter protein-DNA interaction. The Journal of biological chemistry 1992;267(11):7434-43.

Griffiths AA, Andersen PA, Wake RG. Replication terminator protein-based replication fork-arrest systems in various Bacillus species. J Bacteriol 1998;180(13):3360-7.

Helmann TC, Filiatrault MJ, Stodghill PV. Genome-Wide Identification of Genes Important for Growth of Dickeya dadantii and Dickeya dianthicola in Potato (Solanum tuberosum) Tubers. Front Microbiol 2022;13:778927.

Henderson TA, Nilles AF, Valjavec-Gratian M, Hill TM. 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 2001;265(6):941-53.

Hill TM, Pelletier AJ, Tecklenburg ML, Kuempel PL. Identification of the DNA sequence from the E. coli terminus region that halts replication forks. Cell 1988;55(3):459-66.

Hill TM, Tecklenburg ML, Pelletier AJ, Kuempel PL. Tus, the Trans-Acting Gene Required for Termination of DNA-Replication in Escherichia-Coli, Encodes a DNA-Binding Protein. P Natl Acad Sci USA 1989;86(5):1593-1597.

Horiike K, Tojo H, Yamano T, Nozaki M. Interpretation of the stokes radius of macromolecules determined by gel filtration chromatography. J Biochem 1983;93(1):99-106.

Hugouvieux-Cotte-Pattat N, des-Combes CJ, Briolay J, Pritchard L. Proposal for the creation of a new genus Musicola gen. nov., reclassification of Dickeya paradisiaca (Samson et al. 2005) as Musicola paradisiaca comb. nov. and description of a new species Musicola keenii sp. nov. Int J Syst Evol Micr 2021;71(10).

Johnston EB, Kamath SD, Lopata AL, Schaeffer PM. Tus-Ter-lock immuno-PCR assays for the sensitive detection of tropomyosin-specific IgE antibodies. Bioanalysis 2014;6(4):465-476.

Jorgensen SW, Liberti SE, Larsen NB, Lisby M, Mankouri HW, Hickson ID. Esc2 promotes telomere stability in response to DNA replication stress. Nucleic Acids Research 2019;47(9):4597-4611.

Kamada K, Horiuchi T, Ohsumi K, Shimamoto N, Morikawa K. Structure of a replication-terminator protein complexed with DNA. Nature 1996;383(6601):598-603.

Kamath SD, Johnston EB, Koplin JJ, Eckart J, Rolland JM, O'Hehir RE, Schaeffer PM, Allen KJ, Lopata AL. A novel ImmunoPCR based strategy to detect serum concentrations of IgE antibodies specific to the major shrimp allergen tropomyosin in shrimp allergic adults and children. Allergy 2014;69:391-392.

Kono N, Arakawa K, Tomita M. Validation of bacterial replication termination models using simulation of genomic mutations. PLoS One 2012;7(4):e34526.

Kuempel PL, Pelletier AJ, Hill TM. Tus and the Terminators - the Arrest of Replication in Prokaryotes. Cell 1989;59(4):581-583.

Larsen NB, Hickson ID, Mankouri HW. Tus-Ter as a tool to study site-specific DNA replication perturbation in eukaryotes. Cell Cycle 2014;13(19):2994-2998.

Louarn J, Patte J, Louarn JM. Evidence for a fixed termination site of chromosome replication in Escherichia coli K12. J Mol Biol 1977;115(3):295-314.

Midgley-Smith SL, Dimude JU, Taylor T, Forrester NM, Upton AL, Lloyd RG, Rudolph CJ. Chromosomal over-replication in Escherichia coli recG cells is triggered by replication fork fusion and amplified if replichore symmetry is disturbed. Nucleic Acids Res 2018;46(15):7701-7715.

Moreau MJ, Schaeffer PM. Differential Tus-Ter binding and lock formation: implications for DNA replication termination in Escherichia coli. Mol Biosyst 2012a;8(10):2783-91.

Moreau MJ, Schaeffer PM. A polyplex qPCR-based binding assay for protein-DNA interactions. Analyst 2012b;137(18):4111-3.

730 Moreau MJ, Schaeffer PM. Dissecting the salt dependence of the Tus-Ter protein-DNA
731 complexes by high-throughput differential scanning fluorimetry of a GFP-tagged Tus. Mol
732 Biosyst 2013;9(12):3146-54.

733 Moreau MJ, Morin I, Schaeffer PM. Quantitative determination of protein stability and ligand
734 binding using a green fluorescent protein reporter system. Mol Biosyst 2010;6(7):1285-92.

735 Moreau MJJ, Morin I, Askin SP, Cooper A, Moreland NJ, Vasudevan SG, Schaeffer PM.
736 Rapid determination of protein stability and ligand binding by differential scanning
737 fluorimetry of GFP-tagged proteins. RSC Advances 2012;2(31):11892-11900.

738 Morin I, Dixon NE, Schaeffer PM. Ultrasensitive detection of antibodies using a new Tus-
739 Ter-lock immunoPCR system. Mol Biosyst 2010;6(7):1173-5.

740 Morin I, Schaeffer PM, Askin SP, Dixon NE. Combining RNA-DNA swapping and
741 quantitative polymerase chain reaction for the detection of influenza A nucleoprotein. Anal
742 Biochem 2012;420(2):121-6.

743 Mulcair MD, Schaeffer PM, Oakley AJ, Cross HF, Neylon C, Hill TM, Dixon NE. A
744 molecular mousetrap determines polarity of termination of DNA replication in E. coli. Cell
745 2006;125(7):1309-19.

746 Neylon C, Kralicek AV, Hill TM, Dixon NE. Replication termination in Escherichia coli:
747 structure and antihelicase activity of the Tus-Ter complex. Microbiol Mol Biol Rev
748 2005;69(3):501-26.

749 Norouzi M, Panfilov S, Pardee K. High-Efficiency Protection of Linear DNA in Cell-Free
750 Extracts from Escherichia coli and Vibrio natriegens. Acs Synth Biol 2021;10(7):1615-1624.

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

754 Rudolph CJ, Upton AL, Stockum A, Nieduszynski CA, Lloyd RG. Avoiding chromosome
755 pathology when replication forks collide. Nature 2013;500(7464):608-11.

756 Salplachta J, Kubesova A, Horky J, Matouskova H, Tesarova M, Horka M. Characterization
757 of Dickeya and Pectobacterium species by capillary electrophoretic techniques and MALDI-
758 TOF MS. Anal Bioanal Chem 2015;407(25):7625-7635.

759 Samson R, Legendre JB, Christen R, Fischer-Le Saux M, Achouak W, Gardan L. Transfer of
760 Pectobacterium chrysanthemi (Burkholder et al. 1953) Brenner et al. 1973 and Brenneria
761 paradisiaca to the genus Dickeya gen. nov as Dickeya chrysanthemi comb. nov and Dickeya
762 paradisiaca comb. nov and delineation of four novel species, Dickeya dadantii sp nov.,
763 Dickeya dianthicola sp nov., Dickeya dieffenbachiae sp nov and Dickeya zeae sp nov. Int J
764 Syst Evol Micr 2005;55:1415-1427.

765 Schaeffer PM, Headlam MJ, Dixon NE. Protein-protein interactions in the eubacterial
766 replisome. Iubmb Life 2005;57(1):5-12.

767 Seck EH, Senghor B, Merhej V, Bachar D, Cadoret F, Robert C, Azhar EI, Yasir M, Bibi F,
768 Jiman-Fatani AA, Konate DS, Musso D, Doumbo O, Sokhna C, Levasseur A, Lagier JC,
769 Khelaifia S, Million M, Raoult D. Salt in stools is associated with obesity, gut halophilic
770 microbiota and Akkermansia muciniphila depletion in humans. Int J Obes (Lond)
771 2019;43(4):862-871.

Sharma B, Hill TM. 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 1995;18(1):45-61.

Sorenson AE, Schaeffer PM. Electrophoretic Mobility Shift Assays with GFP-Tagged Proteins (GFP-EMSA). Methods Mol Biol 2020a;2089:159-166.

Sorenson AE, Schaeffer PM. High-Throughput Differential Scanning Fluorimetry of GFP-Tagged Proteins. Methods Mol Biol 2020b;2089:69-85.

Toft CJ, Sorenson AE, Schaeffer PM. Rise of the terminator protein tus: A versatile tool in the biotechnologist's toolbox. Analytica Chimica Acta 2022;1213:339946.

Toft CJ, Moreau MJJ, Perutka J, Mandapati S, Enyeart P, Sorenson AE, Ellington AD, Schaeffer PM. Delineation of the Ancestral Tus-Dependent Replication Fork Trap. Int J Mol Sci 2021;22(24).

Wilce JA, Vivian JP, Hastings AF, Otting G, Folmer RH, Duggin IG, Wake RG, Wilce MC. Structure of the RTP-DNA complex and the mechanism of polar replication fork arrest. Nature structural biology 2001;8(3):206-10.

Willis NA, Panday A, Duffey EE, Scully R. Rad51 recruitment and exclusion of non-homologous end joining during homologous recombination at a Tus/Ter mammalian replication fork barrier. Plos Genet 2018;14(7).

Willis NA, Chandramouly G, Huang B, Kwok A, Follonier C, Deng C, Scully R. BRCA1 controls homologous recombination at Tus/Ter-stalled mammalian replication forks. Nature 2014;510(7506):556-9.