Contents

Abbreviations

The standard abbreviations for bases and amino acids were used.

A_{xxx}	absorption at xxx nm
AFM	atomic force microscopy
Ap	ampicillin
Apra	apramycin
ATP	adenosine 5'-triphosphate
bp	base pair
BSA	bovine serum albumin
Cm	chloramphenicol
DNA	deoxyribose nucleic acid
DSB	double-stranded break
dsDNA	double-stranded DNA
dNTP	deoxynucleoside triphosphate
DTT	$_{DL}$-dithiothreitol
EDTA	ethylenediaminetetraacetic acid
EMSA	electrophoretic mobility shift assay
EtBr	ethidium bromide
FPLC	fast performance liquid chromatography
IPTG	isopropyl β-D-thiogalactoside
Km	kanamycin
LB broth	Luria-Burrous broth
LIS	low ionic strength
MA	minimal agar
MC	mitomycin C
nt	nucleotide
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
Pol I-IV	DNA polymerase I-IV
^{r}	resistant
RNA	ribonucleic acid
rpm	rotations per minutes
^{s}	sensitive
SDS	sodium dodecyl sulphate
SSB	the single-strand DNA binding protein
ssDNA	single stranded deoxyribose nucleic acid

TBE	tris/borate/EDTA
Tm	trimethoprim
Tris	tris(hydroxymethyl)aminomethane
UV	ultraviolet
v/v	Volume per volume
wt	wild type
w/v	weight per volume

Acknowledgements

I would like to thank my supervisor, Professor Robert Lloyd, for his sustaining support and advice during these years. I also appreciate the close supervision provided by Tim Moore, Geoff Briggs and Akeel Mahdi. I was also greatly benefited from Christian Rudolf's strict science attitude and Jane Grove's expertise on manipulating cells and words. I would also like to thank Stuart Wood, Amy Upton and Jing Zhang, for giving me the feeling that I was not the only one on boat. I am very grateful to the technical support given by Carol Brown, Lynda Harris and Susan Grey. Thank you to Thorsten Allers not only for his excellent arrangement for seminars, but also for keeping my nearly dead computer running.

I give a special thank to Karen Bunting for invaluable discussions on crystallography and structure of proteins, Kevin Bailey and Susan Liddell for checking my proteins with Mass Spectrometry. I also thank Paul McEwan for collaboration in interpreting the RdgC crystal structure. Thank you to my advisor, Professor Ed Louis and my basketball buddy, Professor Fred Sablitzky for many ideas on my project.

I would never give enough thanks to my parents and my sister for their emotional and financial support. The support of my friends also encoloured my PhD life, especially the support from my flatmate Wei Dong and my girlfriend Guanyi Hou. Finally, I would like to thank the International Office for giving me the International Office China Scholarship for the financial support of my PhD project.

Statement

The work presented in this thesis was carried out solely by the author (unless cited) during three years from October 2005 in the Institute of Genetics, University of Nottingham, whilst registered as a full-time PhD. Student. This thesis is comprised of original work, which has not been presented for examination in any other form. This research was supported by the MRC.

Whilst carrying out the work described in this thesis, I was fully aware of the hazards associated with the materials and techniques I was using, as advised in the Control of Substances Hazardous to Health regulations. The guidelines laid down in these regulations and in the departmental rules were strictly adhered to at all times. I was also aware of, and followed, the regulations concerning the use and disposal of radioisotopes.

Abstract

Previous studies found that RdgC protein plays a role in the DNA repair system in *Escherichia coli*. In *recBC sbcBC* strains, loss of *rdgC* made growth of the strains dependent upon recombination, hence *R*ecombination *D*ependent *G*rowth. RdgC was also found to regulate the activity of RecA, a key protein in recombination, both *in vivo* and *in vitro*. The function of the protein, however, remains unknown.

In this study, I purified and crystallised the RdgC protein. The crystal structure of the protein was then revealed as a homo-dimer, with a head to head, tail to tail organisation, resembling a ring structure. To further investigate how RdgC binds DNA and its *in vivo* functionality, point mutations and chunk deletions were designed and constructed; and I examined all the mutant proteins in DNA binding shift assays *in vitro* and in synthetic lethality assays *in vivo*. A DNA binding model was then proposed based on the results of the DNA binding shift assays. The mutant studies *in vivo* reinforce the idea that the DNA binding activity is crucial for RdgC's function in *Escherichia coli*.

Chapter 1
Introduction

1.1 Overview

Chromosomal DNA encodes almost all biological traits of a living organism. Any changes to it therefore have a potential to interfere with the normal life form. Evolution has introduced sophisticated DNA repair systems to limit this potential to a minimum, among which is DNA recombinational repair. In *Escherichia coli*, this process is catalysed by RecA, the central protein in homologous recombination. The activity of RecA is strictly controlled to avoid unnecessary recombination, which is otherwise dangerous. The RdgC protein is thought to be a negative regulator of RecA by binding to dsDNA, protecting it from being used by an active RecA species.

In this chapter, DNA damage and repair, RecA mediated DNA recombinational repair and its interaction with RdgC are reviewed.

1.2 DNA damage and repair

Unlike damaged proteins that can be replaced by using the source codes of genes, damaged DNA must be repaired.

Mismatch repair (MMR) deals with the rare mismatches left after replication. In *E. coli*, the MMR system distinguishes the template strand from the nascent strand by methylated adenines, and therefore cuts a segment of the unmethylated nascent DNA strand containing the mismatched base. DNA

polymerase III and DNA ligase will fill the gap and finish the repair (Kolodner, 1995).

Base-excision repair (BER) is active against common DNA lesions (such as the products of cytosine and adenine deamination). The repair depends on a variety of DNA glycosylases, each of which is generally specific for one type of lesion. DNA glycosylases recognise and remove the affected base by cleaving the *N*-glycosyl bond, generating an apurinic or apyrimidinic site, commonly referred to as an AP site. A segment of DNA including the AP site is then removed by AP endonucleases, leaving a gap which will be filled by DNA polymerase I and DNA ligase (McCullough et al., 1999).

Nucleotide-excision repair (NER) generally targets at DNA lesions that cause large distortions in the helical structure of DNA. In *E. coli*, a multisubunit enzyme excises a fragment of 12 to 13 nucleotides with the lesion, and the resulting gap is filled by DNA polymerase I and DNA ligase (Sancar, 1996).

Direct repair is special because it does not need a complementary strand to provide information. It applies a variety of enzymes that recognise distinctive DNA lesions, which commonly arise from typical base modifications. Each of these modifications can be reversed by an appropriate class of enzymes. For example, pyrimidine dimers resulting from an ultraviolet light-induced reaction are usually reversed to two pyrimidines by DNA photolyases in *E. coli* (Murphy et al., 2008).

The DNA repair mechanisms described above generally work for lesions in double-stranded DNA, the undamaged strand providing the correct genetic

information to restore the damaged strand to its original state. However, there are also situations where the complementary strand is absent or is itself damaged, which often arises when a replication fork encounters an unrepaired DNA lesion. To overcome it, the DNA repair systems either 'take a guess' over the damaged region to fill random codes in (as applied by error-prone translesion DNA synthesis, often abbreviated TLS), or repair the damaged strand with the genetic information provided by a separate, homologous chromosome. The latter is accurate and undoubtedly favoured wherever possible. In *E. coli*, this process is catalysed by RecA, and is also known as recombinational DNA repair.

1.3 RecA mediated recombination

RecA mediates DNA recombination in prokaryotes as a multi-protein filament covering a segment of single-stranded DNA. Filament formation consists of two steps: nucleation and extension, with the former generally being rate-limiting. Nucleation is more rapid on ssDNA and is greatly slowed if SSB

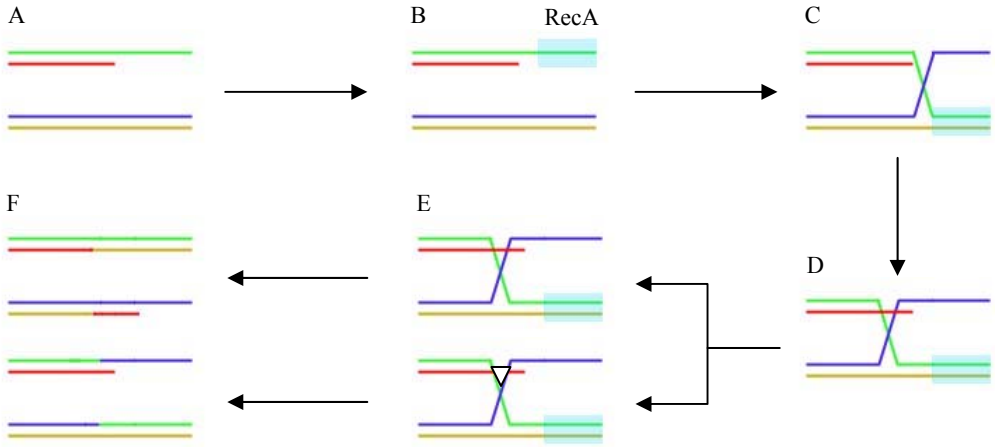


Figure 1-1 RecA mediated homologous recombination. A. ssDNA exposed for RecA binding; B. Filamentation of RecA on the exposed ssDNA; C. Re-pairing mediated by the RecA filament; D. Invasion of the branch into duplex region; E. Two possible resolutions for the recombination intermediate; F. Two corresponding fates of step E.

(Single-Stranded DNA Binding protein) is bound to the ssDNA. Filament extension occurs on the 3'-proximal end, accompanied with dissociation on the opposite end at a slower rate. This filament searches a separate duplex for sequence homologous to the covered ssDNA, and then mediates re-pairing so that the covered ssDNA anneals with its sequence-complementary strand. The resulting DNA branchpoint will be moved by RuvAB or RecG. If the branch is driven into a duplex region of the template DNA, it will form a four-way junction, also known as Holliday junction (Figure 1-1 D). Resolvases such as RuvC may cut the junction, monomerising the recombination intermediate. UvsX in bacteriophage T4 (Kodadek et al., 1988; Beernink and Morrical, 1999), RadA in archaea (Seitz et al., 1998), and Dmc1 (Gupta et al., 2001) and Rad51 (Ogawa et al., 1993; Sung, 1994) in eukaryotes are orthologues of RecA and act in DNA recombination in a very similar way. Its ubiquitous distribution underlines its significance throughout evolution.

Unlike most other DNA repair pathways that deal with problems on double-stranded DNA, recombination mediated by RecA focuses on single-stranded DNA. It is the substrate at which RecA acts—both for DNA recombination and the SOS response. This might raise a potential for initiating recombination whenever ssDNA is exposed. However, any exposed ssDNA is immediately bound and hence protected by the ssDNA binding protein, SSB. This is almost why the RecBCD pathway loads RecA directly to initiate recombination at dsDNA ends.

1.3.1 The RecBCD pathway

RecBCD is a DNA helicase and ATP-dependent exonuclease (ExoV) that acts on blunt or nearly blunt dsDNA ends (Taylor and Smith, 1985). Free DNA duplex ends (DSBs) are caused by a variety of factors. In addition to imperfect templates that replication forks act upon, UV light, ionising radiation, oxygen radicals, DNA-damaging agents and inappropriate chromosomal DNA cleavage may all generate DSBs (Dillingham and Kowalczykowski, 2008).

The resulting DSBs are not always blunt or nearly blunt and therefore need to be trimmed by ssDNA exonucleases (such as SbcB, SbcCD and RecJ), or be filled by polymerases, to create an appropriate template for RecBCD. Once RecBCD is bound to a free DNA end, it translocates along and unwinds the duplex, and at the same time degrades both strands. RecB and RecD are both helicase motors, with RecD being generally faster. RecB also possesses 5' and 3' ssDNA nuclease activities, but is most favourably positioned to cleave the 3' ssDNA tail, and this strand is therefore hydrolysed much more vigorously than the 5' ssDNA tail (Figure 1-2 B). The unwind and degradation of both strands proceed until the RecBCD complex encounters a specific, properly oriented sequence, called Chi sequence (Crossover Hotspot Instigator, 5'-GCTGGTGG-3'), or χ (Sprague et al., 1978; Bianco and Kowalczykowski, 1997). This sequence is recognised and remains tightly bound by the RecC protein. This event prevents the exit of the 3' ssDNA strand from RecC and therefore halting its hydrolysis by RecB. At the same time the 5' ssDNA strand continues to exit and is cleaved more readily by RecB. This change of activity results in a 3' ssDNA tail, a platform on which RecBCD then loads a RecA filament (Figure

1-2 C) (Dixon and Kowalczykowski, 1993; Anderson and Kowalczykowski, 1997; Anderson and Kowalczykowski, 1997; Dillingham and Kowalczykowski, 2008). In this pathway, the ssDNA is never exposed to SSB.

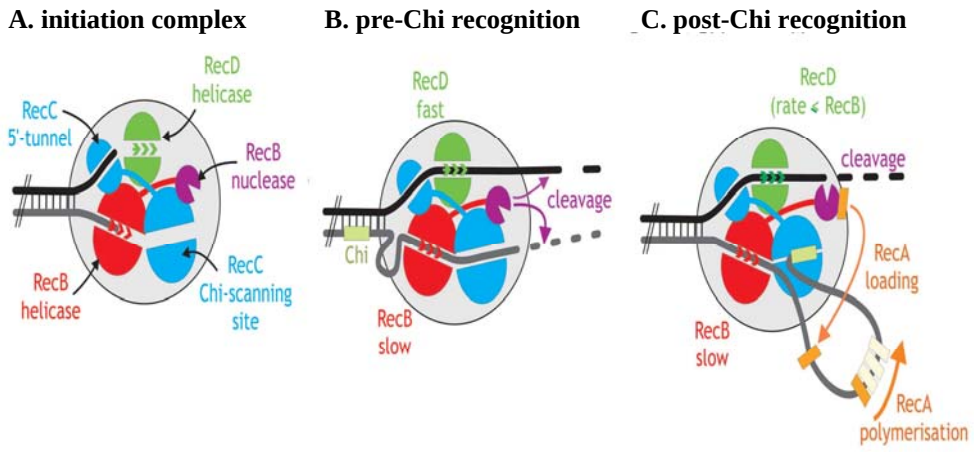


Figure 1-2 Model for RecBCD enzyme mechanism (Dillingham and Kowalczykowski, 2008). The three subunits are colour coded as described above, with important functional regions in the structure also labelled.

In addition to DSBs, RecA can also use ssDNA gaps, but only with the assistance of the RecFOR pathway, which can direct RecA to sites even if they are bound by SSB.

1.3.2 The RecFOR pathway

Under normal conditions, the RecFOR pathway promotes the repair at ssDNA gaps and is less active than the RecBCD pathway. However, in a nucleases depleted background, the RecFOR pathway is also able to process DSBs for loading RecA filament when the RecBCD pathway is not available (Ryder et al., 1996). In fact, the significance of the RecFOR pathway can be evaluated in terms of evolution. The existence of the RecBCD pathway is only restricted in

a few species, whereas RecFOR (in some species RecF or RecO or both are missing) are almost as ubiquitous as the RecA protein (Rocha et al., 2005). Unlike RecBCD working as a single complex, RecR forms a complex either with RecF or with RecO, but not with both at the same time (Webb et al., 1997; Morimatsu and Kowalczykowski, 2003). RecF is a DNA binding protein, with increased affinity for dsDNA (Umezu and Kolodner, 1994; Webb et al., 1995; Webb et al., 1997). RecO binds both dsDNA and ssDNA, and is able to anneal complementary oligonucleotides and catalyse invasion of duplex DNA by a complementary ssDNA (Luisi-DeLuca and Kolodner, 1994; Kantake et al., 2002). The *E. coli* RecR protein does not possess any enzymatic or DNA binding activities, but it increases the affinity of both RecO and RecF for DNA when complexed with them (Umezu and Kolodner, 1994; Webb et al., 1995; Webb et al., 1997). In a classic RecFOR pathway, the RecFR complex probably defines regions for loading of the RecA filament by binding to dsDNA-ssDNA junctions (Sakai and Cox, 2009). Presumably, this activity not only facilitates the initiation of RecA nucleation on ssDNA gaps coated by SSB (Morimatsu and Kowalczykowski, 2003), but also inhibits the extension of RecA filament to dsDNA (Webb et al., 1997). In this context, RecO is also required to act in concert with RecFR (Makharashvil et al., 2009; Sakai and Cox, 2009). However, RecFR was found to bind randomly on dsDNA, with no enhanced affinity to dsDNA-ssDNA junctions. An unknown factor is thus required to guide RecFR (Figure 1-3).

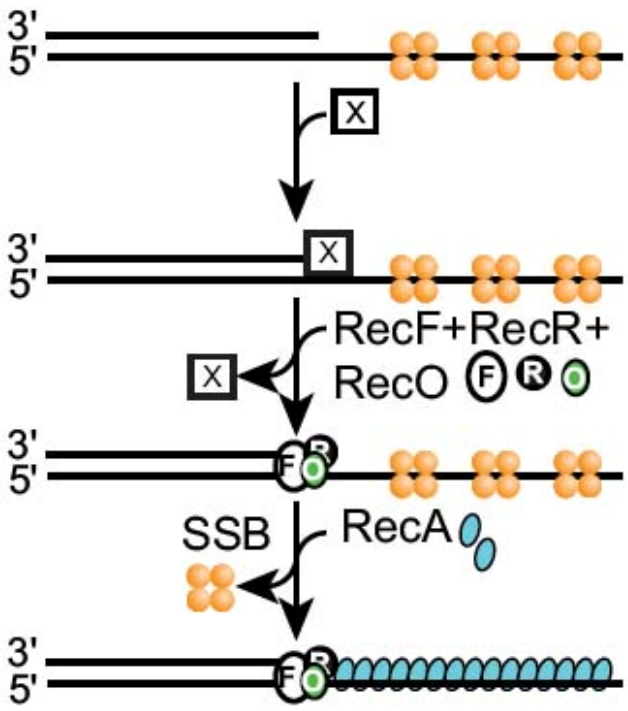


Figure 1-3 Model for RecFOR function on gapped DNA (Sakai and Cox, 2009). An unknown factor (X) is required to recognise the dsDNA-ssDNA junction and guide RecFR to the junction for loading of RecA. This process also requires RecO.

Recent studies have revealed that RecOR is able to mediate RecA loading through an alternative pathway. When there is no proximal duplex, RecOR can facilitate nucleation of RecA filaments onto SSB-coated ssDNA (Sakai and Cox, 2009). RecF is dispensable in this pathway.

In conclusion, the RecFOR proteins are able to load RecA to ssDNA even when it is protected by SSB. This poses a constant threat in that RecA may initiate recombination when it is not required, which is sometimes even dangerous.

1.4 RecA can be toxic

There is growing evidence that a number of factors have evolved to limit recombination in a variety of ways.

1) Remove RecA from ssDNA. UvrD is one of the proteins mediating this strategy (Morel et al., 1993; Veaute et al., 2005; Lestini and Michel, 2007). The UvrD protein is a 3'-5' helicase, and is typically required for NER and MMR (Kuzminov, 1999; Matson and Robertson, 2006).
2) Destabilise RecA-ssDNA filament. The RecX protein falls into this category with its ability of capping the RecA-DNA filament, thus preventing its extension (Drees et al., 2004; Lusetti et al., 2004a; Lusetti et al., 2004b).
3) Avoid exposing ssDNA unnecessarily. Although it is largely achieved by SSB, recent researches suggested PriA and Rep act in a different manner but with a similar effect. PriA is a DNA replication restart protein; the role of Rep is not clear, but is thought to assist replication fork progression (Lane and Denhardt, 1975; Yancey-Wrona and Matson, 1992) and to function in a PriA independent, PriC-Rep replication restart pathway (Sandler, 2000). Both PriA and Rep have 3'-5' helicase activity and are able to unwind the lagging strand at a stalled replication fork, thus retrieving any possible exposed ssDNA at the leading strand (Figure 1-4) (Mahdi et al., 2006).

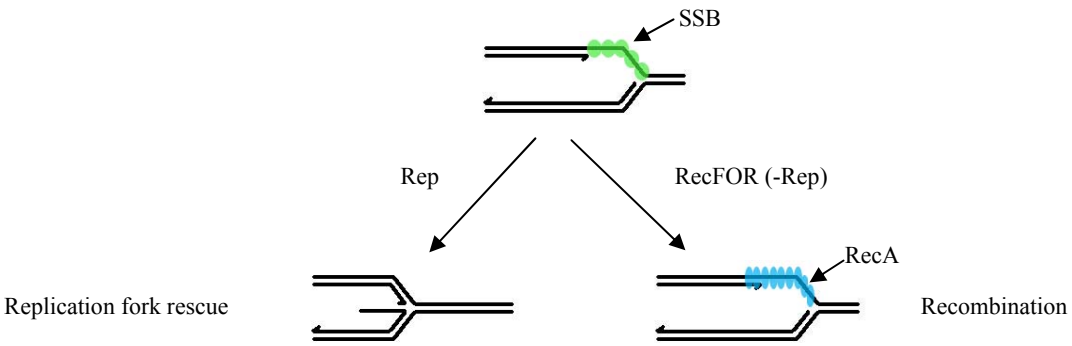


Figure 1-4 A model for how 3'-5' helicase activity of Rep may unwind the lagging strand and close the leading strand gap (the left pathway) to prevent otherwise potentially lethal activity of RecA initiated by RecFOR (the right pathway) (Mahdi et al., 2006). Flips mark the 3' ends.

A number of genetic models fit with the idea that recombination can be toxic. In the absence of UvrD, recombination is enhanced and leads to a requirement for RuvABC to maintain viability (Magner et al., 2007), unless RecA or RecFOR are eliminated. A similar situation has been observed through studies of interactions between mutations in *priA*, *rdgC* and *dnaC*.

Strains carrying null mutations in *priA* have much reduced viability, are highly sensitive to DNA-damaging agents and are defective in recombination (Nurse et al., 1991; Kogoma et al., 1996). This phenotype is presumably a reflection of the failure to promote replication restart when forks encounter lesions in the DNA (Liu and Marians, 1999). It can be suppressed by mutations in *dnaC*, the product of which normally binds the DnaB replicative helicase, directing its loading onto DNA to initiate replication at *oriC*, or to allow replication restart for repair or rescue of damaged or blocked forks.

Studies by Sandler and co-workers revealed that the *priA dnaC* strains rely on the Rep helicase and PriC protein to maintain viability (Sandler, 2000). This was certainly true of the *dnaC809* suppressor (encoding the same substitution

in DnaC as *dnaC810*). The Rep protein was thought to provide an alternative helicase activity that could create a loading pad for DnaB in the absence of PriA helicase activity. PriC is now known to provide an alternative means for loading DnaB (Heller and Marians, 2007). However, further studies by Sandler identified a *dnaC* super-suppressor, *dnaC809,820*, carrying a second substitution that eliminated the requirement for Rep and PriC (Sandler, 2000).

The inactivation of *rdgC* confers no obvious phenotype in otherwise wild type cells (Ryder et al., 1996). However, it eliminates the residual viability of *priA* null cells (Moore et al., 2003) and curtails severely the ability of *dnaC* mutations to suppress the *priA* phenotype (Moore et al., 2003).

dnaC212 is another *dnaC* mutation that can suppress the phenotype of a *priA* null strain. It was isolated by A. V. Gregg in the Lloyd laboratory (Gregg et al., 2002) and studied by Moore and colleagues (Moore et al., 2003). It was discovered that indeed *priA*$^-$ *dnaC212 rdgC*$^-$ strains grew very slowly and readily developed suppressors, which included mutations inactivating or modifying SSB, RecO and RecF (Moore et al., 2003). This observation provided the first hint that RdgC might act to limit unnecessary recombination via RecFOR mediated loading of RecA.

These data point strongly to the idea that RdgC acts in some way to limit RecA activity, and is crucial when the PriC pathway is inactivated in a *priA*$^-$ *dnaC809,820* background. This possibility was further supported when a deletion of *rdgC* was introduced into *priA*$^-$, *priA*$^-$ *dnaC810*, *priA*$^-$ *dnaC809,820* and *priA*$^-$ *dnaC809,820 priC*$^-$ strains, using synthetic lethality construct carrying an unstable plasmid with *priA*$^+$ and *lac*$^+$. Viability of the plasmid free

segregants is revealed when *lac*$^-$ (white) colonies are produced on agar plates supplemented with X-gal and IPTG (Mahdi et al., 2006). This study confirmed that *priA*$^-$ *rdgC*$^-$ cells are inviable (Figure 1-5 N6040). Deletion of *rdgC* also conferred inviability on a *priA*$^-$ *dnaC810* strain (Figure 1-5 N6111), but not on a *priA*$^-$ *dnaC809,820* strain (Figure 1-5 N6457), unless PriC was missing (Figure 1-5 N6538). Furthermore, viability could be restored to a *priA*$^-$ *dnaC810 rdgC*$^-$ strain by eliminating *recA* (Figure 1-5 N6157).

1.5 RdgC

The *rdgC* gene was discovered in 1985 in an attempt to screen for the *sbcCD* gene (Lloyd and Buckman, 1985). RdgC's involvement in DNA recombination became clear first in 1996. Ryder and colleagues found that in strains lacking nucleases, deletion of *rdgC* made the cell's growth dependent on the remaining recombination activity (Ryder et al., 1996). Its high expression level during exponential phase (~1000 molecules per cell) provided another piece of evidence for its interaction with DNA metabolism (Moore et al., 2003). *In vitro* studies then revealed that RdgC is a DNA binding protein (Moore et al., 2003).

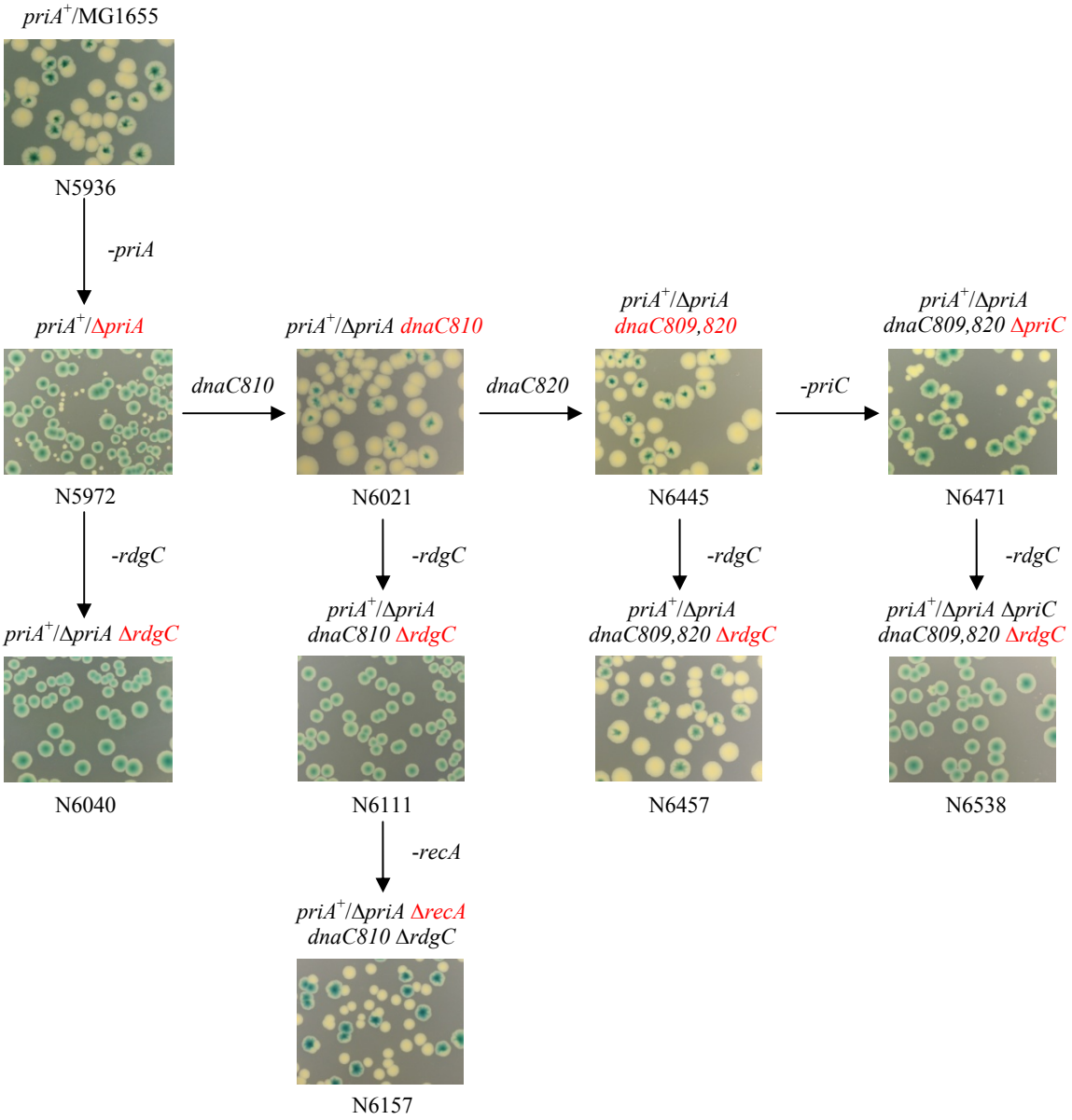


Figure 1-5 Synthetic lethality assays presenting situations where deletion of *rdgC* could have an effect (R. G. Lloyd, unpublished data). Details about how synthetic lethality assay is conducted can be found in Mahdi et al., 2006.

Cells lacking PriA are sick (N5972). It can be overcome by *dnaC810* (N6021) and *dnaC810, 820* (N6445). *dnaC810* is a weak suppressor of Δ*priA*, as deletion of *rdgC* is synthetic lethal with Δ*priA dnaC810* (N6111). This can be overcome either by an additional mutation on *dnaC* to form *dnaC810, 820* (N6457), or deletion for *recA* (N6157). With the strong suppressor (*dnaC810, 820*) of Δ*priA*, cells lacking PriA require either RdgC or PriC to sustain viability; the cells lacking both are not viable (N6538).

1.5.1 RdgC binds DNA

The RdgC protein of *Escherichia coli* is a 34 kDa protein associated with the nucleoid (Ryder et al., 1996; Murphy et al., 1999). In solution, it presents in equilibrium as monomer, dimer and tetramer, with predominant presence as a dimer (Moore et al., 2003; Tessmer et al., 2005; Drees et al., 2006).

Previous studies showed that RdgC is a DNA binding protein without any enzymatic activities (Moore et al., 2003). It appeared that RdgC binds DNA as a dimer (Moore et al., 2003). The dimer was able to bind all DNA structures tested with high affinity, including single stranded DNA, blunt duplex, flayed duplex, three-strand junction, Y-DNA and Holliday junction (Figure 1-6) (Moore et al., 2003). Interestingly, it exhibited a slight preference for dsDNA compared to ssDNA. Nonetheless, RdgC can still bind with high affinity to ssDNA when secondary structures (e.g. hairpin structures) are present (Moore et al., 2003; Drees et al., 2006). Atomic Force Microscopy studies showed an end preference of the protein, but this could not be confirmed by other experiments (Moore, 2002; Tessmer et al., 2005). The minimum lengths of DNA required for stable contact was also determined for both dsDNA (15 bp) and ssDNA (23 nt) (Moore, 2002).

The dimer formation of the protein and its different affinity for dsDNA and ssDNA had implicated more than one DNA binding sites on RdgC. However, the competition assays by introducing poly(dIdC) as a competitor to RdgC:DNA complexes and AFM studies failed to detect a second binding site. The fact that all the DNA was in the form of the highest order complex when

sufficient concentrations of RdgC were used suggests RdgC can diffuse in one-dimension along bound DNA (Moore, 2002).

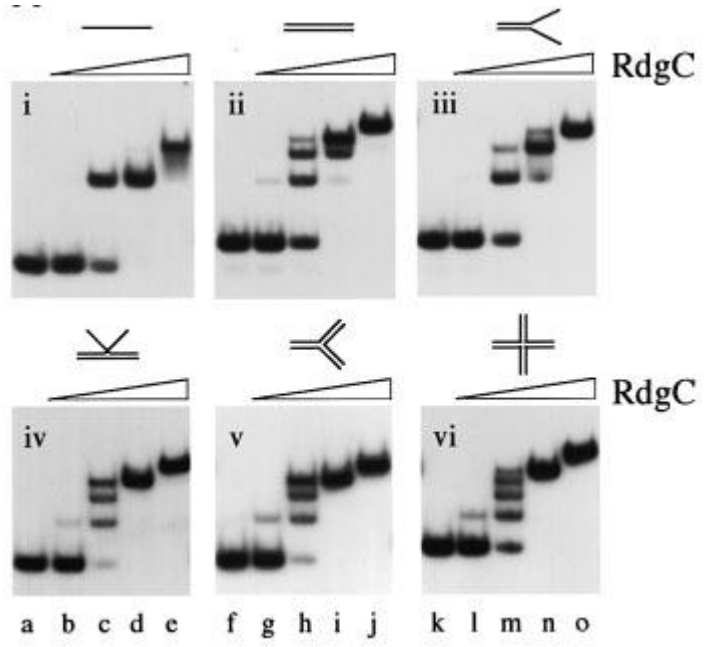


Figure 1-6 Gel assays showing DNA binding activity of RdgC. Binding reactions contained 0.1 nM DNA species, (i) 61 nucleotide ssDNA, (ii) 61 bp dsDNA, (iii) flayed duplex, (iv) three-strand junction, (v) Y-DNA, (vi) Holliday junction J12, and RdgC at 0, 0.5, 5, 50 and 500 nM in lanes a-e, f-j and k-o, respectively. (Moore et al., 2003)

1.5.2 RdgC limits RecA activities *in vitro*

How does RdgC DNA binding limit RecA activities? Drees and colleagues conducted a series of *in vitro* experiments in an attempt to answer this question (Drees et al., 2006). The RecA filament is most stable when bound to ssDNA. This stability decreases when RecA filament is on dsDNA. Although RdgC can bind to ssDNA, its affinity is not high enough for the protein to displace bound RecA. An effective inhibition was only seen when excess RdgC was added prior to the addition of RecA to saturate all ssDNA molecules. This is also

partially due to its high affinity for ssDNA secondary structures. On the other hand, its higher affinity for dsDNA compared to that of RecA enabled RdgC to effectively protect the covered dsDNA from being invaded by RecA:ssDNA filaments. RdgC can even displace RecA proteins that are bound to dsDNA. This re-enforces the idea that the function of RdgC is to prevent or limit unnecessary recombination, by competing for common substrates (Drees et al., 2006). However, given the limited number of RdgC compared to available dsDNA sites that it could bind, it would be almost impossible for RdgC to protect dsDNA from illegitimate recombination if it binds randomly along the chromosome. It has also been proposed that some proteins might act as a guide to direct RdgC. Indeed, several proteins were reported to be able to interact with RdgC in an attempt to reveal protein interactions among essential and also non-essential proteins in *E. coli* (Butland et al., 2005), including the products of *ligA*, *huB*, and *nfi*; but the interactions have not been proved to be genuine and could be mediated by DNA (all the proteins bind to DNA).

It has also been proposed that RdgC might function in different ways as monomers, dimers or tetramers. Disagreeing with Moore et al (2003) who did not observe any bimodal DNA bindings, Drees et al (2006) showed that RdgC bound duplex DNA with higher affinity when protein concentrations were either low (10 nM-100 nM) or high (>1000 nM). When the protein concentrations dropped in between 100 nM and 1000 nM, the rate of increasing bound duplex was reduced. A similar effect was observed under certain conditions when RdgC was added after RecA protein was bound to ssDNA in the absence of SSB. The formation of RecA filaments was presumably

interrupted by regions of secondary structure in the ssDNA, which could be used by RdgC to inhibit RecA activities. This inhibitory effect was evident when RdgC was at low concentrations (0.2 to 0.4 μM) or high concentrations (4 to 16 μM). However, this effect was greatly lessened when the RdgC concentration was at 1 μM. Based on the data, Drees and colleagues suggested that RdgC has high affinity for dsDNA as monomer or tetramer, and this is when RdgC limits RecA activities by binding tightly to dsDNA; when RdgC is present as dimers, the inhibition is limited due to its weakened dsDNA binding (Drees et al., 2006).

1.5.3 RdgC *in vivo*

The idea that RdgC negatively regulates RecA activities by binding to DNA is becoming widely accepted. However, this role is not obvious in an *rdgC* single mutant. Strains deleted for *rdgC* do not show any evident defects in either cell growth or resistance to DNA damages (Ryder et al., 1996). Its *in vivo* interaction with recombination can only be investigated when normal recombinations of replication restart are compromised.

As previously mentioned, cells lacking *priA* are very sick (Figure 1-3 N5972). Since the PriA protein plays a crucial role in DNA replication restart after replication forks abnormally collapse, this suggests such event commonly occurs even under normal growth conditions. Certain *dnaC* mutants such as *dnaC810*, *dnaC212*, and *dnaC809,820* can suppress the loss of *priA* by actively using the PriC-Rep pathway for replication restart. However, all the viabilities more or less require the presence of RdgC, as discussed previously. For instance, *priA*$^-$ *rdgC*$^-$ and *priA*$^-$ *dnaC810 rdgC*$^-$ strains are dead (Moore et al.,

2003); *priA⁻ dnaC212 rdgC⁻* strains are very sick and readily develop suppressors (Moore et al., 2003); and *priA⁻ dnaC809,820 rdgC⁻* has to rely on PriC for cells growth (Figure 1-3; R. G. Lloyd, unpublished data). The toxic effects of the RecFOR mediated loading of RecA have been confirmed at least in two of the situations: *priA⁻ dnaC810 rdgC⁻* (Figure 1-3; R. G. Lloyd, unpublished data) and *priA⁻ dnaC212 rdgC⁻* (Moore et al., 2003). In addition to *recF*, *recO* and *recR*, the suppressors for *priA⁻ dnaC212 rdgC⁻* also included mutations on *ssb* and *rpoB*. The C-terminus of SSB is known to interact with many proteins, including RecO (Hobbs et al., 2007) and PriA (Cadman and McGlynn, 2004). Some SSB C-terminus mutants have been reported to perform much of the wild type function, but are resistant to RecFOR mediated loading of RecA. *rpoB* encodes the β subunit of RNA polymerase, which attaches to DNA during transcription. This process is much slower than replication, potentially causing pathogenic encounter between transcription and replication. The mutations on *rpoB* presumably weaken the interaction between the RNA polymerase and the transcribing DNA, thus reducing the chance of halting replication forks (Moore, 2002).

The interaction between RdgC and RecA is rather perplexing in strains lacking RecBCD and SbcBC. SbcB and SbcC digest exposed DNA ends and attack hairpin structures (Connelly and Leach, 1996), and is problematic when RecBCD is not available to process and protect DNA ends. The cells lacking both nuclease systems are as viable as the wild type. However, if the strain is further deprived of *rdgC*, the cell's viability relies on the RecFOR pathway to

load RecA (Ryder et al., 1996). This indicates that RdgC acts in a pathway parallel to RecFOR in DNA maintenance.

To sum up, RdgC is a DNA binding protein, without any preference for specific structures or sequences, although it binds dsDNA slightly better than ssDNA. This DNA binding protein is dispensable for cell growth under normal conditions, but plays a vital role to antagonise RecA toxicity when replication restart systems are compromised. The question is: how does RdgC achieve these effects? I set out to try and answer this question by studying the structure of the protein and its mode of DNA binding. For this purpose, I purified the RdgC protein in large scale, and crystallised it with a variety of DNA substrates. Its crystal structure was resolved to 2.4 Å. Several *rdgC* mutants were constructed based on the structural information; and their *in vivo* synthetic lethality assays were combined with their gene products *in vitro* DNA binding analyses to elevate our knowledge to a new level on RdgC's role in DNA recombinational repair.

Chapter 2
Materials and methods

2.1 Materials

2.1.1 Buffers and solutions

All buffers and solutions are given as working concentrations. The concentration of the stock solution, when made, is detailed in parenthesis. After manufacture, buffers and solutions were sterilised either by autoclaving at 121 ºC for 15 minutes, or by filtering through a Whatman 0.45 μm filter where necessary.

All buffers used in this work are listed below:

TBE – 90 mM Tris-borate, 1 mM EDTA (stock solution, 10×)

GBB (Gel Binding Buffer) – 50 mM Tris·HCl pH8.0, 100 μg/ml BSA, 2 mM $MgCl_2$, 6% (v/v) glycerol, 1 mM DTT (stock solution, 5×)

LIS (Low Ionic Strength) – 6.7 mM Tris·HCl pH 8.0, 3.3 mM sodium acetate, 2 mM $MgCl_2$ (stock solution, 20×)

MC buffer – 100 mM $MgSO_4$, 5 mM $CaCl_2$

SSC – 150 mM $MgCl_2$, 15 mM sodium citrate, pH 7.0 (stock solution, 20×)

SDS PAGE running buffer – 0.1% (w/v) SDS, 1.44% (w/v) glycine, 0.3% (w/v) Trizma base (stock solution, 10×)

SDS PAGE loading buffer – 50 mM Tris·HCl pH 6.8, 100 mM DTT, 2% (w/v) SDS, 0.1% (w/v) bromphenol blue, 10% (v/v) glycerol (stock solution, 5×)

Disruption buffer – 50 mM Tris·HCl pH7.5, 1 mM EDTA, 1 mM DTT

Buffer A – 50 mM Tris·HCl pH 7.5, 1 mM EDTA, 1 mM DTT (10 mM DTT for purification of Seleno-methionine derivatives)

Buffer B – 20 mM Tris·HCl pH 7.5, 0.1 M NaCl (extra 2 mM DTT for purification of Seleno-methionine derivatives)

2.1.2 Microbial growth media and supplements

Yeast extract, tryptone and Bacto-agar were all obtained from Difco. Liquid and solid media used for the growth of *E. coli* were prepared according to the following standard recipes, and then sterilised by autoclaving at 121 ºC, for 15 minutes. Media were supplemented with antibiotics as required.

2.1.2.1 LB media

LB (Luria-Burrous) broth contained 5 g/l yeast extract, 10 g/l tryptone, 0.5 g/l NaCl and 0.08 g/l NaOH made up in distilled water to a final pH 7.5. LB agar plates were supplemented with 15 g/l of Bacto-agar.

2.1.2.2 Mu media

Mu broth contained 5 g/l yeast extract, 10 g/l tryptone and 10 g/l NaCl made with distilled water to a final pH 7.5. Mu agar plates were supplemented with 10 g/l Bacto-agar. Mu overlays were supplemented with either 4 or 6 g/l Bacto-agar.

2.1.2.3 Minimal media

Minimal 56/2 salts media (Willetts et al., 1969) contained 2.64 g/l KH_2PO_4, 4.3 g/l Na_2HPO_4, 0.1 g/l $MgSO_4 \cdot 7H_2O$, 0.1 g/l $(NH_4)_2SO_4$, 0.005 g/l $Ca(NO_3)_2$ and 0.0025 g/l $FeSO_4 \cdot 7H_2O$ in distilled water. For minimal salts solid media, 56/2 salts were used at double strength and Bacto-agar added to 15 g/l. 56/2 salts media was further supplemented with thiamine (1 μg/ml), a carbon source (glucose (3.3 mg/ml) or maltose (3.3 mg/ml)) and any specific requirements as described below.

2.1.2.4 Antibiotic solutions

Antibiotic stocks were made in sterile distilled water unless otherwise stated. Stock solutions of the following were kept at 4 ºC: chloramphenicol (2 mg/ml), ampicillin (4 mg/ml), carbenicillin (4 mg/ml), kanamycin (4 mg/ml) and streptomycin (20 mg/ml). 3 ml aliquots of tetracycline (2 mg/ml) were stored at -20 ºC. Trimethoprim (1 mg/ml) and rifampicin (10 mg/ml) were dissolved in methanol and stored at -20 ºC. Antibiotics were used at the following working concentrations:

Chloramphenicol	20 μg/ml
Ampicillin	40 μg/ml
Carbenicillin	40 μg/ml
Kanamycin	25 μg/ml
Streptomycin	100 μg/ml

Tetracycline	20 μg/ml
Trimethoprim	20 μg/ml
Rifampicin	10-100 μg/ml as indicated (e.g. Rif = 50 μg/ml)

The antibiotics could be used on rich or minimal plates, with the exception of trimethoprim which is only selective on minimal plates.

2.1.3 Strains, bacteriophages and plasmid used

The *E. coli* strains used in this study are listed in Table 2.1, and plasmids in Table 2.2. Bacteriophage P1 *vir* was used in this study for transductions (Miller, 1972). If derived in this work, the shorthand under source or reference in Table 2.1 indicates the construction parameters, including the recipient strain, P1 donor strain, type of selection and any other comments.

Δ*rdgC*::*dhfr* or Δ*rdgC2*::*dhfr* was introduced to replace the wild type *rdgC* in the chromosome. The replacing *dhfr* encodes resistance to trimethoprim, which was used to select *rdgC*$^-$ cells. $rdgC_{mutant}$-*cat* was consequently engineered to replace Δ*rdgC*::*dhfr* or Δ*rdgC2*::*dhfr*. The *cat* gene immediately downstream of the linked alleles encodes resistance to chloramphenicol, which was used to select $rdgC_{mutant}$ cells. Details of the method can be found in section 2.2.9.

Table 2.1 Strains used in this study

Strain	Relevant genotype	Source or reference
a) MG1655 derivatives		
MG1655	wild type[1]	(Bachmann, 1996)
N4586	*ΔrdgC*::*dhfr*	P1·JP947 × MG1655 to Tm^r
N5286	*ΔxonA300*::*cat*	P1·STL2694 × MG1655 to Cm^r

N5500	*priA300*	(Mahdi et al., 2006)
N5521	*priA300 dnaC810 zji-202*::Tn*10*	P1·DIM167 × N5500 to Tc^r
N5539	*dnaC810 zji-202*::Tn*10 priA2*::*kan*	P1·PN105 × N5521 to Km^r
N5935	*ΔpriA*::*apra*	Original recombineering strain
N5949	*ΔpriA*::*apra dnaC*(unknown)	UV^r derivative of N5935
YJ012	*ΔrdgC*::*dhfr ΔxonA300*::*cat*	P1· N5286 × N4586 to Cm^r
YJ015	*ΔrdgC*::*dhfr ΔxonA300*::*cat* pT7pol26	pT7pol26 × YJ012 to Km^r

b) AB1157 derivatives

AB1157[2]		(Bachmann, 1996)
JP947	*ΔrdgC*::*dhfr*	A gift from Joe Peters

c) TB28, F^- (plasmid-free) derivatives

TB28	Δ*lacIZYA*	(Bernhardt and de Boer, 2004)
AM1833	$rdgC_{R118C}$-*cat*	This work[3]
AM1834	$rdgC_{P76G,L116T,\Delta(77-115)}$-*cat*	This work
AM1835	$rdgC_{P76G,L116T,\Delta(77-115),R118C}$-*cat*	This work
AM1836	$rdgC_{H222A}$-*cat*	This work
AM1837	$rdgC_{K227A}$-*cat*	This work
AM1838	$rdgC_{R118A}$-*cat*	This work
AM1839	$rdgC_{K211A}$-*cat*	This work
AM1840	$rdgC_{E218R}$-*cat*	This work
AM1887	$rdgC_{wt}$-*cat*	This work
AM1905	$rdgC_{F120T}$-*cat*	This work
AM1917	$rdgC_{F120S}$-*cat*	This work
AM1929	*ΔrdgC2*::*dhfr*[4]	This work
AM1931	*N.rdgC-cat*	This work
AM1950	$rdgC_{R97S,K98Q,K100Q,K101E}$-*cat*	This work
AM2029	$rdgC_{Q212A}$-*cat*	This work
YJ025	$rdgC_{R118C,F120T}$-*cat*	This work

d) TB28, pAM374 (pRC7, $priA^+$) derivatives

N5936	Δ*lacIZYA*	(Mahdi et al., 2006)
N5972	Δ*priA*::*apra*	P1·N5949 × N5936 to $Apra^r$
N6021	Δ*priA*::*apra dnaC810 zji-202*::Tn*10*	P1·N5539 × N5972 to Tc^r
N6040	Δ*priA*::*apra ΔrdgC*::*dhfr*	P1·N4586 × N5972 to Tm^r
N6111	Δ*priA*::*apra dnaC810 zji-202*::Tn*10* Δ*rdgC*::*dhfr*	P1·N4586 × N6021 to Tm^r
N6131	*priA2*::*kan dnaC810 zji-202*::Tn*10*	P1·PN105 × N6021 to Km^r
N6143	*priA2*::*kan dnaC810 zji-202*::Tn*10* *ΔrdgC*::*dhfr*	P1·N4586 × N6131 to Tm^r
N6157	*priA2*::*kan dnaC810 zji-202*::Tn*10* *ΔrdgC*::*dhfr ΔrecA*::*apra*	P1·N6124 × N6143 to $Apra^r$
N6445	Δ*priA*::*apra dnaC809,820 zji-202*::Tn*10*	P1·N6424 × N6021 to Tc^r
N6457	Δ*priA*::*apra dnaC809,820 zji-202*::Tn*10*	P1·N4586 × N6445 to Tm^r

	ΔrdgC::*dhfr*	
N6471	Δ*priA*::*apra dnaC809,820 zji-202*::Tn*10 priC303*::*kan*	P1·N6424 × N6445 to Km^r
N6538	Δ*priA*::*apra dnaC809,820 zji-202*::Tn*10 ΔrdgC*::*dhfr priC303*::*kan*	P1·N6424 × N6457 to Km^r
AM1843	Δ*priA*::*apra* $rdgC_{R118C}$-*cat*	P1·AM1833 × N5972 to Cm^r
AM1844	Δ*priA*::*apra* $rdgC_{P76G,L116T,\Delta(77-115)}$-*cat*	P1·AM1834 × N5972 to Cm^r
AM1845	Δ*priA*::*apra* $rdgC_{P76G,L116T,\Delta(77-115),R118C}$-*cat*	P1·AM1835 × N5972 to Cm^r
AM1846	Δ*priA*::*apra* $rdgC_{H222A}$-*cat*	P1·AM1836 × N5972 to Cm^r
AM1847	Δ*priA*::*apra* $rdgC_{K227A}$-*cat*	P1·AM1837 × N5972 to Cm^r
AM1848	Δ*priA*::*apra* $rdgC_{R118A}$-*cat*	P1·AM1838 × N5972 to Cm^r
AM1850	Δ*priA*::*apra* $rdgC_{E218R}$-*cat*	P1·AM1840 × N5972 to Cm^r
AM1853	Δ*priA*::*apra dnaC810 zji-202*::Tn*10* $rdgC_{R118C}$-*cat*	P1·AM1833 × N6021 to Cm^r
AM1854	Δ*priA*::*apra dnaC810 zji-202*::Tn*10* $rdgC_{P76G,L116T,\Delta(77-115)}$-*cat*	P1·AM1834 × N6021 to Cm^r
AM1855	Δ*priA*::*apra dnaC810 zji-202*::Tn*10* $rdgC_{P76G,L116T,\Delta(77-115),R118C}$-*cat*	P1·AM1835 × N6021 to Cm^r
AM1856	Δ*priA*::*apra dnaC810 zji-202*::Tn*10* $rdgC_{H222A}$-*cat*	P1·AM1836 × N6021 to Cm^r
AM1857	Δ*priA*::*apra dnaC810 zji-202*::Tn*10* $rdgC_{K227A}$-*cat*	P1·AM1837 × N6021 to Cm^r
AM1858	Δ*priA*::*apra dnaC810 zji-202*::Tn*10* $rdgC_{R118A}$-*cat*	P1·AM1838 × N6021 to Cm^r
AM1860	Δ*priA*::*apra dnaC810 zji-202*::Tn*10* $rdgC_{E218R}$-*cat*	P1·AM1840 × N6021 to Cm^r
AM1911	Δ*priA*::*apra* $rdgC_{F120T}$-*cat*	P1·AM1905 × N5972 to Cm^r
AM1912	Δ*priA*::*apra dnaC810 zji-202*::Tn*10* $rdgC_{F120T}$-*cat*	P1·AM1905 × N6021 to Cm^r
AM1913	Δ*priA*::*apra N.rdgC-cat*	P1·AM1931 × N5972 to Cm^r
AM1914	Δ*priA*::*apra dnaC810 zji-202*::Tn*10 N.rdgC-cat*	P1·AM1931 × N6021 to Cm^r
AM1918	Δ*priA*::*apra ΔrdgC2*::*dhfr*	P1·AM1929 × N5972 to Tm^r
AM1919	Δ*priA*::*apra dnaC810 zji-202*::Tn*10 ΔrdgC2*::*dhfr*	P1·AM1929 × N6021 to Tm^r
AM1935	Δ*priA*::*apra* $rdgC_{K211A}$-*cat*	P1·AM1839 × AM1918 to Cm^r
AM1936	Δ*priA*::*apra dnaC810 zji-202*::Tn*10* $rdgC_{K211A}$-*cat*	P1·AM1839 × AM1919 to Cm^r
AM1948	Δ*priA*::*apra* $rdgC_{R97S,K98Q,K100Q,K101E}$-*cat*	P1·AM1950 × AM1918 to Cm^r
AM1949	Δ*priA*::*apra dnaC810 zji-202*::Tn*10* $rdgC_{R97S,K98Q,K100Q,K101E}$-*cat*	P1·AM1950 × AM1919 to Cm^r
AM1963	Δ*priA*::*apra* $rdgC_{F120S}$-*cat*	P1·AM1917 × AM1918 to Cm^r
AM1964	Δ*priA*::*apra dnaC810 zji-202*::Tn*10* $rdgC_{F120S}$-*cat*	P1·AM1917 × AM1919 to Cm^r
AM2046	Δ*priA*::*apra dnaC809,820 zji-202*::Tn*10 priC303*::*kan* $rdgC_{wt}$-*cat*	P1·AM1887 × YJ026 to Cm^r
AM2048	Δ*priA*::*apra dnaC809,820 zji-202*::Tn*10 priC303*::*kan* $rdgC_{K227A}$-*cat*	P1·AM1837 × YJ026 to Cm^r

AM2049	Δ*priA*::*apra dnaC809,820 zji-202*::Tn*10 priC303*::*kan rdgC*$_{R118A}$-*cat*	P1·AM1838 × YJ026 to Cm^r
YJ026	Δ*priA*::*apra dnaC809,820 zji-202*::Tn*10 priC303*::*kan ΔrdgC2*::*dhfr*	P1·AM1929 × N6471 to Cm^r
YJ027	Δ*priA*::*apra dnaC809,820 zji-202*::Tn*10 priC303*::*kan rdgC*$_{K211A}$-*cat*	P1·AM1839 × YJ026 to Cm^r
YJ028	Δ*priA*::*apra dnaC809,820 zji-202*::Tn*10 priC303*::*kan rdgC*$_{Q212A}$-*cat*	P1·AM2029 × YJ026 to Cm^r
YJ029	Δ*priA*::*apra dnaC809,820 zji-202*::Tn*10 priC303*::*kan rdgC*$_{E218R}$-*cat*	P1·AM1840 × YJ026 to Cm^r
YJ030	Δ*priA*::*apra dnaC809,820 zji-202*::Tn*10 priC303*::*kan rdgC*$_{H222A}$-*cat*	P1·AM1836 × YJ026 to Cm^r
YJ033	Δ*priA*::*apra dnaC809,820 zji-202*::Tn*10 priC303*::*kan rdgC*$_{R118C}$-*cat*	P1·AM1833 × YJ026 to Cm^r
YJ034	Δ*priA*::*apra dnaC809,820 zji-202*::Tn*10 priC303*::*kan rdgC*$_{P76G,L116T,Δ(77-115)}$-*cat*	P1·AM1834 × YJ026 to Cm^r
YJ035	Δ*priA*::*apra dnaC809,820 zji-202*::Tn*10 priC303*::*kan rdgC*$_{P76G,L116T,Δ(77-115),R118C}$-*cat*	P1·AM1835 × YJ026 to Cm^r
YJ036	Δ*priA*::*apra dnaC809,820 zji-202*::Tn*10 priC303*::*kan rdgC*$_{F120T}$-*cat*	P1·AM1905 × YJ026 to Cm^r
YJ037	Δ*priA*::*apra dnaC809,820 zji-202*::Tn*10 priC303*::*kan rdgC*$_{F120S}$-*cat*	P1·AM1917 × YJ026 to Cm^r
YJ038	Δ*priA*::*apra dnaC809,820 zji-202*::Tn*10 priC303*::*kan rdgC*$_{R97S,K98Q,K100Q,K101E}$-*cat*	P1·AM1950 × YJ026 to Cm^r
YJ039	Δ*priA*::*apra dnaC809,820 zji-202*::Tn*10 priC303*::*kan rdgC*$_{R118C,F120T}$-*cat*	P1·YJ025 × YJ026 to Cm^r
YJ040	Δ*priA*::*apra dnaC809,820 zji-202*::Tn*10 priC303*::*kan N.rdgC-cat*[5]	P1·AM1931 × YJ026 to Cm^r
YJ041	Δ*priA*::*apra rdgC*$_{Q212A}$-*cat*	P1·AM2029 × AM1918 to Cm^r
YJ043	Δ*priA*::*apra rdgC*$_{R118C,F120T}$-*cat*	P1·YJ025 × AM1918 to Cm^r
YJ044	Δ*priA*::*apra dnaC810 zji-202*::Tn*10 rdgC*$_{Q212A}$-*cat*	P1·AM2029 × AM1919 to Cm^r
YJ046	Δ*priA*::*apra dnaC810 zji-202*::Tn*10 rdgC*$_{R118C,F120T}$-*cat*	P1·YJ025 × AM1919 to Cm^r
YJ047	Δ*priA*::*apra rdgC*$_{wt}$-*cat*	P1·AM1887 × AM1918 to Cm^r
YJ048	Δ*priA*::*apra dnaC810 zji-202*::Tn*10 rdgC*$_{wt}$-*cat*	P1·AM1887 × AM1919 to Cm^r

e) BL21 (DE3) derivatives

BL21 (DE3)	F$^-$ *ompT hsdS*$_B$ (*r*$_B^-$ *m*$_B^-$) *gal dcm* (DE3) pLysS	(Studier and Moffatt, 1986)
YJ014	*ΔrdgC*::*dhfr*	P1·N4586 × BL21 (DE3) to Tm^r

f) Other strains

DIM167	DM4000[6] *priA2*::*kan dnaC810 zji-202*::Tn*10*	(Moore, 2002)
N5946	DM4000 *priC303*::*kan priB202 dnaC809,820*	A gift from Steve Sandler
N6419	*dnaC809,820 zji-202*::Tn*10*	T. Moore, unpublished

N6424	*priC303*::*kan priB202 dnaC809,820 zji-202*::Tn*10*	P1·N6419 × N5946 to Tc^r
PN105	*priA2*::*kan sulA*	A gift from Ken Marians
STL2669	(Δ*recA-srlR*)306::*Tn*10, *xonA2*	A gift from Susan Lovett
STL2694	*ΔxonA300*::*cat thr-1 leuB6 proA2 supE44 kdg51 rfbD1 araC14 lacY1 galK2 xyl-5 mtl-1 tsx-33 rpsL31 rac*$^-$	(Viswanathan and Lovett, 1998)
W3110	wild type	(Bachmann, 1996)

1 F^- λ^- *ilvG*$^-$ *rfb-50 rph-1*. MG1655 is the 'wild type' K-12 strain used in this study

2 F- λ- rac- *thi-1 hisG4* Δ*(gpt-proA)62 argE3 thr-1 leuB6 kdgK51 rfbD1 araC14 lacY1 galK2 xylA5 mtl-1 tsx-33 supE44(glnV44) rpsL31 (strR)*

3 From recombineering of *rdgC* in TB28 pKD46 (see methods section 2.2.9)

4 Δ*rdgC2*::*dhfr* is total deletion of *rdgC*, whilst Δ*rdgC*::*dhfr* is partial deletion of *rdgC*.

5 *rdgC* cloned from *Neisseria meningitidis*, a gift from Tom Baldwin, of the Meningococcal Virulence and Vaccine Development Group, Nottingham

6 *E. coli B F*$^-$ Δ*(pro-lac)*$_{XIII}$ *hisG4 argE3 thr-1 ara-14 xyl-5 mtl-1 rpsL31 sfiA*::*Mu-d*(Ap, *lac*, *B*::Tn*9*) (Sandler, 1996)

The plasmids used in this study were constructed from pRC7, pET22b and pT7-7 (Table 2.2).

Table 2.2 Plasmids used in this study

Name	Description	Source or reference
pRC7		(Bernhardt and de Boer, 2004)
pAM374	*priA*$^+$	(Mahdi et al., 2006)
pKD46	Temperature sensitive plasmid at 37 °C and encoding phage λ Red recombinase	(Datsenko and Wanner, 2000)
pT7-7		(Tabor and Richardson, 1985)
pGS853	*rdgC*$_{wt}$	(Moore et al., 2003)
pGB043	*rdgC*$_{P76G,L116T,Δ(77-115)}$	This work*
pGB045	*rdgC*$_{R118C}$	This work
pGB046	*rdgC* $_{P76G,L116T,Δ(77-115),R118C}$	This work
pGB049	*rdgC*$_{R118A}$	This work
pGB051	*rdgC*$_{F120S}$	This work
pGB053	*rdgC*$_{F120T}$	This work
pGB055	*rdgC*$_{R97S,K98Q,K100Q,K101E}$	This work

pYJ003	*rdgC*$_{H222A}$	This work
pET22b		Novagen
pYJ001	*rdgC*$_{wt}$	This work
pGB047	*rdgC*$_{R118C}$	*Nde*I-*Hind*III fragment from pGB045
pGB048	*rdgC* $_{P76G,L116T,\Delta(77-115),R118C}$	*Nde*I-*Hind*III fragment from pGB046
pGB050	*rdgC*$_{R118A}$	*Nde*I-*Hind*III fragment from pGB049
pGB052	*rdgC*$_{F120S}$	*Nde*I-*Hind*III fragment from pGB051
pGB054	*rdgC*$_{F120T}$	*Nde*I-*Hind*III fragment from pGB053
pGB056	*rdgC*$_{R97S,K98Q,K100Q,K101E}$	*Nde*I-*Hind*III fragment from pGB055
pGB065	*rdgC*$_{R119C,F120T}$	This work
pYJ007	*rdgC*$_{K227A}$	This work
pYJ006	*rdgC*$_{H222A}$	*Nde*I-*Hind*III fragment from pYJ003
pYJ013	*rdgC*$_{E218R}$	This work
pYJ014	*rdgC*$_{K211A}$	This work
pYJ017	*rdgC*$_{Q212A}$	This work
pLysS	Expresses T7 lysozyme, Cm^r	Stratagene
pT7Pol26	Expresses T7 polymerase, Km^r	(Mertens et al., 1995)

* Construction of these plasmids are detailed in section 2.2.7

2.1.4 Oligonucleotides

DNA oligonucleotides were synthesised by phosphoramidite chemistry and prepared by MWG. They were supplied lyophilised, and required no further purification when used to prime PCR. When used in EMSAs, they were purified on a sequencing gel beforehand. The oligonucleotides used in this study are listed below.

Table 2.3 Oligonucleotides

Name	Sequence 5’ – 3’	Length (nt)
YJ1	AATGTAATCGTCTATGACGTT	21
YJ2	AACGTCATAGACGATTACATT	21
YJ3	GTAATCGTCTATGACGTT	18
YJ4	AACGTCATAGACGATTAC	18
YJ5	ATCGTCTATGACGTT	15
YJ6	AACGTCATAGACGAT	15
YJ7	TGTAATCGTCTATGACGTT	19
YJ8	AACGTCATAGACGATTACA	19
YJ9	TAATCGTCTATGACGTT	17
YJ10	AACGTCATAGACGATTA	17
YJ11	TGTAATCGTCTATGACGTTTTTTAACGTCATAGACGATTACA	42
YJ12	TAATCGTCTATGACGTTTTTTAACGTCATAGACGATTA	38
YJ13	GCTTCAATGTGATTGGTGATGCGTTCGCTGGTCAGATCTTGTTTC	45
YJ14	GAAACAAGATCTGACCAGCGAACGCATCACCAATCACATTGAAGC	45
YJ17	CACTTTTCCGGCTTCAATGGCATTGGTGATCTCTTCGC	38
YJ18	GCGAAGAGATCACCAATGCCATTGAAGCCGGAAAAGTG	38
YJ19	GCCAGTTTAGTCACCACTGCTCCGGCTTCAATGTG	35
YJ20	CACATTGAAGCCGGAGCAGTGGTGACTAAACTGGC	35
YJ21	GCTGGTCAGATCTTGTGCCTTCGCGCGGATCAC	33
YJ22	GTGATCCGCGCGAAGGCACAAGATCTGACCAGC	33
YJ23	GAGAAGCTTACGTTGTGCTTCGCCACC	27
YJ24	GGAGATATACATATGCTGTGG	21
YJ25	GGAAGGGCATATGTGGTTCAAAAATTTAATGACTTA	36
YJ26	CACAAGCTTATAATCTCGCTTTTTCGCCAC	30
YJ27	GTTAGCAGCCGGATCCTCAGTG	22
YJ28	CGTTATCATGCCCGCTAAATTTAACGACAAGGCCGTGGAAATTATCATGT GGTTCAAAAATTTAATG	67
YJ29	CTTCGCTGGTCAGATCTGCTTTCTTCGCGCGGATC	35
YJ30	GATCCGCGCGAAGAAAGCAGATCTGACCAGCGAAG	35
TMRC1	TGACCTCGTCCAGTTCACG	19

TMRC1R	CGTGAACTGGACGAGGTCA	19
TMRC1-MM1	TGACCTCGTACAGTTCACG	19
TMRC1-MM3	TGACCTCGAAAAGTTCACG	19
TMRC1R-D1	CGTGAACTGACGAGGTCA	18
TMRC2-DUMB	CAGTTCACGTTTTCGTGAACTGGACGAGGTCTTTTGACCTCGTC	44
TMRC3-DUMB	CCGGTCACGTTTTCGTGACCGGGCCGAGGTCTTTTGACCTCGGC	44
RGL13	GACGCTGCCGAATTCTGGCTTGCTAGGACATCTTTGCCCACGTTGACCC	49
RGL17	GGGTCAACGTGGGCAAAGATGTCCTAGCAAGCCAGAATTCGGCAGCGTC	49
GB-P29	GTTATCTGCGCGCGCAAAGAAGAAAAAATCCTCGGTACCGGC	42
GB-P30	GCCGGTACCCCGCGTGCTTTCAGCCGTTTTAGCCAG	36
GB-P31	CAAATTGTTATCTGCGCGCGCAAAGAAGAAAAAGGTACCCCG	42
GB-P32	GCCGGTACCTTCAGCCGTTTTAGCCAGACAATGATGTGGATCG	43
GB-P34	CGGCTGAAAGCACACGGCAGCAGAGA	26
GB-P35	CGGCTGAAAGCACACGGGGTACCGAG	26
GB-P37	TTTTAGCCAGACAATGATGTGGATCGACACGG	32
GB-P41	CGGCAGCAGAGAGTGCAGCACTTCGTCTTTCAGC	34
GB-P42	GCGGCTTTCAGCCGTTTTAGCCAGACAATGATGTGGATCG	40
GB-P45	CTAGCAGCCGTTTTAGCCAGACAATGATGTGGATCG	36
GB-P46	CACGCGGCAGCAGAGAGTGCAGCACTTCG	29
GB-P47	GTCCCGTTTTAGCCAGACAATGATGTGGATCGACACGG	38
GB-P48	GTCGCACGCGGCAGCAGAGAGTGCAGC	27
GB-P51	CGATCTGCAGCTGACTCGCCTGTTCC	26
GB-P52	CGGCCTGCAGGAAACCGAAAAAGATTCGC	29
GB-P77	CGACTCACTATAGGGAGACCACAACGG	27
GB-P78	GGGACGTCGCACACGGCAGCAGAGAGTGC	29
RdgC/Nm-5’	CGTTATCATGCCCGCTAAATTTAACGACAAGGCCGTGGAAATTATCATGTGGTTCAAGCAGATTAG	66
Pro/rdgC-5’	ATCATGCCCGCTAAATTTAACGACAAGGCCGTGGAAATTATCATGCTGTGGTTCAAAAATTTA	63
Tm/rdgC-5’	ATCATGCCCGCTAAATTTAACGACAAGGCCGTGGAAATTATCTGCAAGCAGGATAGACGGC	61
Tm/rdgC-3’	ACGCGGCAGGCGTCGCATCCGGCATTAAAGGAAAATCAGCAATCAGCATCCAATGTTTCCG	61
RdgC/Cm-3’	GCCTGGTGTGGCTTCGTACGCCGGATAAGACGCGGCAGGCGTCGATCTCAAGAAGATCATCTT	63

2.2 Methods

2.2.1 Growth of bacterial strains

2.2.1.1 Solid media

Agar plates were dried for 10-15 minutes at 70 ºC in a Leec drying cabinet prior to use. To obtain single colonies, inocula of a strain were streaked on the surface of a plate with a Tungsten loop, which was sterilised by heating to incandescence in a Bunsen flame. For precise amounts, samples from liquid cultures were pipetted directly onto the surface. For analysing a large number of strains simultaneously, cells were inoculated in regular arrays with a platinum tip (gridded). After 8-12 hours growth, the grids were replicated onto suitable test plates by adherence to a sterile velvet. Plates were incubated in a Leec incubator at 37 ºC, or as indicated.

2.2.1.2 Liquid media

Primary overnight cultures for routine use were prepared by suspending a single colony in 5 ml of LB broth in 15 ml screw capped tubes. Cultures were typically grown at 37 ºC overnight in a Leec incubator, with gentle rotation. Strains were stored like this for up to two weeks at 4 ºC. For long term storage, 2.5 ml fresh overnight culture was mixed with 1.5 ml 80% glycerol and stored at -20 ºC.

2.2.2 Manipulation of bacterial strains

2.2.2.1 Transfer of mutations by P1 transduction

Phage P1 *vir* stocks were grown on the appropriate donor strain and used to transduce the selected marker into the desired recipient strains. Phage stocks were made by adding 10^7-10^8 plaque forming units (pfu) from a wild type strain (usually P1·W3110) into pre-incubated host strains in 8 ml Mu broth with A_{650} of 0.3 to 0.4. Host strains were treated with 0.1 ml 0.5 M $CaCl_2$ for 5 minutes prior to the addition of the phage. The cultures were incubated at 37 ºC with vigorous shaking until lysis was appropriate (denoted by slight clearing and flocculation). The remnant cells were treated by adding 0.5 ml chloroform. The lysates were then centrifuged at 10,000 rpm for 15-20 minutes at 4 ºC in 15 ml Corex glass tubes. The supernatant having the phage was transferred into sterile culture tubes with 0.5 ml chloroform, and stored at 4 ºC. 8 ml experimental cultures of recipient strains were grown by diluting overnight cultures approximately 20 fold into fresh Mu broth supplemented with antibiotics needed and incubated at 37 ºC in a shaking water bath (Grants Instruments) with vigorous aeration to an A_{650} of 0.6. The cells were then pelleted by centrifugation and resuspended in 2 ml MC buffer (100 mM $MgSO_4$, 5 mM $CaCl_2$). 0.2 ml samples of the cells were then mixed with 50-200 µl of phage P1 *vir* grown on the appropriate donor strains and the mixture incubated for 20-25 minutes at 37 ºC before adding 0.2 ml 1 M NaCitrate to prevent further phage infection. For transduction of antibiotic resistance markers, the transduced cells were mixed with 2.5 ml molten 0.6% overlay agar and plated directly on suitable selection plates. The plates were typically

incubated at 37 ºC for 24-72 hours. For linked markers, the transductants were gridded onto LB plates as described, and the resulting array was screened for segregation of the relevant genotype. Transductants were then purified on selective plates and used to make overnight cultures. The genotype of the transduced strain was verified by diagnostic plate tests and by PCR mediated sequencing where necessary.

2.2.2.2 Transformation of plasmids

Heat shock transformation

E. coli cells to be transformed with plasmid DNA were grown to an A_{650} of 0.6 in 8 ml of LB broth. Cells were harvested by centrifugation in an SS34 rotor at 6000 rpm, 4 ºC for 10 minutes and resuspended in 1 ml 0.1 M $CaCl_2$ on ice. 200 µl of these cells were added to plasmid DNA as required and incubated on ice for 20 minutes. The mix was then heat shocked in a 42 ºC water bath for 2 minutes and returned immediately to ice. 1 ml of LB buffer was added and the cells incubated at 37 ºC for 1 hour. The cells were harvested by high speed centrifugation in a benchtop centrifuge, resuspended in 100 µl LB broth and spread onto LB agar plates supplemented with appropriate antibiotic selection. Plates were then incubated at 37 ºC for overnight.

Electroporation

An overnight culture of the parental strain was diluted 1:100 in 50 ml of LB broth with an appropriate amount of ampicillin and IPTG. Cells were grown to an A_{650} of 0.8 and harvested after 15 minutes incubation on ice. The pellet was resuspended in ice-cold 1 mM HEPES buffer then centrifuged. This step was

repeated twice and the final pellet was resuspended in 1 mM HEPES buffer with 10% glycerol. Cell suspension was stored at -70 °C as 100 µl aliquots.

1 µl of plasmid DNA was mixed with 40 µl of electrocompetent cells in an electroporation cuvette. Cells were pulsed at 1.8 kV and immediately recovered by adding 1 ml of SOC medium after electroporation. The mixture was transferred to a tube and recovered in a shaking water bath at 37 °C for 1 hour. Cells were collected and plated on appropriate selective plates. Next day colonies were collected in LB broth and mixed with 3 ml 80% glycerol per 5 ml of cell suspension and stored at -20 °C.

2.2.3 Assays on *E. coli* strains

2.2.3.1 Checking strain genotypes

It was often necessary to verify strain genotypes after transduction. When screening large numbers of candidates, e.g. moving linked genes, colonies were inoculated into regular arrays on a master plate, then replica plated using sterile velvets onto appropriate selection media. Patches displaying the required phenotype were purified from the master plate onto LB agar, grown as an overnight culture and re-tested. If fewer candidates required testing, e.g. after the inheritance of genes containing an antibiotic resistance marker, 2-10 µl of overnight culture was tested directly by streaking on the appropriate selection media.

2.2.3.2 Measuring sensitivity to DNA damage

Semi-quantitative analysis of DNA damage sensitivity was achieved by streaking 10 µl of overnight culture onto LB agar supplemented with and without MC at 0.2 µg/ml or 0.5 µg/ml. A duplicate set was UV irradiated at a dose rate of 1 J/m^2/sec, at a peak output of 245 nm, for 30 sec (MC containing plates) and 60 sec (LB agar only plates). Sensitivity to MC and UV was scored after overnight incubation, at 37 ºC, by comparing the growth of test strains to either a wild type, or parent, control strain.

2.2.4 Preparation and analysis of DNA

2.2.4.1 Extraction of plasmids

Plasmid DNA was purified from overnight cultures of strains using either a Qiagen mini or midi prep kit. Cells were harvested by centrifugation and resuspended in 50 mM Tris·HCl, pH 8.0, 10 mM EDTA, 100 µg/ml RNase A. Cell lysis was performed by adding an equal volume of lysis solution (200 mM NaOH, 1% SDS) and allowing the reaction to proceed at room temperature for 5 minutes. Addition of ~1.4 volumes of neutralisation buffer (3 M KAc, pH 5.5), followed by centrifugation, separated cell debris and chromosomal DNA from the plasmid DNA. For the mini prep purification, the plasmid DNA was absorbed to a silica gel QIAprep column, washed with a high salt buffer to remove nuclease activity and a buffer containing 70% ethanol and then eluted with 30-50 µl of TE buffer. For the midi prep purification, the plasmid DNA was absorbed to a QIAgen tip column, equilibrated with 750 mM NaCl, 50 mM MOPS, pH 7.0, 15% ethanol, 0.15% Triton X-100, prior to loading DNA

extract. The DNA was washed with a similar buffer, but containing 1M NaCl and eluted with 1.25 M NaCl, 50 mM Tris-HCl, pH 8.5, 15% ethanol. The DNA was precipitated with 0.7 volumes isopropanol and pelleted by centrifugation. The pellet was washed with 70% ethanol and the DNA resuspended in 70 - 200 µl of TE buffer.

2.2.4.2 Restriction enzyme digestion

Restriction endonuclease digests were routinely conducted in 20 µl reactions containing 0.2 -1 µg DNA in the appropriate buffer with 1 unit of restriction enzyme. Digests were incubated at the required temperature (as recommended by the supplier), for 2 hours or more.

2.2.4.3 Cloning sequences into plasmid vectors

Appropriately digested, and purified DNA fragments, were ligated using T4 DNA ligase according to the manufacturer's recommendations. Vector DNA was dephosphorylated using Antarctic phosphatase (New England Biolabs) in accordance with the manufacturer's recommendations. Fragment concentrations were quantified using a Beckman Coulter DU 530 Spectrophotometer, such that the ligation reaction contained a three-fold molar excess of insert DNA ends to vector ends. Reactions were completed in a final volume of 10 µl for 15 min at room temperature. For improved efficiency, both cohesive-end and blunt-end ligation reactions were performed overnight at 4°C. Ligation reactions were transformed into competent DH5α, and transformants obtained under the relevant selection.

2.2.4.4 PCR amplification of DNA

Amplification of DNA followed an adapted procedure previously described (Mullis and Faloona, 1987; Saiki et al., 1988). Pairs of 5' and 3' primers, the required DNA template, dNTPs and Taq, Dynazyme or Phusion polymerase (New England Biolabs) were used with the appropriate buffers provided. PCR products were analysed by gel electrophoresis and purified by isolating the product from an agarose gel and purifying using the Qiagen Gel Extraction kit.

2.2.4.5 Labelled DNA substrates

End labelling

Oligonucleotides were provided without a 5' terminal phosphate group, and were labelled by attaching a [^{32}P] phosphate to the 5' terminus. The reaction mixture consisted of 100-500 ng DNA, 1 µl T4 Kinase, 1-2 µl [γ^{32}P]-ATP and a final concentration of 1× forward reaction buffer (70 mM Tris·HCl, pH 7.6, 10 mM $MgCl_2$, 100 mM KCl, 1 mM mercaptoethanol) in a total volume of 20 µl. The reaction was incubated for 1 hour at 37 °C, and halted by incubating at 65 °C for 15 minutes. The unincorporated, labelled ATP was removed by using Biospin P6 spin columns (Biorad).

Annealing

Oligonucleotides were annealed in 20 µl reactions, in a screw capped eppendorf, containing 0.2 µM of each oligo and 1× SSC. The mixture was placed in a hot block at 100 °C for 5 min, after which it was switched off and allowed to cool to room temperature. Salt was removed from the reaction using

a BIO-RAD P6 chromatography column, with buffer exchanged for distilled water. Oligonucleotides obtained from MWG-Biotech AG were phosphorylated after annealing in a 50 µl reaction containing 10 units T4 polynucleotide kinase (PNK), 1× T4 ligase buffer incubated at 37 °C for 1 hour. Reactions were heat inactivated at 65 °C for 20 min. Oligonucleotides were purified from the reaction using a QIAquick nucleotide removal kit (QIAGEN), in accordance with manufacturer's specifications, and eluted into 30 µl EB (10 mM Tris·HCl pH 8.5).

2.2.4.6 Gel electrophoresis

Agarose gels

DNA was typically resolved on a 1% (w/v) TBE agarose gel. Before loading, DNA was mixed with approximately 0.2 volumes Ficoll gel loading buffer (or 0.2 volumes of 80% glycerol for quantification). 0.2 µg/ml ethidium bromide was added to gels to allow visualisation (or gels were stained in TBE buffer added with 2.5‰ SYBR Green (Invitrogen) for quantification) of DNA fragments under UV light using a BIO-RAD Gel Doc EQ and Quantity One 1-D Analysis Software. Gels were run as standard at 90 V for approximately 1 hour. DNA fragments were sized using 0.25 µg 1 kb ladder (New England Biolabs).

Native polyacrylamide gel analysis

Radiolabelled DNA substrates were purified and analysed by native gel electrophoresis. Acrylamide gels contained polyacrylamide/bisacrylamide

(29:1) and were polymerised with 0.1% ammonium persulphate and 0.05% TEMED. Gels were subjected to electrophoresis in a Biorad Protean II apparatus in running buffer identical to the gel buffer. Gels were dried by placing onto Whatman 3MM filter paper and dried on a vacuum slab drier. The dried gels were exposed to X-ray film or a storage phosphor screen (Molecular Dynamics).

Sequencing gels

Single strand DNA oligos were purified on sequencing gels before application in *in vitro* DNA binding assays. Sequencing gels contained 12% polyacrylamide/bisacrylamide (29:1), 1× TBE and 46% (w/v) urea, and were polymerised with 0.1 ammonium persulphate and 0.05% TEMED. Electrophoresis performed at 2000 V for 3 hours. Gels were dried by placing onto Whatman 3MM filter paper and dried on a vacuum slab drier. Bands containing the appropriate DNA were visualised by UV shadowing and extracted of DNA by soaking the sliced gel containing the DNA with Elution Buffer (Qiagen). Purified DNA samples were quantified by absorbance at A_{260}.

2.2.4.7 Extraction of DNA from agarose gel

DNA products were extracted from gel segments using a QIAquick Gel Purification kit (in accordance with the manufacturer's specifications). DNA was eluted from the spin column and resuspended in 20 µl TE buffer, EB buffer or water.

2.2.5 Preparation and analysis of proteins

2.2.5.1 Protein overexpression

Standard procedures (apply for all proteins except for Seleno-methionine derivatives)

Overexpression of cloned genes in *E. coli* was tested on a small scale before large-scale purifications. Samples from fresh 5 ml overnight cultures, made from freshly transformed colonies, were used to inoculate 8 ml of Mu broth; the appropriate selection was maintained at all stages. The secondary cultures were grown to an A_{650} of 0.4 - 0.5 at 37 °C, and if selection was maintained by the β-lactamase gene, carbenicillin was added to a final concentration of 0.25 mg/ml. Chloramphenicol or kanamycin was added to maintain plasmid pLysS or pT7Pol26, respectively. A sample of the culture was removed as a zero timepoint, the remainder split into two, and inducer added to half. Unless otherwise stated in the text, induction was with 1 mM IPTG for the T7 promoter driven systems. Other promoter systems used are described in the text. Overexpression was typically for 3 hours. Total cell protein was analysed by SDS PAGE, using 1 ml of culture spun down into 150 ml 1× Disruption buffer. Overexpression on a preparative scale was performed using 1 L baffle flasks containing 400 ml Mu broth, inoculated with 3 - 5 ml of a fresh culture. Details are otherwise as for small scale overexpression.

Overexpression for Seleno-methionine (SeMet) derivatives

This protocol is based on (Doublie, 1997), and adjusted according to procedure from the laboratory of Jonas Emsley (Nottingham). Samples from fresh 5 ml overnight cultures, made from freshly transformed colonies, were used to inoculate 1 L baffle flasks containing 300 ml 56/2 and 3 ml 20% glucose; the appropriate selection was maintained at all stages. The secondary cultures were grown to an A_{600} of 0.3 at 37 °C, and metabolic inhibition was achieved by adding L-Lysine, L-Phenyalanine and L-Threonine at 100 mg/l, and L-Isoleucline, L-Leucine and L-valine at 50 mg/l, and finally L-SeMet at 50 mg/l; all solutions were syringe filtered before addition. After a 15 minutes period of incubation, a sample of the culture was removed as a zero timepoint, the remainder was induced with 1 mM IPTG for the T7 promoter driven systems. Overexpression was typically for 15 hours. Total cell protein was analysed by SDS PAGE, as for the standard procedures.

2.2.5.2 Protein purification

E. coli cultures containing overexpressed protein were harvested by centrifugation and the cell pellet resuspended in Buffer A (5-10 ml per g cells). Cell lysis was completed by sonication using a Soniprep 150 (MSE): 20 ml volumes were subjected to 6×30 second pulses at 8 microns on ice. Soluble protein and cell debris were separated by centrifugation (30,000 RCF, 20 minutes, 4 °C), and samples of each analysed by SDS PAGE. The soluble fraction was filtered through a 0.45 μm filter (Millipore) and subjected to further purification steps using various columns.

2.2.5.3 Column chromatography

Column chromatography was performed on FPLC (GE Healthcare) following standard procedures. All columns used in this study were purchased from GE Healthcare. Column chromatography buffers (section 2.1.1.2) were prepared and filtered under vacuum. All purification steps were performed at 4 °C or on ice, from the point at which cells were lysed. UV absorbance at 280 nm was used to detect protein peaks during column chromatography. SDS PAGE and the Bradford assay were used for subsequent protein analysis. Protein samples were filtered through 0.45 μm filters before loading onto the column to prevent blockages.

2.2.5.4 SDS-PAGE

Proteins were separated by one-dimensional electrophoresis using the Novex (Invitrogen) system throughout. The protein gels were separated into resolving, and stacking components. The resolving gel contained 12.5% acrylamide/ bis-acrylamide 29:1, 0.375 M Tris·HCl pH 8.8, and 0.1% (w/v) SDS. The stacking gel contained 5% acrylamide/ bis-acrylamide 29:1, 0.125 M Tris·HCl pH 6.8, and 0.1% SDS. Both portions of the gel were polymerised using 0.08% ammonium persulphate and 0.08% tetramethylethylene diamine (TEMED). Protein samples were prepared by mixing the sample and 1/2 volume of 3× SDS loading buffer, and heating to 95 °C for 5 minutes. Protein samples were loaded onto the gel, and run at 180 V/40 mA for 80 minutes in SDS PAGE running buffer. Gels were stained using Thermo Scientific GelCode Blue Safe

Protein Stain; excess stain was removed by soaking in water from 2 hours to overnight with gentle agitation.

2.2.5.5 Measurement of protein concentration

Protein concentrations were estimated using a Biorad Protein Assay. The A_{595} of a series of 1-24 μg/ml aliquots of BSA protein standard were determined, and compared to the A_{595} of the sample protein.

2.2.6 Crystallography

2.2.6.1 Protein preparation

The protein sample for crystallography was purified by standard procedures, and was finally exchanged in buffer B by gel-filtration column. It was kept at 4 ºC or on ice before the subsequent crystallisation, or kept at -80 ºC for long term storage.

2.2.6.2 DNA preparation

Oligonucleotides were bought from MWG. For ssDNA co-crystallisation, the oligonucleotides were mixed with the proteins directly. For dsDNA co-crystallisation, the oligonucleotides were annealed in 200 μl reactions, in a screw capped eppendorf, containing 0.5 mM of each oligo and 1× SSC. The mixture was placed in a PCR reactor starting at 95 ºC for 60 seconds, then entering a slope with decreasing temperature at 0.5 ºC per 12 seconds until it was 70 ºC. A less steep slope followed at 0.5 ºC per 60 seconds until it was 40 ºC, then the temperature was decreased at 0.5 ºC per 12 seconds again until it

was 4 ºC, at which point the reaction was held. The sample was run on a 15% polyacrylamide TBE gel with an un-annealed single strand oligonucleotide as a reference at 100 V for 2 hours. The gel was wrapped in Saran film and covered by a 2-ply silver foil to protect the majority amount of the annealed sample from the subsequent UV shadowing for locating the band. The un-irradiated portion of the band was excised and soaked with EB buffer at 4 ºC for overnight. The eluted DNA was concentrated to 100-200 µl by centrifugation in a pre-washed Vivaspin 6 5000 MWCO column at 6000 RCF, 8 ºC, or by using a Vivapore 7500 MWCO at 4 ºC. The DNA concentration was determined by the absorbance at A_{260}.

2.2.6.3 Crystallisation

Protein and DNA samples were mixed in ice cold buffer B and incubated on ice for 5 minutes, and were centrifuged at 13,000 rpm on a benchtop at 4 ºC for 10 minutes. The supernatant was transferred to a new tube, if there was pellet on the bottom of the tube. The samples were either crystallised manually, or by the Matrix Hydra high throughput crystallisation robot.

High throughput crystallisation

100 µl from a mother liquids set chosen for crystallisation screen was injected into each reservoir of a 96-cell CrystalQuick plate (Greiner Bio-one). 0.6 µl of each mother liquid was then moved to the three sample wells elevated in each cell to mix with 0.6 µl crystallisation samples. The plate was then sealed carefully by Crystal Clear Sealing Tape (Hampton Research) to completely

separate each cells. The plates were typically incubated at 10 ºC for 2-7 days and the results were rated by the visual inspection.

Sitting drop vapour diffusion in large scale

Mother liquids were added to a 24-cell Cryschem plate (Hampton Research). 500 µl of each mother liquid was added to appropriate reservoirs. 2 to 4 µl of each mother liquid was moved to the sample well by a Gilson P10 pipette. Same amount of crystallisation sample was added to mix with the mother liquid in each sample well. The plate was sealed with Crystal Clear Sealing Tape (Hampton Research) to separate each cells. The plates were incubated at 10 ºC, 13 ºC or 4 ºC.

Crystal dissolving

The liquid soaking the selected crystals was removed by a Gilson P10 pipette. The mother liquid from the reservoir in the same cell was used to wash the crystal for 4-5 times, and finally the crystal was dissolved with 5 µl of water. The protein was digested by Proteinase K at room temperature for 20-30 minutes.

2.2.7 Plasmid constructions and site-directed mutagenesis

The plasmids carrying mutant *rdgC* genes (Table 2-4) were generally constructed in two different ways. The pYJ plasmids were constructed with site-directed mutagenic PCR. Forward and backward primers were designed to overlap each other to introduce the base substitutions as required. The PCR reactions were catalysed by Phusion. The conditions of the PCR reactions were

designed based on QuikChange® Site-Directed Mutagenesis Kit (Stratagene). The PCR products were digested by *Dpn* I restriction enzyme to remove parental supercoiled DNA, and were transformed into DH5α. After an overnight growth at 37 ºC, transformants were selected and the mutant plasmids were extracted and sequenced for confirmation of the base substitutions.

The pGB plasmids were PCR constructed with the help from Dr G. S. Briggs, by using two adjacent primers. Since the Phusion enzyme used in the PCR reactions does not add phosphate at product ends, the PCR products were phosphorylated before religation. In this case, a unique diagnostic restriction site was always introduced by the primers to ensure that the correct plasmids were collected. For chunk deletions ($rdgC_{\Delta F}$ and $rdgC_{\Delta F \cdot R118C}$), two identical restriction sites were introduced by primers to both ends of the desired deletion region. The PCR products were digested by the appropriate enzyme and religated afterwards.

2.2.8 *In vitro* DNA binding assays

2.2.8.1 Linear DNA binding assays

Labelled DNA and protein were mixed on ice in GBB, and binding was allowed to proceed on ice for 15 minutes. Samples were run on a 4% polyacrylamide gel made with the appropriate LIS buffer and polymerised with 0.1% AMPS and 0.05% TEMED. Electrophoresis was completed at 160 V for 90 minutes, and the gels were dried on filter paper before being analysed by

phosphorimagery or autoradiography. The data were collected by using the Storm, and analysed by using ImageJ (Abramoff et al., 2004).

2.2.8.2 Plasmid DNA binding assays

Plasmid DNA and protein were mixed and incubated as in 2.2.8.1. Samples were run on 1% agarose gel. Electrophoresis was completed at 80 V for 90 minutes. Gels were stained by SYBR Green and visualised under UV light.

2.2.9 Chromosome Recombineering

As illustrated in Figure 2-1, plasmids carrying mutant *rdgC* genes of interests (Table 2-4) were digested at the *Hind*III site immediately downstream of *rdgC*, where a *cat Hind*III-*Hind*III fragment encoding chloramphenicol resistance was ligated. The region of *rdgC-cat* was amplified by PCR. PCR primers were designed to contain ~25 bases homology to either ends of the target genes (usually including a resistance marker gene) on a plasmid, and ~40 bases homologous to chromosomal gene insertion points. The plasmid was usually linearised by an appropriate restriction enzyme before the PCR. PCR products were removed for template by *Dpn* I and were dialysed on nitrocellulose filter disc against 1 mM HEPES for 30 minutes. The PCR products were then transformed into Δ*rdgC*::*dhfr* or Δ*rdgC2*::*dhfr* strains harbouring the pKD46 by electroporation (refer to 2.2.2.2; note that the transforming DNA were PCR products, rather than plasmids), and recombineered to replace the chromosomal Δ*rdgC*::*dhfr* or Δ*rdgC2*::*dhfr* with the help of pKD46. pKD46 is a temperature sensitive plasmid that encodes the phage λ Red recombinase and can be removed by growing the cells at 37 ºC for overnight (Datsenko and

Wanner, 2000). For the same reason, the electroporated cells were recovered at 30 ºC for 2 hours, instead of 37 ºC for 1 hour in the common procedure (see 2.2.2.2 for details). The recombineered sequences in the chromosome were confirmed by PCR mediated sequencing using primers Pro/rdgC-5' and rdgC/Cm-3'. Chromosomal *rdgC* was deleted in the same way by introducing Δ*rdgC*::*dhfr* or Δ*rdgC2*::*dhfr*, which was confirmed by PCR mediated sequencing using primers Tm/rdgC-5' and Tm/rdgC-3'. *rdgC* from *Neisseria meningitidis* was also engineered to replace the native *rdgC* in *E. coli* with the same method. Its presence in the chromosome was confirmed by PCR mediated sequencing using primers RdgC/Nm-5' and RdgC/Cm-3'.

Table 2.4 *rdgC* mutants

Amino acid mutation	Base changes	PCR primers	Plasmid construct	Recombineered strain
K211A	AAA→GCA	YJ21/YJ22	pYJ014	AM1839
Q212A	TCG→GCA	YJ29/YJ30	pYJ017	AM2029
E218R	GAG→CGC	YJ13/YJ14	pYJ013	AM1840
H222A	CAC→GCC	YJ17/YJ18	pYJ006 pYJ003	AM1836
K227A	AAA→GCA	YJ19/YJ20	pYJ007	AM1837
R118C	CGT→GTG	GB-P37/34	pGB047 pGB045	AM1833
R118A	CGT→GCG	GB-P41/42	pGB050 pGB049	AM1838
F120T	TTC→ACG	GB-P47/48	pGB054 pGB053	AM1905
F120S	TTC→AGC	GB-P45/46	pGB052 pGB051	AM1917
R118C·F120T	GCG→GTG TTC→ACG	GB-P77/78*	pGB065	YJ025
P76G,L116T,Δ(77-115)	CCG→GGT CTG→ACC	GB-P29/30	pGB043	AM1834
P76G,L116T,Δ(77-115),R118C	CCG→GGT CTG→ACC CGT→GCG	GB-P37/35	pGB048 pGB046	AM1835

R97S,K98Q,K100Q,K101E	CGT→AGT AAG→CAG AAG→CAG AAA→GAA	GB-P51/52	pGB056 pGB055	AM1950

***** The primers introduced base substitutions for F120T; R118C was borrowed from pGB045 as a NdeI/AatII fragment.

2.2.10 Synthetic lethality assays

The assay used an unstable *lac*$^+$ plasmid, pRC7, carrying a wild type gene of interest to cover a deletion allele in the chromosome of a *lac*$^-$ strain. If any further mutation is lethal with the covered mutation, plasmid-free segregants fail to grow and only plasmid carrying cells forming blue colonies are seen on an indicator plate containing IPTG and X-gal. IPTG induces the expression of the *lac* operon on the pRC7 plasmid, and the expression product digests X-gal on the plate, transforming the cells having the plasmid to blue. If the further mutation is not lethal with the covered mutation, plasmid free segregants form healthy white (*lac*$^-$) colonies.

Stock cultures of strains carrying pAM374 (a pRC7 derivative, encoding PriA) were diluted 80-fold in LB broth and grown with no ampicillin selection to A_{650} 0.4 before plating dilutions on LB or minimal agar supplemented with X-gal and IPTG. Plates were photographed and scored after 48 hours (if LB) or 72 hours (if minimal) of incubation. White colonies formed by plasmid-free cells could be subcultured on LB or minimal agar where appropriate. The method details can be found in (Mahdi et al., 2006).

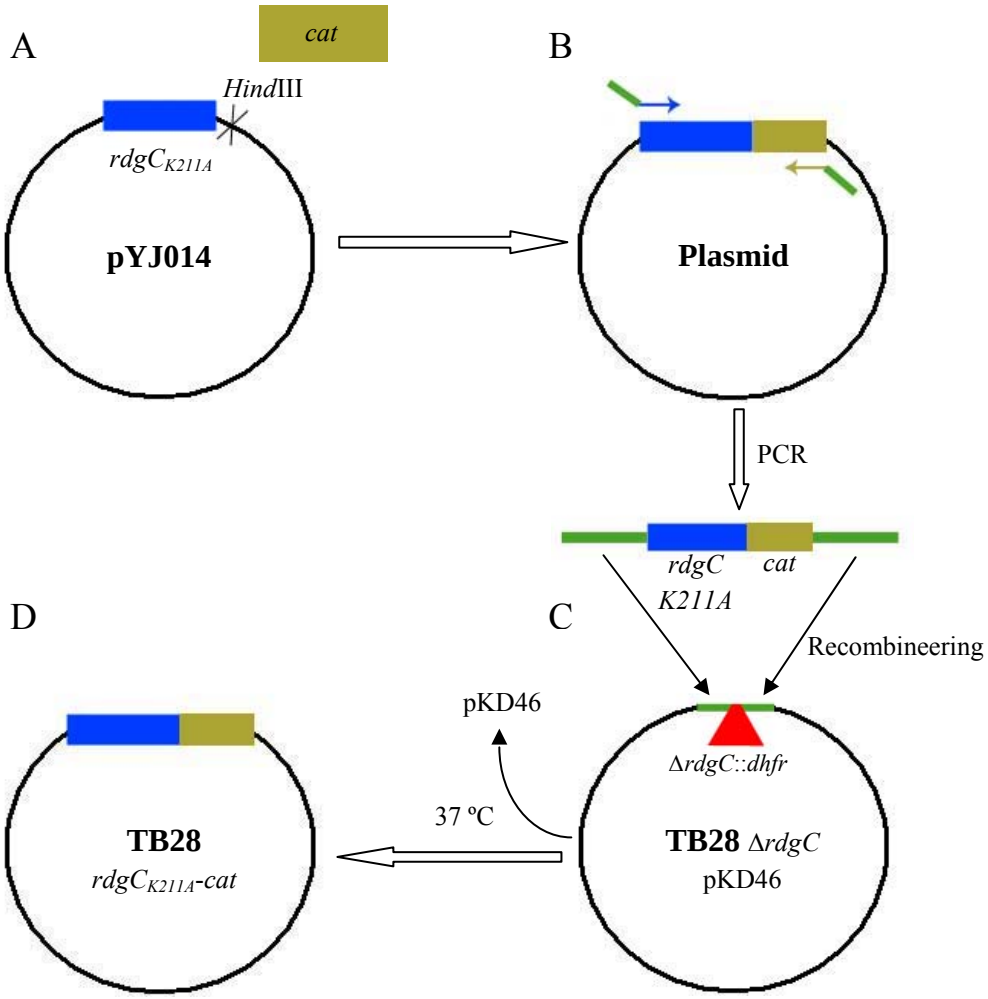


Figure 2-1 General procedure of recombineering
This is an example to use $rdgC_{K211A}$ from a plasmid to replace Δ*rdgC*::*dhfr* in the chromosome of a TB28 derivative. All the other mutant *rdgC* were recombineered in the chromosome following the same procedure. The green bars indicate homologous sequences.

A. A *Hind*III-*Hind*III fragment of *cat* gene was ligated immediately downstream of $rdgC_{K211A}$ in a plasmid (usually a pET22b derivative).
B. The region of $rdgC_{K211A}$-*cat* was PCR amplified with primers containing sequences homologous to Δ*rdgC*::*dhfr* flanking sequences in the chromosome of a TB28 derivative.
C. pKD46 encoding protein phage λ Red recombinase mediated recombination that would replace Δ*rdgC*::*dhfr* with $rdgC_{K211A}$-*cat*.
D. The recombinants were incubated at 37 ºC for overnight. The temperature sensitive plasmid pKD46 was lost.

Chapter 3
Structural analysis of RdgC

3.1 Background

RdgC's association with DNA recombinational repair is thought to be achieved by binding to DNA, however, its DNA binding does not show any sequence or structural preference. This poses the question of how RdgC achieves its effect on the recombination reaction.

Attempting to answer this question, I overexpressed, purified and crystallised the protein. Its seleno-methionyl derivative was also crystallised for phase determination (selenium replacing the sulphur in all methionine residues of the protein, therefore giving an opportunity to determining the phase). The structure of RdgC was finally determined by X-ray crystallography.

3.2 Crystallisation of RdgC

3.2.1 Cloning and Overexpression of native RdgC

The *rdgC*$_{wt}$ gene had previously been cloned into pT7-7, generating pGS853, which can be used to overexpress RdgC (Moore, 2002). For a tight control of overexpression, the *rdgC* coding sequence was subcloned as an *Nde*I-*Hind*III fragment in pET22b, forming pYJ001. pYJ001 was transformed either into BL21 (DE3) pLysS cells or STL2669 pT7Pol26 cells. Overexpression was carried out as described (see 2.2.5.1). The protein was visible without purification after overexpression (Figure 3-1).

3.2.2 Purification of native RdgC

The cells expressing RdgC were lysed by sonication on ice. RdgC in the supernatant was precipitated by performing a two-step ammonium sulphate precipitation. The supernatant from the 0 – 40% ammonium sulphate cut was collected for a second ammonium sulphate cut from 40 – 65%. The precipitate containing RdgC was resuspended in buffer A containing 0.3 M NaCl and cleaned by a 0.45 micron filter before it entered column chromatography stages. The first column to purify RdgC was a heparin column (HiTrap^{TM} Heparin HP 5 ml (GE Healthcare)) and it removed most impurity in the solution (Figure 3-2). Fractions containing RdgC were applied to a Q sepharose (HiTrap^{TM} Q Sepharose HP 5 ml (GE Healthcare)). After this stage the solution was applied to the heparin column for concentration due to the 5 ml loading limit of the following gel filtration column (Sephacryl 16/60 S200HR). After purification by the gel filtration column, the protein was concentrated and quantified on the assumption that the protein is a dimer (see 1.5), with a total mass of 68 kDa. Note that RdgC for crystallisation was not dialysed against 50% glycerol as it might hamper crystal growth.

3.2.3 Overexpression of seleno-methionyl RdgC

The protein was overexpressed from *E. coli* BL21 (DE3) carrying pLysS and pYJ001 (as native RdgC) using feedback inhibition of methionine biosynthesis, as described (Doublie, 1997). The detailed procedure can be found in 2.2.5.1.

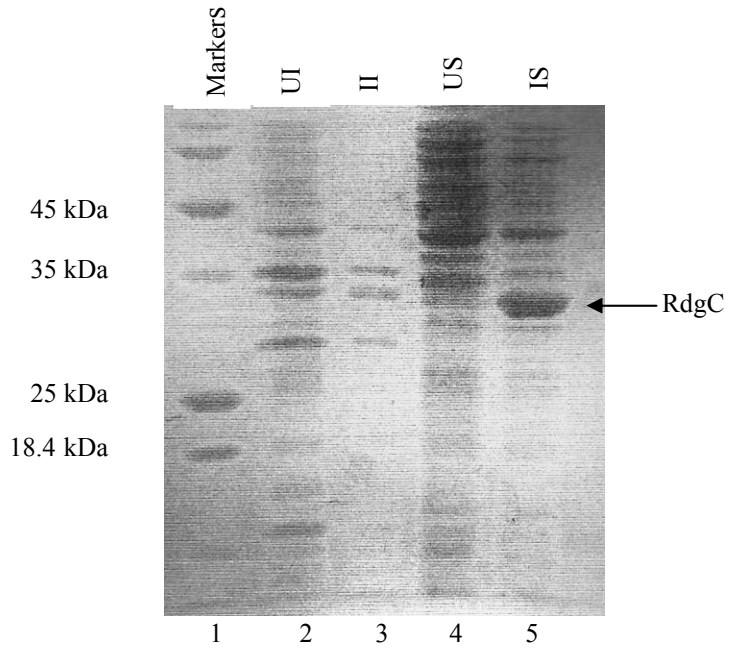


Figure 3-1 12.5% SDS-PAGE gel showing overexpression of *Escherichia coli* wild type RdgC. Lane 1, molecular weight markers; Lane 2, UI: insoluble lysate from uninduced cells; Lane 3, II: insoluble lysate from induced cells; Lane 4, US: soluble lysate from uninduced cells; Lane 5, IS: soluble lysate from induced cells.

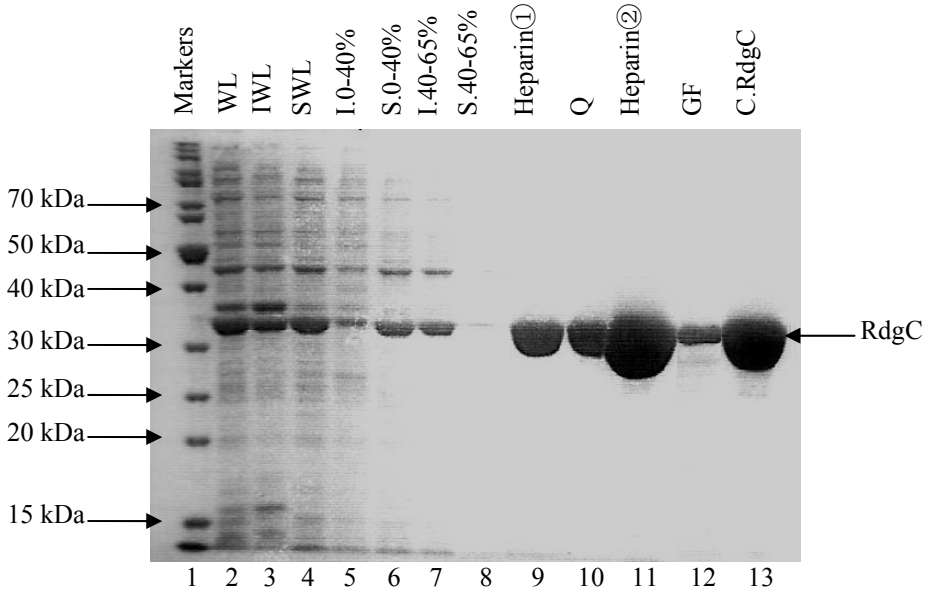


Figure 3-2 12.5% SDS-PAGE gel showing ammonium precipitation and column stages of purification of *Escherichia coli* wild type RdgC

Lane 1, molecular weight markers; Lane 2, whole cell lysate from induced cells; Lane 3, insoluble whole cell lysate; Lane 4, soluble whole cell lysate; Lane 5, insoluble fraction of 0-40% $(NH_4)_2SO_4$ cut; Lane 6, soluble fraction of 0-40% $(NH_4)_2SO_4$ cut; Lane 7 insoluble fraction of 40-65% $(NH_4)_2SO_4$ cut; Lane 8 soluble fraction of 40-65% $(NH_4)_2SO_4$ cut; Lane 9-12, collected fractions from a series of columns as indicated above the lanes. Note that Heparin① indicates the Heparin column was used for purification, whereas Heparin② was used for concentrating. GF stands for Gel-filtration; Lane 13, Concentrated RdgC.

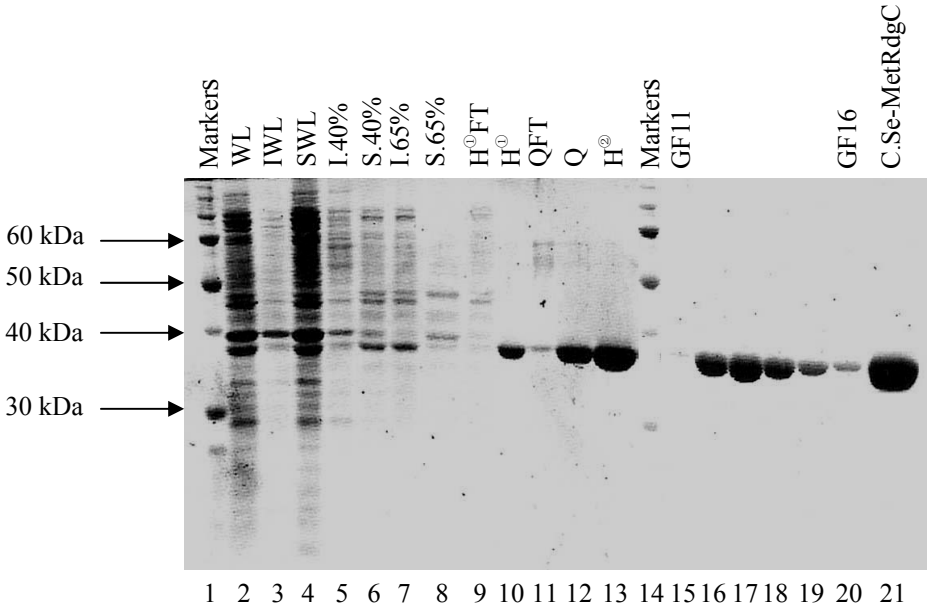


Figure 3-3 12.5% SDS-PAGE gel showing ammonium precipitation and column stages of purification of Se-Met RdgC

Lane 1 and 14, molecular weight markers; Lane 2, whole cell lysate from induced cells; Lane 3, insoluble whole cell lysate; Lane 4, soluble whole cell lysate; Lane 5, insoluble fraction of 0-40% $(NH_4)_2SO_4$ cut; Lane 6, soluble fraction of 0-40% $(NH_4)_2SO_4$ cut; Lane 7 insoluble fraction of 40-65% $(NH_4)_2SO_4$ cut; Lane 8 soluble fraction of 40-65% $(NH_4)_2SO_4$ cut; Lane 9-13, 15-20 collected fractions from a series of columns as indicated above the lanes. Note that Heparin① indicates the Heparin column was used for purification, whereas Heparin② was used for concentration. GF stands for Gel-filtration; Lane 21, Concentrated Se-Met RdgC.

3.2.4 Purification of seleno-methionyl RdgC

The whole procedure was performed as for *E. coli* native RdgC (3.2.3) (Figure 3-3). Incorporation of Se-Met was confirmed by the protein's molecular weight, obtained by mass spectrometry.

3.2.3 Crystallisation trials

3.2.3.1 High throughput crystallisation

The three samples prepared for crystallisation screening were *E. coli* native RdgC alone, *E. coli* native RdgC with YJ1 (a 21 nt ssDNA) and *E. coli* native RdgC with YJ5 (a 15 nt ssDNA). The concentrations in the crystallisation mix of protein and DNA were both 100 μM. When RdgC was crystallised alone, the portion for DNA was replaced by Elution Buffer (Qiagen).

The Matrix Hydra high throughput crystallisation robot was applied to operate crystallisation on 96-cell crystallography plates. The three mother liquid suites selected for crystallisation screening were Classic Suite, MPDs Suite and PEGs Suite (Hampton research). Crystals were grown at 10 ºC by sitting drop vapour diffusion, with a 1:1 ratio of complex and reservoir solution in each drop. The results of the trials are shown in Table 3-1, Table 3-2 and Table 3-3, respectively.

Table 3-1 The Classics Suite					
Mother liquid			Crystal forms		
Salt	Buffer	Precipitant	RdgC	RdgC:YJ1	RdgC:YJ5
Null	0.1M HEPES sodium salt pH 7.5	2 %(v/v) PEG 400, 2.0 M Ammonium sulfate		Big size, cube like crystals	
0.2 M Ammonium acetate	0.1 M tri-Sodium citrate pH 5.6	30 %(w/v) PEG 4000	Needles		Needles
0.2 M Ammonium sulphate	0.1 M MES pH 6.5	20 %(w/v) PEG 5000 MME	Needles		Needles

Table 3-2 The MPDs Suite					
Mother liquid			Crystal forms		
Salt	Buffer	Precipitant	RdgC	RdgC:YJ1	RdgC:YJ5
0.2 M di-Ammonium phosphate	Null	40 %(v/v) MPD	Big size, cube like crystals		Needles

Table 3-3 The PEGs Suite					
Mother liquid			Crystal forms		
Salt	Buffer	Precipitant	RdgC	RdgC:YJ1	RdgC:YJ5
0.2 M Magnesium nitrate	Null	20 %(w/v) PEG 3350	Needles and plates	Needles and plates	
0.2 M Magnesium sulfate	Null	20 %(w/v) PEG 3350	Plates	Plates	

3.2.3.2 Sitting drop vapour diffusion in large scale

Crystallisation in large scale was operated on 24-cell crystallography plates. *E. coli* native RdgC was crystallised with different lengths and structures of DNA substrates (see Table 3-4; the sequences of the composing oligonucleotides can be found in Table 2-3).

Based on the results of high through-put crystallisation, the protein:DNA complexes (the DNA molecule was either ssDNA or blunt duplex) and the protein alone were crystallised in $CaCl_2$ or $MgCl_2$ as salt, HEPES (pH 6.5-8.5) as buffer, and PEG (400, 1000, 2000, or 4000) as precipitant. Crystals were grown at 10 ºC, 13 ºC, or 4ºC.

After many trials, the optimal conditions for crystallisation were identified as when RdgC co-crystallised with YJ7/8 at 10 ºC, from 200 mM $CaCl_2$, 100 mM HEPES pH 7.0-7.5, and 15-20 % PEG 1000. Crystals under such conditions grew up to 400 × 200 × 200 μM (Figure 3-4 A). Crystallisation in the absence of DNA produced smaller and non-isomorphous crystals, and diffracted to only 2.8 Å.

Since the blunt dsDNA and ssDNA molecules applied in crystallisation could not be located in the structure (see below for details), I also made some unusual DNA substrates in the hope to trigger a distinctive interaction between RdgC and DNA, or that the conformation of the DNA molecules could help themselves arrange in the crystal in a unique manner. The expected structures of the unusual DNA substrates include duplexes with a bubble, or hairpin at

one or both ends, overhangs, and a duplex with a kink (Table 3-4).

Unfortunately, the DNA still remained invisible in the crystal structure.

Table 3-4 DNA substrates for crystallography

Name	Length	Structure
YJ1	21 nt	
YJ3	18 nt	
YJ5	15 nt	
YJ7	19 nt	
YJ1/2	21 bp	
YJ3/4	18 bp	
YJ5/6	15 bp	
YJ7/8	19 bp	
YJ9/10	17 bp	
TMRC1/1R	19 bp	
YJ1/10	17 bp with 4 nt overhangs	
YJ1/6	15 bp with 6 nt overhangs	
YJ2/5	15 bp with 6 nt overhangs	
YJ2/9	17 bp with 4 nt overhangs	
TMRC1M1/1R	19 bp with 1mismatch bubble	
TMRC1M3/1R	19 bp with 3 mismatch bubble	
TMRC2-DUMB	18 bp with 4 nt hairpins at both ends	
TMRC3-DUMB		
TMRC1M1/1-D1	18 bp with a kink.	
YJ11	18 bp with a 4 nt hairpin at one end	
YJ12	17 bp with a 4 nt hairpin at one end	

In order to confirm the presence of DNA in the crystals, some washed and dissolved crystals were analysed by gel electrophoresis on a 15 % polyacrylamide TBE gel. As illustrated in Figure 3-4 B, DNA existence was confirmed both in the drop solvent and the crystals. Its lacking density for DNA in the Fo-Fc difference map may be related to the mode of RdgC:DNA binding.

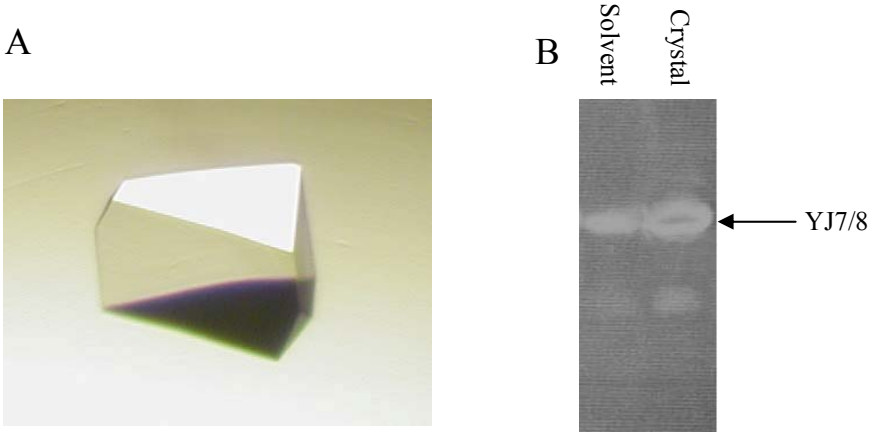


Figure 3-4 RdgC crystal and EMSA assay proving the presence of DNA in the crystal

A. *E. coli* native RdgC and YJ 7/8 (19 bp dsDNA) crystallised to 400 x 200 x 200 µM
B. Washed crystal of (Se-Met RdgC:YJ 7/8) was solved and run on a 15% polyacrylamide TBE gel. Lane 1, drop solvent; Lane 2, washed and solved crystal.

3.3 X-ray diffraction

Crystals for diffraction were harvested by Dr Geoffrey Briggs. Data for the Se-Met derivative were collected on ID14-4 at the European Synchrotron Radiation Facility with a wavelength of 0.98 Å. Data processing was performed by Paul McEwan (Nottingham). The structure presented in section 3.4 below is for the Se-Met RdgC co-crystallised with YJ7/8 (19 bp blunt dsDNA), which was solved to 2.4 Å.

The packing of RdgC molecules in the crystal lattice allows holes to form long and continuous pores (Figure 3-5 A). However, no additional density was observed in the Fo-Fc difference map that could be used to locate the DNA, despite the evidence of its presence in the crystal (3.2.3.2).

3.4 Structure of RdgC

The data generated were used to solve the structure of RdgC. This part of the project was conducted by Dr Geoffrey Briggs in collaboration with Paul McEwan, although I was closely associated with the structure determination throughout.

In agreement with previous studies (Moore et al., 2003), RdgC crystallises as a dimer, with a head to head, tail to tail organisation (Figure 3-5 B and C). The

dimer forms a ring structure, with a 30 Å diameter central hole (C alpha to C alpha).

The RdgC protein is a unique fold, as revealed by the Dali server (Holm and Sander, 1993). The overall protein fold adopts a toroidal structure mimicking a horseshoe, with two gates closing the open end (Figure 3-5 B). The inner face consists of 8 anti-parallel β strands, with one of the dimer interfaces (horseshoe interface) falling between the 4^{th} and 5^{th} strands. They are covered by 2 anti-parallel helices on the outer surface. The horseshoe interface is held together by hydrogen bonds between the β strands and between a triad of highly conserved residues at the ends of the helices: Q212 and E218 from one monomer and K227 from the other. The conserved histidine, H222, completes the dimer interface (Figure 3-6 and 3-7 A).

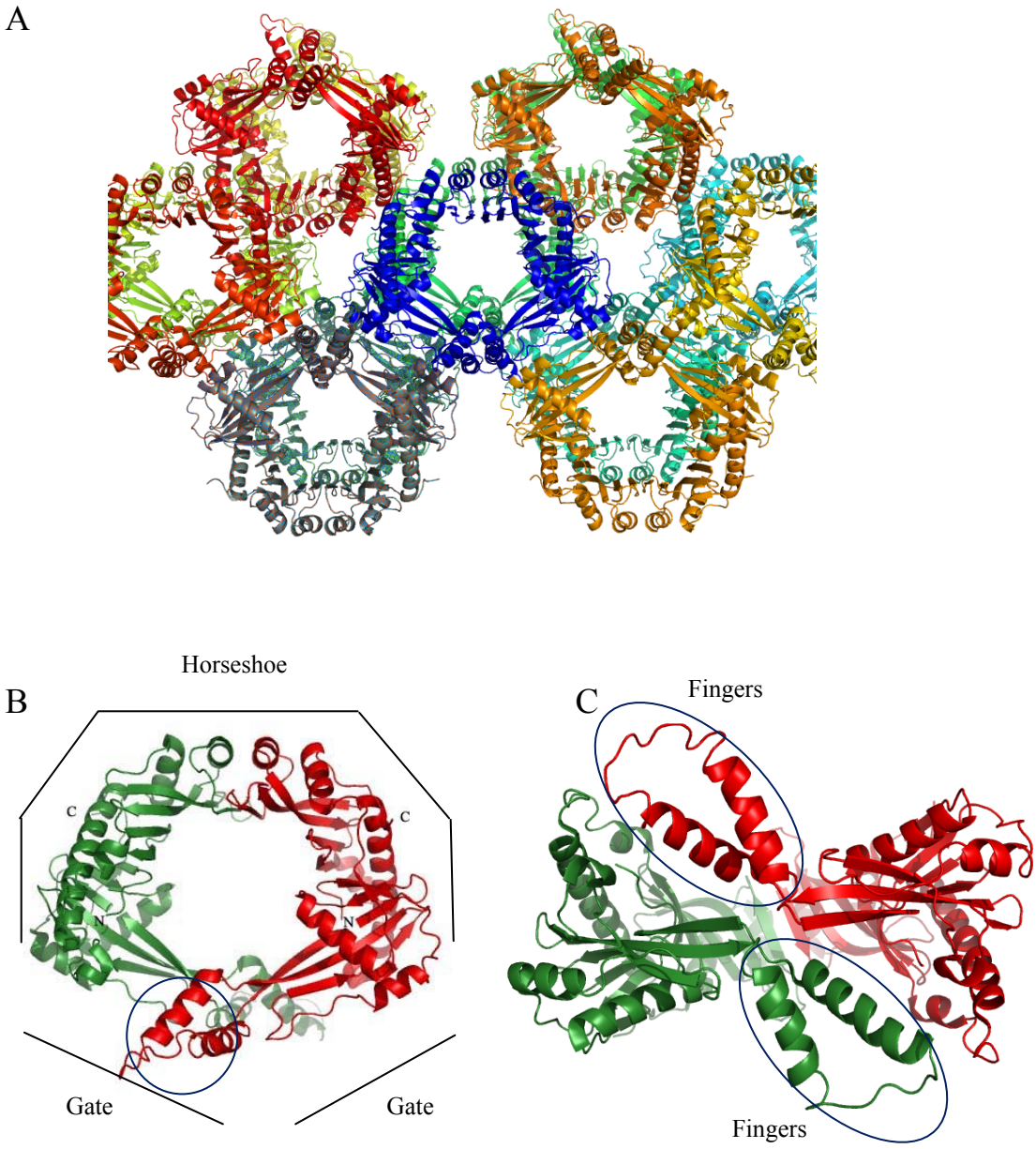


Figure 3-5 Overall structure of RdgC

A. Crystal packing of RdgC rings to from long pores through the lattice.
B. Ribbon diagram showing the ring structure of the RdgC dimer and identifying the horseshoe and gate regions. The two chains are differentiated by shades of red and green. Note that the N-termini are buried in the molecule, and the C-termini are exposed outside. The open circle marks the location of one of the finger domains.
C. Horizontally rotated ribbon diagram showing the bottom and the finger domains.

The second dimer interface lies where the two regions of extended β strand forming the gates overlap. Unlike the horseshoe dimer interface, the gate interface is primarily maintained by hydrophobic interaction, where a conserved phenylalanine, F120, on one chain sits in a pocket formed by highly conserved residues on the other chain: I74, L75, P76, V79, L115, L116, R118, A119 and F120, and *vice versa* (Figure 3-6 and 3-7 C). The hydrogen bond formed between the guanidinium and carbonyl groups of the R118 residues provides additional stabilisation (Figure 3-7 D).

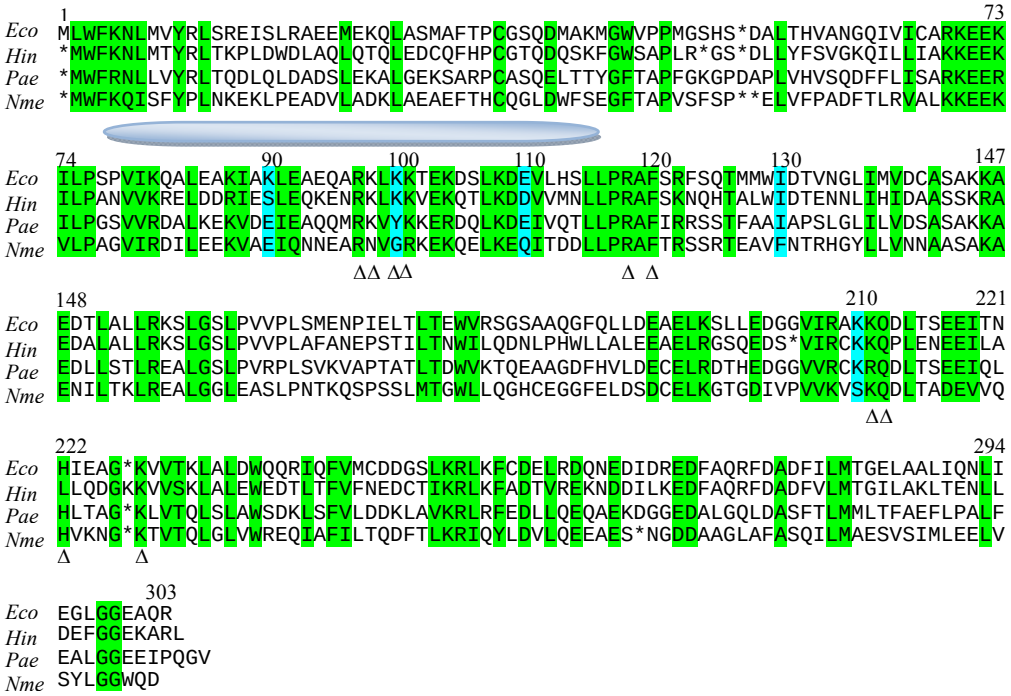


Figure 3-6 Alignment of RdgC

Eco=E. coli; *Hin=Haemophilus influenzae*; *Pae=Pseudomonas aeruginosa*; *Nme=Neisseria meningitidis*. Highly conserved residues have green shading. Residues with triangles participated in mutagenesis studies. Residues shaded in light blue correspond to the numbering on their top. The grey-blue bar marks the residues underneath composing the finger domain, from S77 – L115.

Two anti-parallel α helices extended from the gate interface project perpendicular to the ring on either side, dubbed as 'finger' domains. They range from S77 to L115 and contain a high proportion of conserved lysines, rendering a highly electropositive upper surface. The turn between the two helices is also electro-positively charged, attributing to an arginine (R97) and three lysines (K98, 100, 101) (Figure 3-7 B). The two finger domains do not have identical conformations in the crystal structure, a possible consequence of non-specific crystallographic contacts. It may also reflect their high B factors, and it is possible that some flexibility is necessary for protein function.

The net charge of RdgC is highly negative at physiological pH, since a monomer has 27 glutamate, 20 aspartate, 23 lysine, 16 arginine and 4 histidine residues. Therefore the net charge of a dimer would be -8 if all histidines are positively charged. However, the charge distribution is non-uniform, with most negatively charged residues lying on the outer surface and most positively charged residues sitting around the wall of the centre hole (Figure 3-8 A).

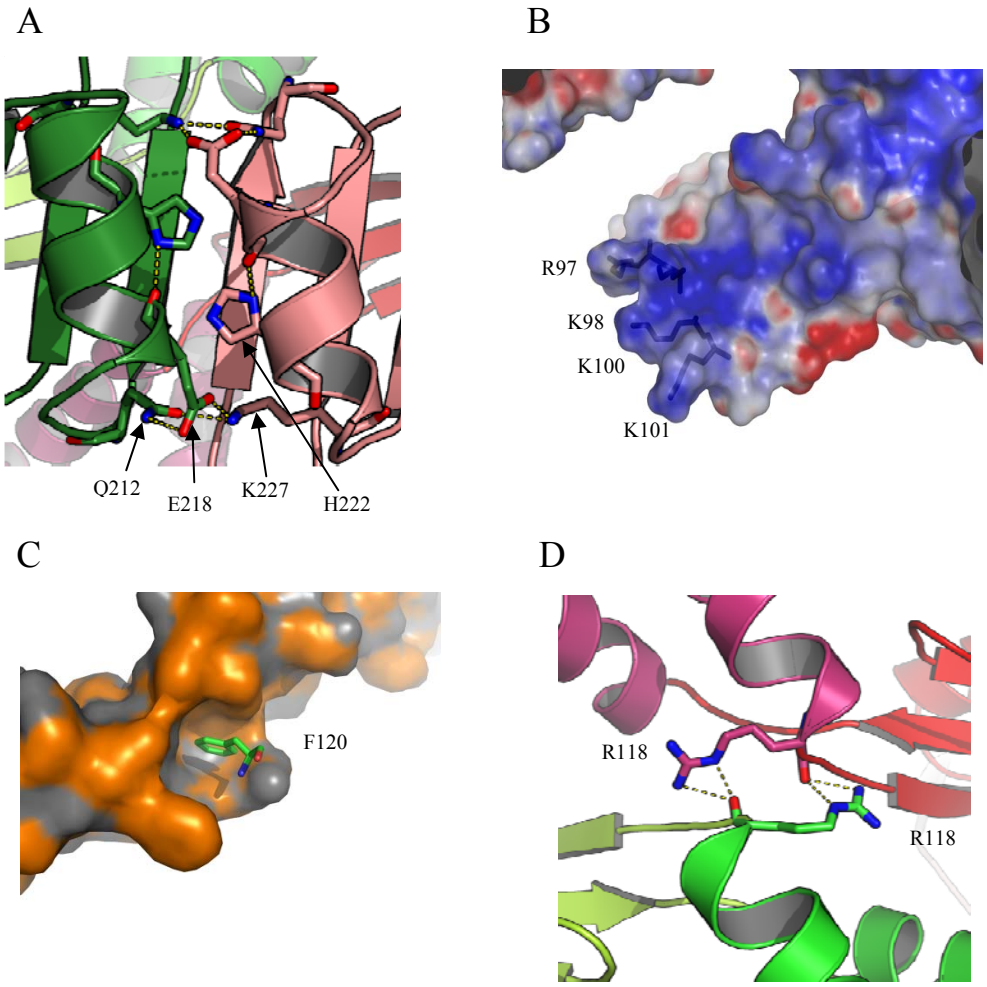


Figure 3-7 Conserved residues at the dimer interfaces and finger domains (Briggs et al., 2007)

A. A triad of conserved residues forming two hydrogen bond networks at the horseshoe interface.
B. An electrostatic surface representation showing the four basic residues at the tip of the finger domain.
C. A hydrophobic surface representation showing the pocket inserted with F120 at the gate interface.
D. R118 forming hydrogen bonds across the gate interface.

3.5 Discussion

X-ray crystallography revealed that the RdgC dimer adopts a novel structure. Despite the evidence supporting the presence of DNA in the crystals, it could not be located in the structure of the DNA binding protein. This can be explained by its non-specific DNA binding as suggested by previous research (Moore et al., 2003).

However, several lines of data suggest RdgC binds DNA by encircling it in the central hole of the dimer.

1) The diameter of the central hole is 30 Å, which is readily able to accommodate a dsDNA (25 Å diameter in width).
2) The central hole of the dimer is positively charged, which can directly interact with the negatively charged DNA backbone (Figure 3-8 A, B). This also explains limited preference on DNA structure and sequence.
3) The proposed DNA binding channel is 40 Å from the front to the back of the ring and 55 Å from fingertip to fingertip. This area can interact with approximately 18 bp of DNA, in agreement of the size (15-20 bp) required to make a stable complex *in vitro* (Moore et al., 2003) (Figure 3-8 C, D).
4) The crystal packing of RdgC leaves only the central hole for accommodating the DNA (Figure 3-5 A).

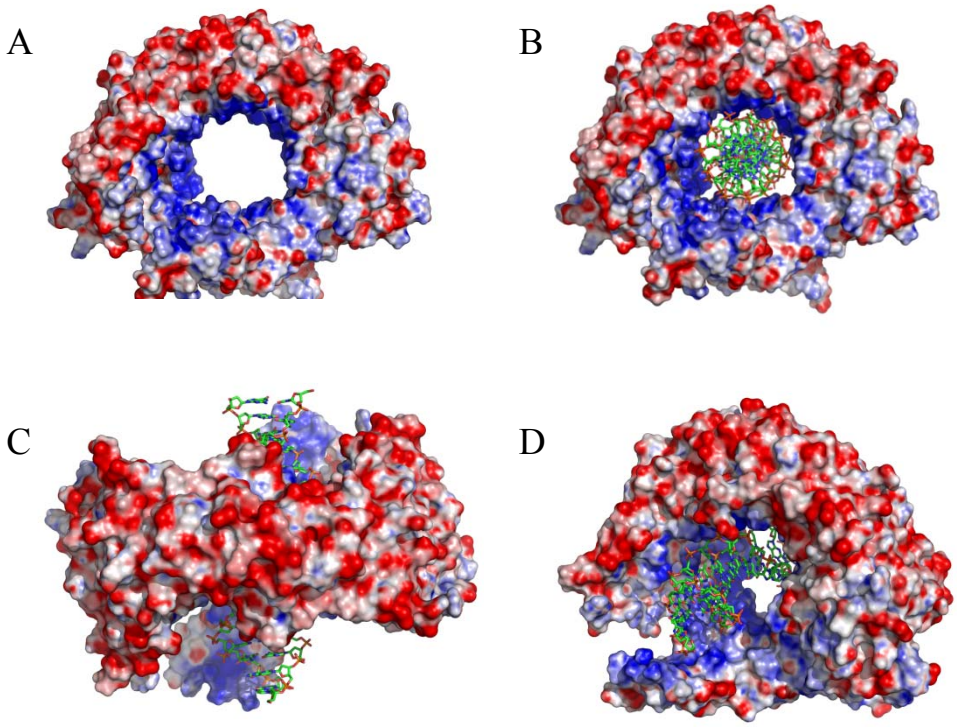


Figure 3-8 Electrostatic surface representations of RdgC and modelling of DNA binding (Briggs et al., 2007)

A. The ring of the RdgC dimer has an electropositive inner surface (blue) and an electronegative outer surface (red).

B. C. and D. Three orientations of RdgC with duplex DNA modelled in the positive central hole.

Although a DNA duplex could fit in the central hole as modelled in Figure 3-8 B, C and D, it was not directly obvious how RdgC might load on DNA. The following chapters describe my attempts to unravel the mechanism of DNA loading.

Chapter 4
The horseshoe interface

4.1 Background

The ring structure of RdgC suggests a simple mode of DNA binding – the ring encircles duplex DNA, thus limiting its exposure to RecA activity. This may explain how RdgC prevents illegitimate DNA recombination mediated by RecA. The active RecA:ssDNA filament searches homologous duplex and promotes strand exchanges between the two DNA molecules. The RdgC ring might cover the target dsDNA to prevent the recombination, or even halt it by "robbing" the dsDNA from the RecA:ssDNA filament.

The question to this point is how RdgC loads onto DNA.

a) Does it act only at DNA ends, in which case the closed ring might simply slide onto the DNA? However, this is not likely the case if RdgC blocks RecA loading at DNA gaps. Or
b) Does the ring open to allow loading at any point on DNA?

Preliminary studies using Atomic Force Microscopy suggest a possible preference for DNA ends (Tessmer et al., 2005). However, other data suggest RdgC could also bind circular plasmid DNA (Tessmer et al., 2005), which appears only achievable by opening the ring.

I examined this aspect of RdgC in more details with mutagenesis assays. This chapter will focus on the studies concerning the horseshoe interface of the ring.

4.2 *In vitro* DNA binding assays

4.2.1 Mutant design

The horseshoe interface is held together by hydrogen bonds between the β strands, and two sets of salt bridges formed in a triad of highly conserved residues at the ends of the helices: Q212 and E218 from one monomer and K227 from the other. H222 may also be involved. Since any change in the β strands might disrupt the whole structure of the RdgC dimer, I only generated mutants for Q212 (pYJ017:RdgC$_{Q212A}$), E218 (pYJ013:RdgC$_{E218R}$), K227 (pYJ007:RdgC$_{K227A}$) and H222 (pYJ006:RdgC$_{H222A}$). Another residue, K211, appears to be able to interact with the backbone of bound DNA duplex, thus the construction of RdgC$_{K211A}$ (pYJ014).

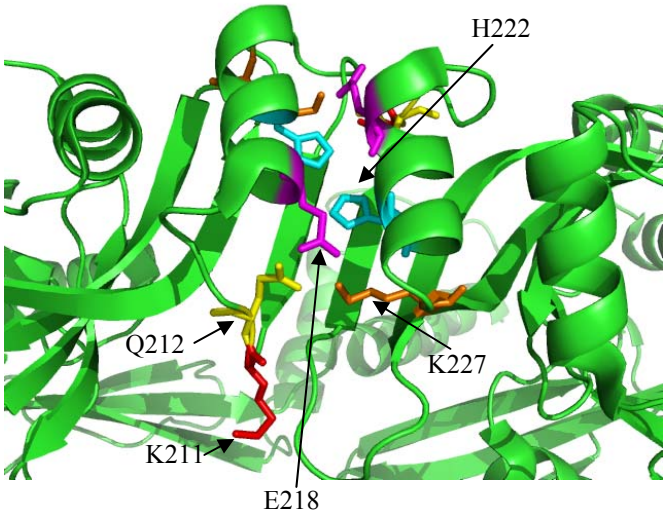


Figure 4-1 Ribbon diagram illustrating the horseshoe interface of RdgC dimer and the residues investigated in mutant studies.

4.2.2 Cloning, overexpression and purification

4.2.2.1 Cloning

Primers were designed to introduce base substitutions, on a plasmid with *E. coli* native *rdgC* gene fragment cloned between an *Nde*I site and a *Hind*III site. Normally the template plasmid was pYJ001, a pET22b derivative. After mutagenic PCR on pYJ001, the product was digested with Dpn I restriction enzyme to remove methylated parental DNA template. The treated sample was transformed into competent DH5α cells. After proliferation of the cells, the plasmids of interest were purified, sequence checked and stored at -20 ºC.

4.2.2.2 Overexpression

In order to avoid chromosomally encoded wild type RdgC, the chromosomal *rdgC* gene was deleted in strain MG1655 to form N4586. Δ*xonA300* was introduced to N4586 to form YJ012. The cells were then transformed with plasmid pT7pol26 to form YJ015 as an overexpression strain for mutant RdgC proteins. Another overexpression strain served for the same purpose stemmed from BL21 (DE3) pLysS, the chromosomal *rdgC* of which was deleted and replaced with *dhfr* to form YJ014. The plasmid having the mutant *rdgC* of interest was then transformed into an expression strain. The production of mutant RdgC proteins did not differ much in the two backgrounds (data not shown).

4.2.2.3 Purification

Mutation of residues potentially affecting RdgC's DNA binding affinity might be expected to affect the affinity of the protein for a heparin column. Therefore the mutant proteins were purified on a linear gradient of 0.0-1.0 M NaCl in Buffer A (see Appendix 1). For the same reason, when concentrated by the Heparin column, the protein sample was washed in Buffer A without NaCl. Other than that purification was performed as for *E. coli* native RdgC.

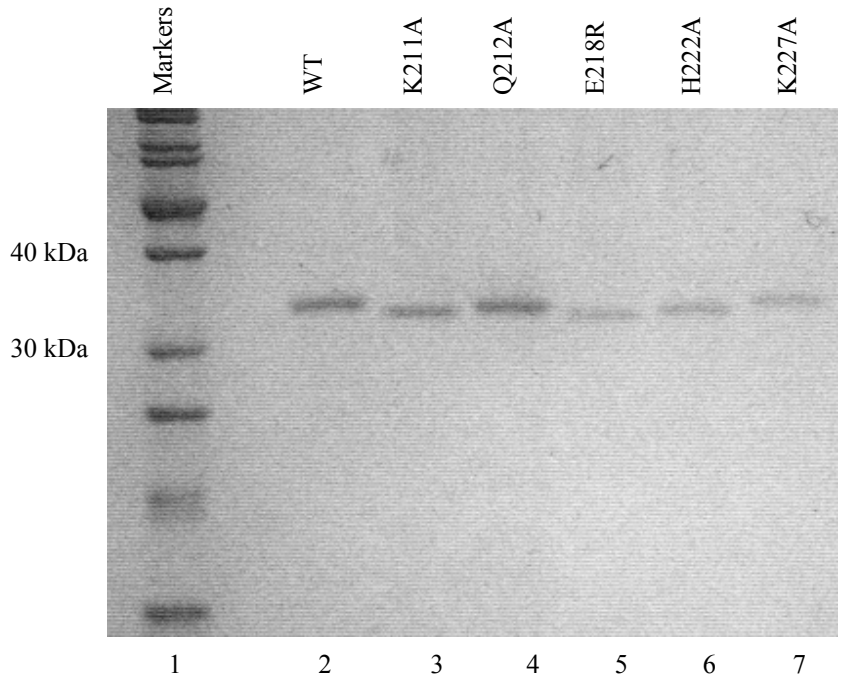


Figure 4-2 12.5% SDS-Page gels showing purified mutant RdgC proteins with *E. coli* native RdgC

Lane 1, Molecular weight markers; Lane 2, wild type RdgC; Lane 3-7, RdgC proteins with mutations as labelled on the top of each lane.

4.2.3 Single-stranded DNA binding assays

In general, all tested mutants and the wild type RdgC were able to shift the 49 nt ssDNA in electrophoretic mobility shift assays (EMSAs). The results show that RdgC formed two distinct complexes with the ssDNA oligo, RGL13 (Figure 4-3 A, wild type). This is consistent with the RdgC structure, which could interact with approximately 18 bp of dsDNA. Although RdgC requires a slightly longer ssDNA for stable contact (Moore et al., 2003), the 49 nt RGL13 is supposed to interact with two RdgC dimers simultaneously when RdgC is in excess (as illustrated on the right of WT in Figure 4-3 A).

The horseshoe interface mutants exhibited different abilities to shift ssDNA. Among them, $RdgC_{Q212A}$ and $RdgC_{H222A}$ were able to shift the ssDNA to the same extent as the wild type. $RdgC_{K211A}$ bound with only slightly lower affinity, and as $RdgC_{Q212A}$ and $RdgC_{H222A}$, it could shift the ssDNA as a sharp band. In contrast, $RdgC_{E218R}$ and $RdgC_{K227A}$ could not form a stable complex with RGL13 even in great protein excess. This resulted in massive complex dissociation during electrophoresis (Figure 4-2 A, bottom right, Lane 4). To be specific, wild type RdgC, $RdgC_{Q212A}$ and $RdgC_{H222A}$ started to shift 0.2 nM RGL13 at 1 nM (RdgC dimers:DNA≈5:1); $RdgC_{K211A}$ started at 3 nM RdgC dimers:DNA≈15:1); $RdgC_{E218R}$ and $RdgC_{K227A}$ barely started at 20 nM (RdgC dimers:DNA≈100:1) (Figure 4-3 B).

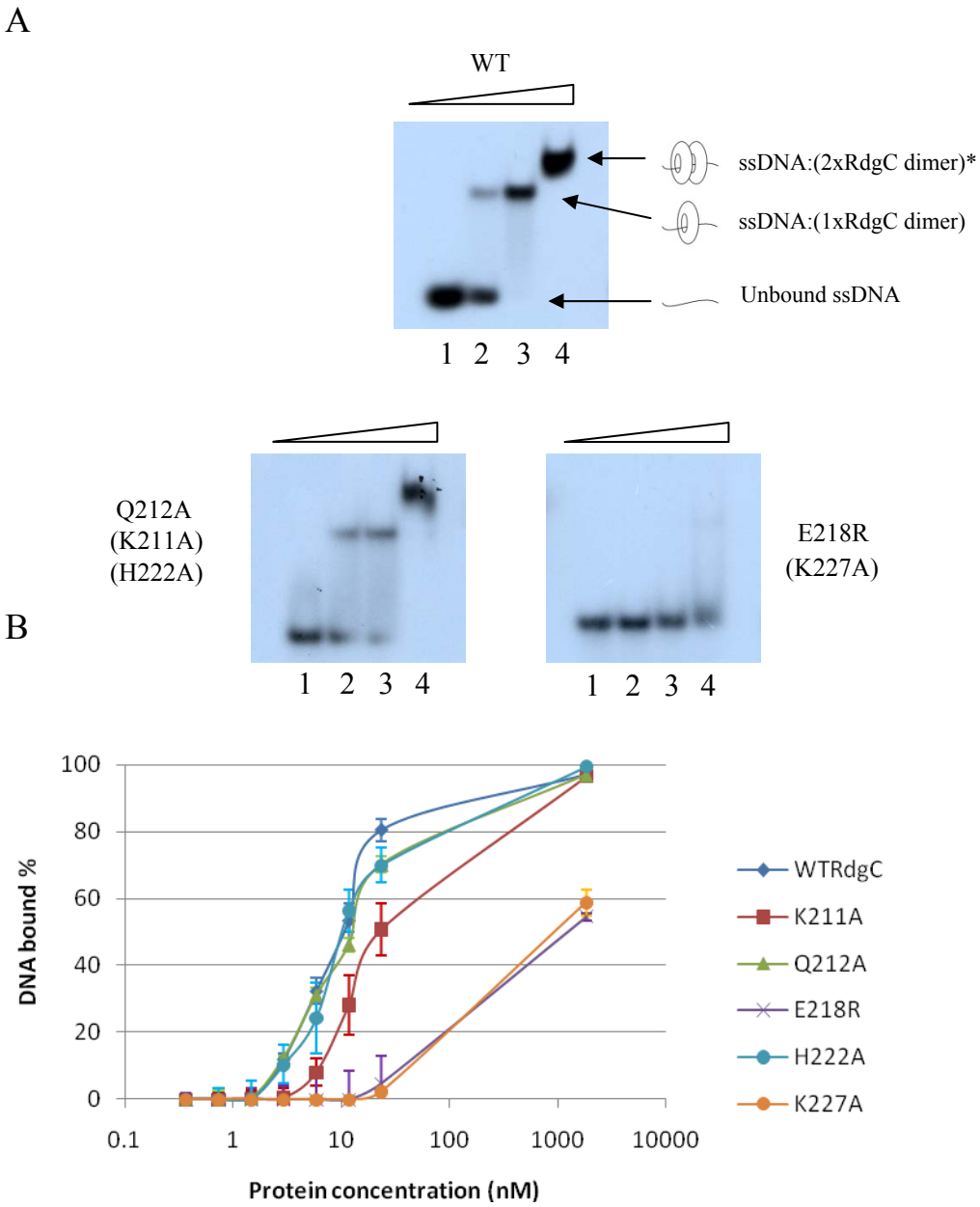


Figure 4-3 RdgC binds ssDNA

A. EMSA assays showing ssDNA binding of wild type RdgC and RdgC mutants with mutation at the horseshoe interface. DNA substrate was RGL13 (49 nt), at 0.2 nM. Protein concentrations: Lane 1, 0 nM; Lane 2, 5.92 nM; Lane 3, 23.68 nM; Lane 4, 1850 nM. $RdgC_{H222A}$ and $RdgC_{K211A}$ had similar shift pattern to the one of $RdgC_{Q212A}$; $RdgC_{K227A}$ had similar shift pattern to the one of $RdgC_{E218R}$.

B. Binding isotherm for ssDNA binding (RGL13) of the RdgC proteins, assembled from at least 2 EMSA's. For the reactions, 0.2 nM RGL13 was used and the protein titrations were at 0, 0.37, 0.74, 1.48, 2.96, 5.92, 11.84, 23.68 and 1850 nM. The error bars stand for standard deviation at each value point.

4.2.4 Double-stranded DNA binding assays

The assays were performed as for ssDNA binding, except that the DNA template was replaced by a linear and blunt-ended dsDNA of 49 bp (RGL13/17).

E. coli RdgC can form stable complex with dsDNA longer than 15 bp (Moore et al., 2003), and therefore can saturate RGL13/17 (49 bp dsDNA) with 3 dimers (Figure 4-4 A, WT). Not surprisingly, the mutants exhibited similar relative DNA affinities for RGL13/17 as for RGL13 (Figure 4-4 A), indicating that RdgC uses the same binding channel for dsDNA and ssDNA. In agreement with previous study (Moore et al., 2003), all of the proteins formed complexes more stable with RGL13/17 than with RGL13.

To be specific, wild type RdgC, $RdgC_{Q212A}$ and $RdgC_{H222A}$ showed signs of shift at 0.3 nM (RdgC dimers:DNA≈2:1); $RdgC_{K211A}$ started at 1 nM (RdgC dimers:DNA≈8:1); $RdgC_{E218R}$ and $RdgC_{K227A}$ started at 6 nM (RdgC dimers:DNA≈30:1) (Figure 4-3 B).

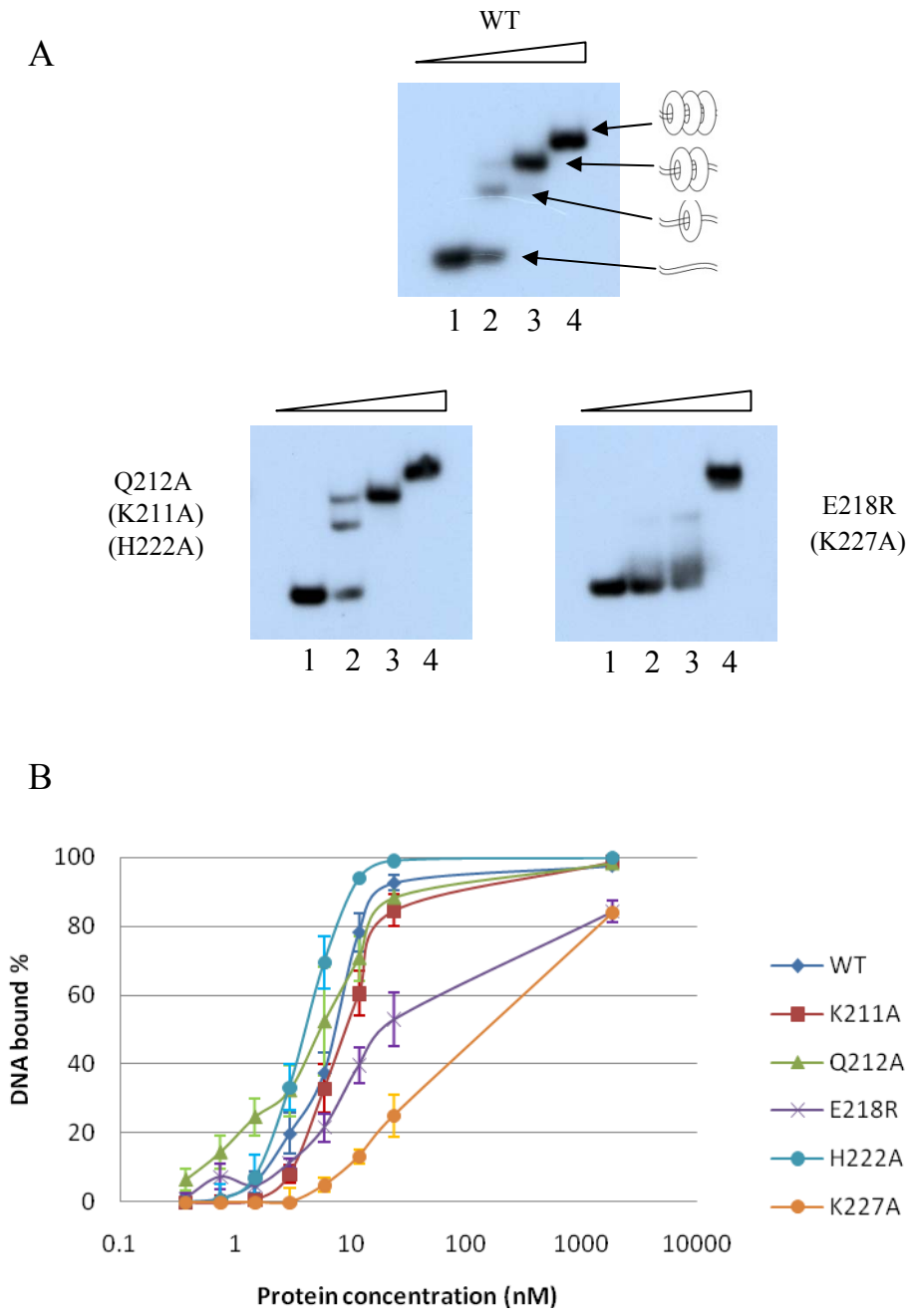


Figure 4-4 RdgC binds dsDNA

A. EMSA assays showing dsDNA binding of wild type RdgC and RdgC mutants with mutation at the horseshoe interface. DNA substrate was RGL13/17 (49 bp), at 0.2 nM. Protein concentrations: Lane 1, 0 nM; Lane 2, 5.92 nM; Lane 3, 23.68 nM; Lane 4, 1850 nM. $RdgC_{H222A}$ and $RdgC_{K211A}$ had similar shift pattern to the one of $RdgC_{Q212A}$; $RdgC_{K227A}$ had similar shift pattern to the one of $RdgC_{E218R}$.

B. Binding isotherm for dsDNA binding (RGL13/17) of the RdgC proteins, assembled from at least 2 EMSAs. For the reactions, 0.2 nM RGL13/17 was used and the protein titrations were at 0, 0.37, 0.74, 1.48, 2.96, 5.92, 11.84, 23.68 and 1850 nM. The error bars stand for standard deviation of each value point.

4.2.5 Circular DNA binding assays

Circular DNA binding assays were performed on 1% agarose gels. A 3000 bp dsDNA plasmid, pGEM-7zf(+) (Promega), was selected as DNA substrate at a fixed concentration value of 4.0 ng/µl. The plasmid DNA was mixed with RdgC protein at 0, 2.3, 4.6, 9.2, 18.5, 37.0, 74.0, 148 and 1590 ng/µl. Note the several weak bands on top of the single sharp strong band in each lane, including the lane free of protein. This was due to the different conformations the plasmid adopted in the reaction buffer, which led to the different migration rates seen on the agarose gels (Figure 4-5).

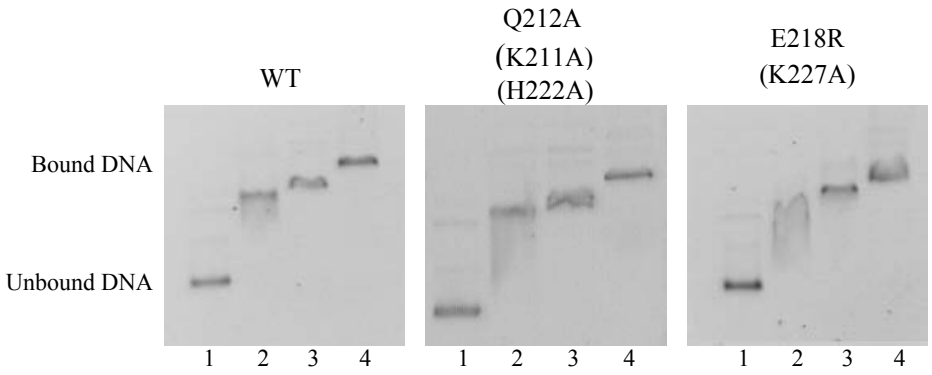


Figure 4-5 RdgC binds circular DNA

The DNA substrate was pGEM-7zf(+) at 4.0 ng/µl. RdgC protein concentrations at Lane 1-Lane 4 were 0, 18.5, 74 and 1590 ng/µl, respectively.

Every horseshoe interface mutant was able to slow the migration rate of pGEM-7zf(+) (Figure 4-5). The data were not quantified due to the limited accuracy of the stain method used and the different conformations the plasmid

adopted in the reaction buffer. However, it is clear that the mutations did not appear to affect circular DNA binding of RdgC to a very large extent.

4.3 *in vivo* synthetic lethality assays

The *in vitro* analyses revealed that the mutant protein retained the ability to bind DNA. To investigate whether they also retain the ability to function *in vivo*, I introduced the mutations into the *E. coli* chromosome at the native site, and exploited a synthetic lethality assay to examine their ability to function.

The assay used an unstable *lac*$^+$ plasmid, pRC7, carrying a wild type gene of interest to cover a deletion allele in the chromosome of a *lac*$^-$ strain. If any further mutation is lethal with the covered mutation, plasmid-free segregants fail to grow and only plasmid carrying cells forming blue colonies are seen on an indicator plate containing IPTG and X-gal. IPTG induces the expression of the *lac* operon on the pRC7 plasmid, and the expression product digests X-gal on the plate, transforming the cells having the plasmid to blue. If the further mutation is not lethal with the covered mutation, plasmid free segregants form healthy white (*lac*$^-$) colonies. The method details can be found in (Mahdi et al., 2006) and in Materials and Method, section 2.2.10.

It is known that deletion of *rdgC* is synthetically lethal in strains with problems in DNA replication restart systems, such as *priA*$^-$, *priA*$^-$ *dnaC810* and *priA*$^-$, *dnaC809*,*820 priC*$^-$ strains (See section 1.5.1 for details). Therefore, a pRC7 derivative that harbours a wild type copy of *priA* (pAM374) was introduced

into strains with the genotype and the wild type *rdgC* gene then replaced with the mutant allele in question to reveal the *in vivo* functions of each mutant RdgC.

Nonetheless, problems caused by impaired DNA repair systems can often be attenuated by slow growth conditions on minimal medium. I therefore also tested all the mutants under this condition.

4.3.1 Strains construction

The three strains for the assays: *priA*$^-$, *priA*$^-$ *dnaC810* and *priA*$^-$, *dnaC809*,*820* *priC*$^-$, each carrying pAM374, were first deleted for *rdgC* using a Δ*rdgC*::*dhfr* (or Δ*rdgC2*::*dhfr*) allele, which encodes a trimethoprim resistance protein. Thereafter, they were transduced with P1 phage grown on strains carrying the mutant *rdgC* gene in question located very tightly to a downstream *cat* cassette encoding resistance to chloramphenicol. Cm^r transductants were purified and checked to confirm that they were sensitive to trimethoprim, but resistant to chloramphenicol, i.e. the Δ*rdgC*::*dhfr* (or Δ*rdgC2*::*dhfr*, where appropriate) allele had been replaced by the *cat* linked mutant *rdgC* gene encoding the protein of interest. Wild type *rdgC* linked to *cat* was also transduced in the same way for positive controls.

4.3.2 RdgC mutants in *priA*$^-$ strains

For convention throughout the thesis, a strain with genotype$_{chromosome}$ harbouring a pRC7 derivative carrying a wild type gene$_{vector}$ is referred to as gene$_{vector}^+$/genotype$_{chromosome}$. The *priA*$^-$ *lac*$^-$ strain having pAM374 was thus

designated *priA*$^{+}$/*priA*$^{-}$ *lac*$^{-}$. Since strains for synthetic lethality assays are always *lac*$^{-}$ in the chromosome, it was omitted in most cases and the strain was therefore also known as *priA*$^{+}$/*priA*$^{-}$.

The *priA*$^{+}$/*priA*$^{-}$ strain (Figure 1-3, N5972) demonstrated the severe problems encountered by the cells without PriA with predominant solid blue colonies and a small number of tiny white colonies on LB indicator plates. This remaining viability was susceptible to further deletion of *rdgC* in the chromosome, as demonstrated by *priA*$^{+}$/*priA*$^{-}$ *rdgC*$^{-}$, with only solid blue colonies viable (Figure 4-5, AM1918). The genes carrying mutant *rdgC* were transduced into *priA*$^{+}$/*priA*$^{-}$ *rdgC*$^{-}$. The wild type *rdgC* gene was also transduced for positive control (Figure 4-6, YJ047). After 2 days of growth on X-gal IPTG LB plates, none of the horseshoe interface mutants appeared to be able to substitute for wild type RdgC in the *priA*$^{+}$/*priA*$^{-}$ strains (Figure 4-6 A), indicating that they had a reduced capacity to function *in vivo*.

Since such morbidity caused by abnormal DNA replication restart machinery may be alleviated under slow growth conditions, the same mutant strains were also examined on minimum media supplemented with X-gal and IPTG. After 3 days of 37 ºC incubation, the positive control YJ047 did produce relatively more white colonies with improved health (54% of overall colonies were white), despite their smaller overall sizes than their LB media cultured counterparts (Figure 4-6 B, YJ047).

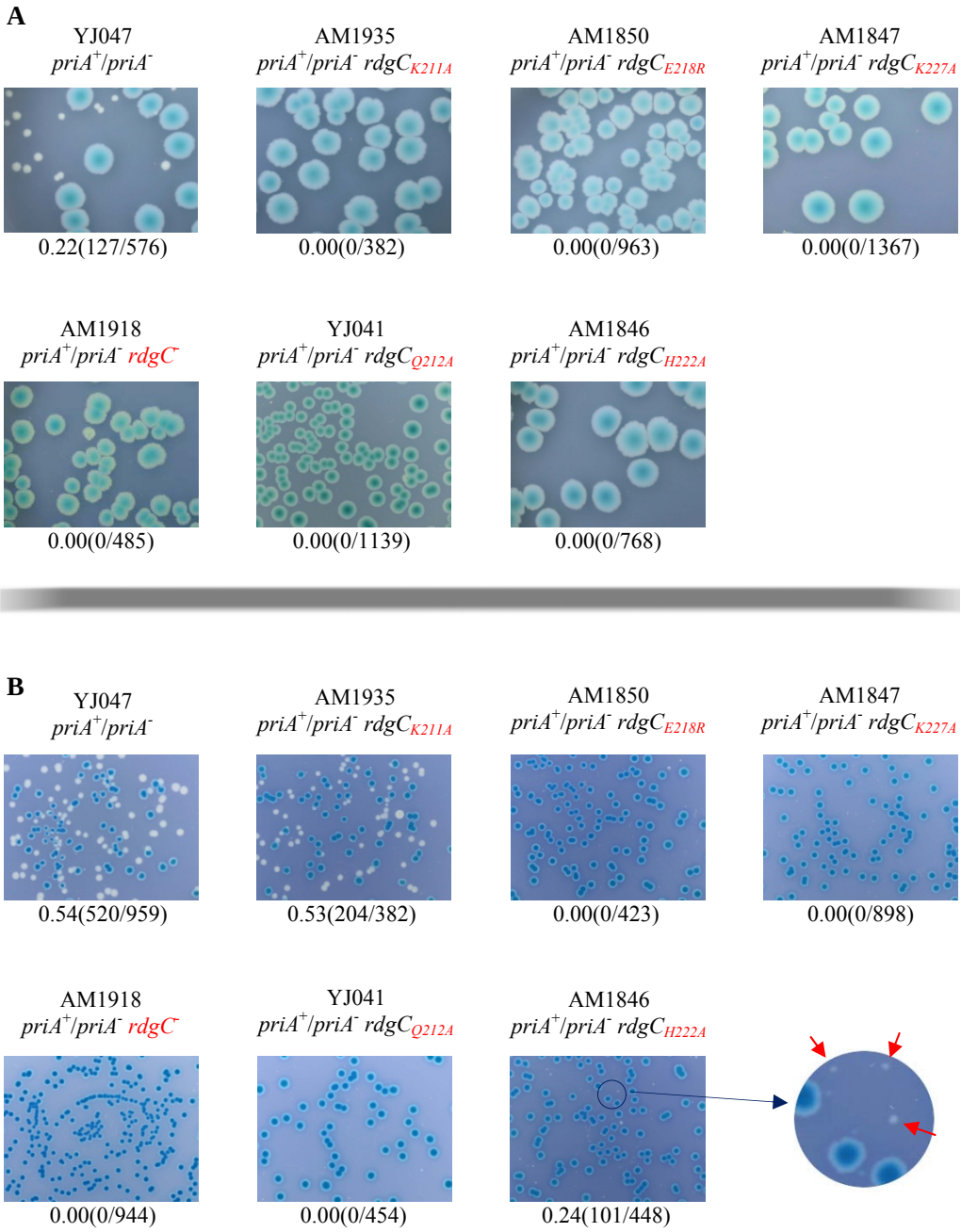


Figure 4-6 Synthetic lethality assays in *priA*$^{+}$/*priA*$^{-}$ for the horseshoe interface mutants. On the top of each picture are numbers and genotypes of the strains; on the bottom are numbers indicating the fraction of white colonies. The actual numbers of white/total colonies are shown in parentheses. This convention is adopted in the ensuing figures.

A. Synthetic lethality assays on LB media containing X-gal and IPTG
B. Synthetic lethality assays on minimum media containing X-gal and IPTG

The circular figure on the bottom, right to AM1846 is a magnification showing tiny white colonies on the plate, as indicated by red arrows.

Interestingly, the horseshoe interface mutants showed different levels of capacity to substitute for wild type RdgC under these conditions of slower growth. The inviability seen in *priA*$^{+}$/*priA*$^{-}$ *rdgC*$^{-}$ could not be rescued by $RdgC_{Q212A}$, $RdgC_{E218R}$ or RdC_{K227A}. $RdgC_{H222A}$ was able to improve viability, and a small fraction of tiny white colonies (24% of white + blue) emerged after 3 days of growth. $RdgC_{K211A}$ was found to function as well as the wild type protein. The growth of the white colonies was almost as good as the wild type in terms of both size and number of colonies (53% in $rdgC_{K211A}$, 54% in $rdgC_{WT}$ were white colonies) (Figure 4-6 B). This ability of the K211A mutant protein shows some activity relative to the other mutants as it is clearly no better at binding DNA than for instance $RdgC_{Q212A}$ or $RdgC_{H222A}$. The failure of $RdgC_{K227A}$ and $RdgC_{E218R}$ to function is not surprising given the reduced affinity for DNA oligos (Figures 4-3 and 4-4).

4.3.3 RdgC mutants in *priA*$^{-}$ *dnaC810* strains

A *priA* null strain is very sick due to its inability to load DnaB at stalled or rescued replication forks. However, a *dnaC810* mutation overcomes this by means which are not fully understood, but which permit DnaB loading in the absence of PriA. In a synthetic lethality assay, this is reflected in improved growth of white colonies (compare strain YJ048 in Figure 4-7 A with strain YJ047 in Figure 4-6 A). However, *dnaC810* cannot rescue cells lacking both PriA and RdgC (Figure 4-7 A, AM1919). Thus, *priA*$^{-}$ *dnaC810* cells provide a very sensitive assay for RdgC activity.

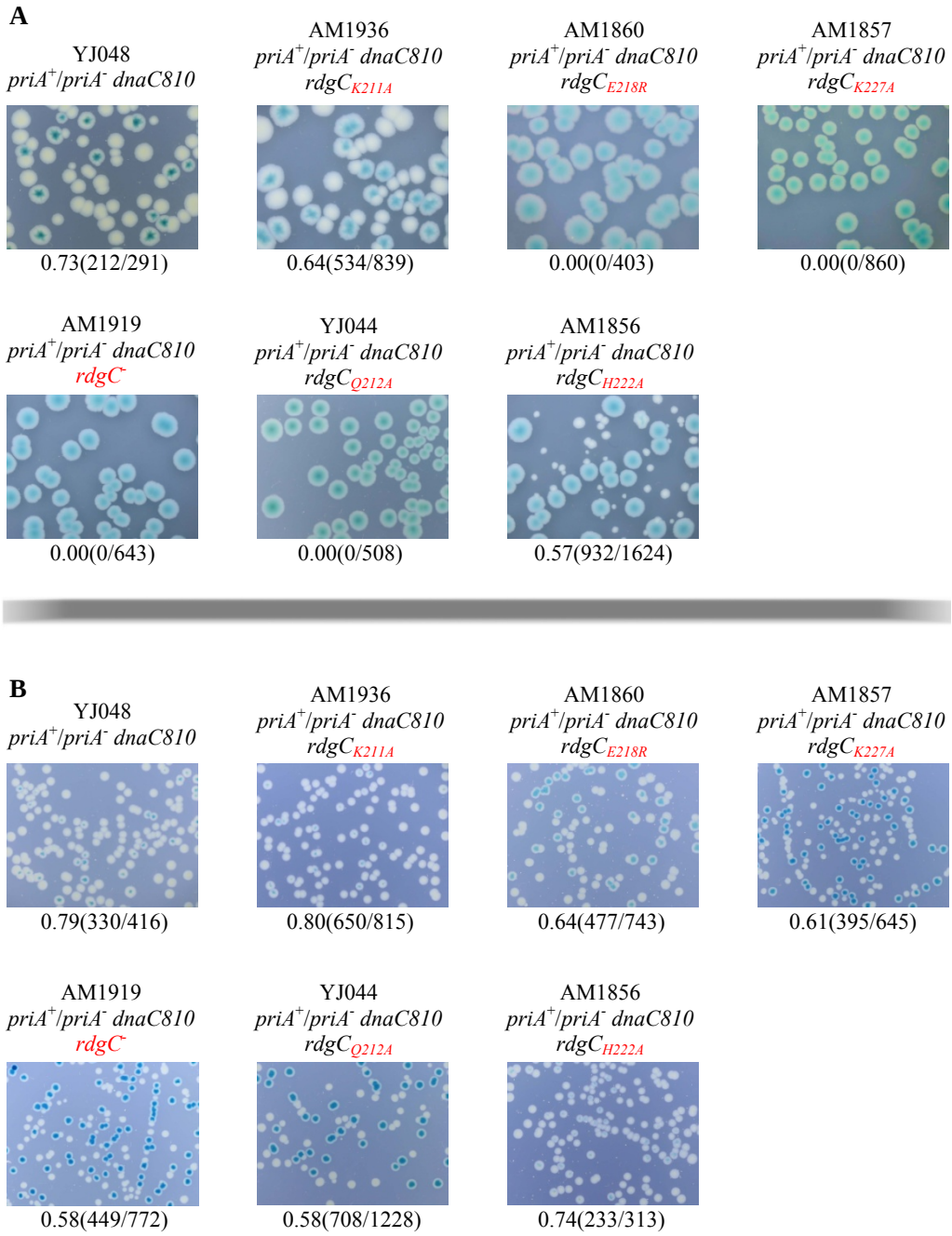


Figure 4-7 Synthetic lethality assays in *priA*$^{+}$/*priA*$^{-}$ *dnaC810* for the horseshoe interface mutants.

A. Synthetic lethality assays on LB media containing X-gal and IPTG
B. Synthetic lethality assays on minimum media containing X-gal and IPTG

In this assay, $RdgC_{K211A}$, and $RdgC_{H222A}$ to a lesser extent, were able to substitute for wild type RdgC to allow growth of plasmid-free segregants (white colonies) on LB indicator plates (Figure 4-7 A).

Rather strikingly, the segregation of healthy white colonies revealed that RdgC is not required to maintain the viability of *priA⁻ dnaC810* cells on minimal medium (Figure 4-7 B, AM1919). It is therefore not surprising that constructs expressing any one of the six mutant RdgC proteins segregate healthy white colonies, i.e. the assay on minimal agar is non-informative in this case.

4.3.4 RdgC mutants in *priA⁻ dnaC809, 820 priC⁻* strains

As described earlier, *dnaC809,820* is a super-suppressor of *priA⁻* that enables it to grow without PriC, provided RdgC is present (Figure 1-3; R. G. Lloyd, unpublished). On LB indicator plates, a *priA⁺/priA⁻ dnaC809,820 priC⁻* construct segregates solid blue colonies accompanied by smaller white colonies. However, without wild type RdgC present, it forms only blue colonies (Figure 4-8 A, strain AM2046 and YJ026). The same is true on minimal indicator medium (Figure 4-8 B).

All of the horseshoe interface mutants appeared to have little activity in this background, behaving much the same as in a *priA⁺/priA⁻* background (compare Figure 4-8 A with Figure 4-6 A). The strain encoding $RdgC_{K211A}$ (YJ027) produces some white colonies (12%, compared with 44% with wild type RdgC, but these are extremely tiny, Figure 4-7 A). Again, as with the constructs in the *priA⁺/priA⁻* background (Figure 4-6 B), the $RdgC_{K211A}$ and $RdgC_{H222A}$ proteins showed some activity in the *priA⁺/priA⁻ dnaC809,820 priC* background when

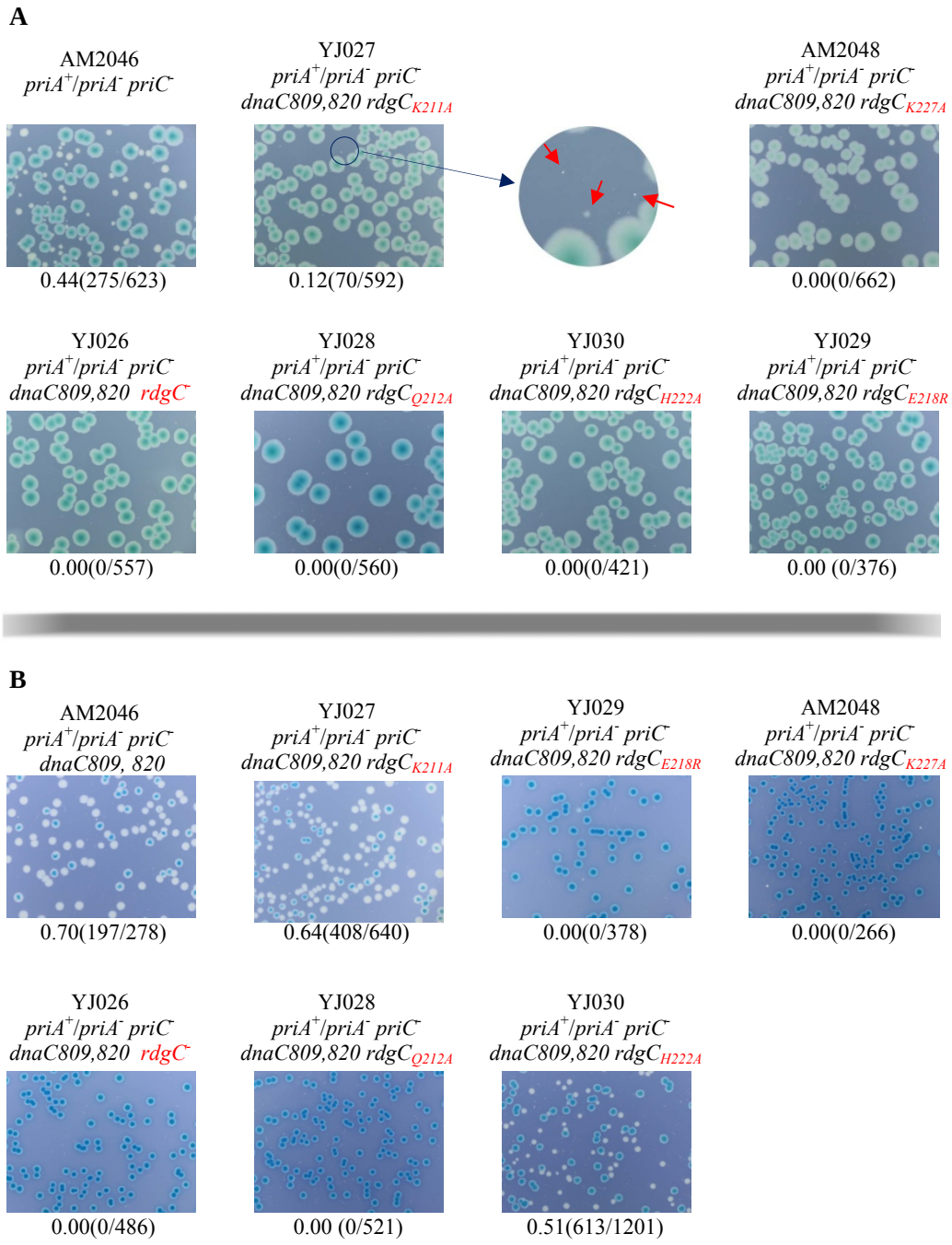


Figure 4-8 Synthetic lethality assays in *priA*$^{+}$/*priA*$^{-}$ *dnaC809,820* for the horseshoe interface mutants.

A. Synthetic lethality assays on LB media containing X-gal and IPTG
B. Synthetic lethality assays on minimum media containing X-gal and IPTG

The circular figure right to YJ047 is a magnification showing tiny white colonies on the plate, as indicated by red arrows.

plated on minimal agar (Figure 4-8 B), but none of the others showed any activity.

4.4 Discussion

4.4.1 Some mutations at the horseshoe interface negatively affected RdgC ability to bind DNA

Wild type RdgC is able to bind ssDNA, dsDNA and circular plasmid DNA substrates with high affinities. Binding is weakened by some of the mutations of conserved residues at the horseshoe interface. Wild type RdgC and most of the mutants represented a sigmoidal binding curve in dsDNA and ssDNA binding assays, indicating some level of cooperativity of DNA binding. The non-sigmoidal binding curves of $RdgC_{K227}$ and $RdgC_{E218R}$ to ssDNA (Figure 4-3) and $RdgC_{K227}$ to dsDNA (Figure 4-4) should be dealt with with care, since the data for the mutants were not comprehensive enough to draw a faithful binding curve. In fact, Moore's research suggested a very weak cooperative DNA binding of RdgC by using a Scatchard plot of binding (Moore, 2002). Similar situations were found for most of the rest of the mutants and will not be iterated.

The change in E218 or K227 appeared to have the most severe effect, with much reduced affinity for ssDNA and dsDNA. Nevertheless, both of the mutants were still able to bind ssDNA and dsDNA at high concentrations. However, there were obvious smears sometimes observed tailing the high concentration protein shift bands (Figure 4-3 A, E218R and data not shown).

This is a sign of unstable protein:DNA contact that caused continuous dissociation when the complex was shifting in the gel. This suggests the two residues are important to hold the horseshoe tightly, which is crucial for linear oligo binding.

The change in Q212 or H222 ($RdgC_{Q212A}$ and $RdgC_{H222A}$) did not appear to disturb RdgC's DNA binding ability in all three experiments. It appears that Q212 is not as predicted to interact with E218 and K227 to form triads and the two H222 do not form a hydrogen bond, the interactions are less important for DNA binding than expected, or the assays failed to detect the existing DNA binding defect brought by Q212A and H222A.

The mutation on K211 slightly reduced RdgC's affinity for dsDNA and ssDNA, indicating the residue is hardly involved in the DNA backbone contact.

It is also worth to note that no gradual change in ssDNA binding assays was observed from 1 bound dimer per DNA molecule to 2 dimers per DNA molecule, as I saw in dsDNA binding assays (compare Figure 4.3 A to Figure 4.4 A). It is therefore possible that the slower migration rate was due to 3 or more dimers complexed with 1 DNA molecule. However, the shift and the minimum length requirement for stable ssDNA binding (23 nt, Moore et al., 2003) support that the higher molecular weight species reflects 2 dimers complexed with 1 DNA molecule. This is also applied to the rest of the mutants described in following two chapters.

Circular DNA binding did not appear to be much affected by the horseshoe interface mutations. Taken together with the ssDNA and dsDNA assays, it

appears that the instability with short linear oligos is because the dimer ring with the mutations could not bind tightly and easily slide off the oligos. Once they attach to circular DNA, they move freely along it like beads on a necklace.

4.4.2 The mutations at the horseshoe interface interrupted RdgC's *in vivo* functions

In all the three strain backgrounds tested, *priA*$^{+}$/*priA*$^{-}$ (YJ047) and *priA*$^{+}$/*priA*$^{-}$ *dnaC810* (AM2046) appeared to provide the most stringent test of RdgC activity.

From the assays with these strains, it could be concluded that the Q212, E218 and K227 substitutions inactivated RdgC's ability to support growth of the disabled strains. $RdgC_{H222A}$ did retain some function, but it was so weak that in most conditions tested the strain having it either had no improvement or improved only slightly. $RdgC_{K211A}$ was surprisingly to be the best able to support growth, given its medium rank in the *in vitro* DNA binding assays (4.4.1). However, even this protein is clearly not fully functional.

Chapter 5
The gate interface

5.1 Background

The gate interface is supposed to be held together primarily by hydrophobic interactions by sitting a phenylalanine (F120) residue from one monomer into a hydrophobic pocket on the other monomer. An additional hydrogen bond formed between R118 from each monomer completes the interface (Figure 5-1). This is different from the way the horseshoe interface is maintained. This asymmetry might imply different functionalities of the two interfaces.

This chapter focuses on the studies on the gate interface.

5.2 *In vitro* DNA binding assays

5.2.1 Mutant design

$RdgC_{F120S}$ and $RdgC_{F120T}$ were constructed to disrupt the hydrophobic interactions that are postulated to hold the gate interface closed. $RdgC_{R118A}$ and $RdgC_{R118C}$ were designed in the hope to study the function of the hydrogen bond at the interface by either losing it ($RdgC_{R118A}$), or further strengthening it ($RdgC_{R118C}$). The cysteine residues introduced in both side chains were expected to form a disulphide bond, a covalent interaction which is more stable than the wild type's hydrogen bond. R118→C, F120→T double mutant was also generated to better understand the hydrophobic interactions.

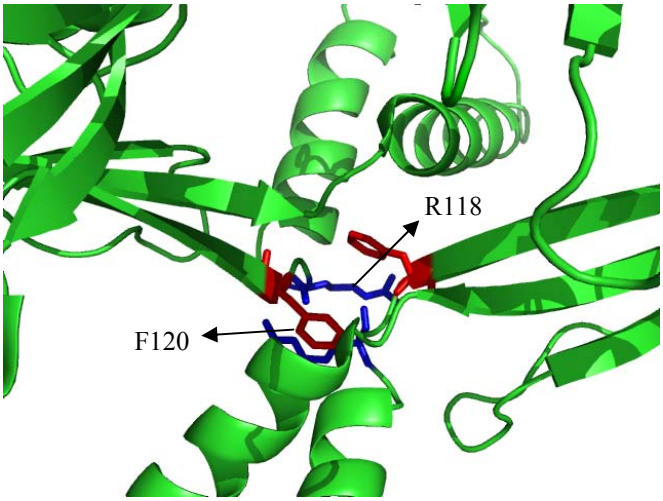


Figure 5-1 Ribbon diagram illustrating the gate interface of RdgC dimer and the residues participated in mutant studies. Residues shaded in red are F120; residues shaded in blue are R118.

5.2.2 Cloning, overexpression and purification

The pYJ001 or pGS853 derivatives having the mutations of interests were designed and created as described in Section 2.2.7. Overexpression and purification of the proteins were performed as for the horseshoe interface mutants (4.2.2.2, 4.2.2.3) (Figure 5-2).

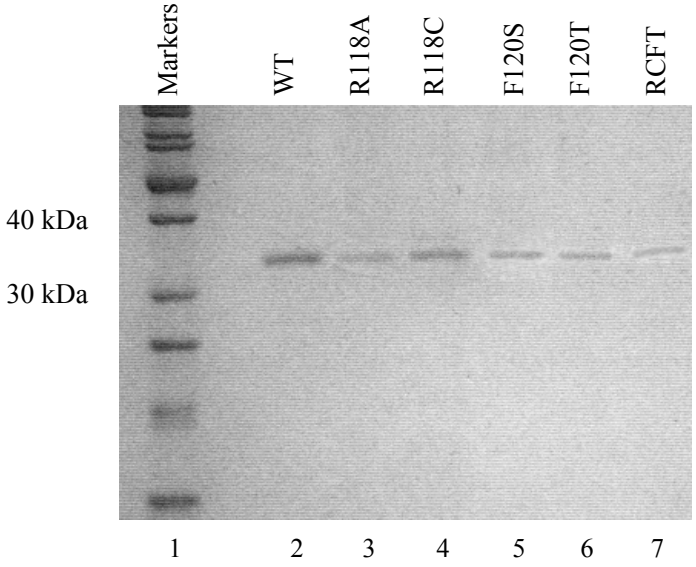


Figure 5-2 12.5% SDS-Page gels showing purified mutant RdgC proteins with *E. coli* native RdgC

Lane 1, Molecular weight markers; Lane 2, wild type RdgC; Lane 3-7, RdgC proteins with mutations as labelled. RCFT=R118C·F120T.

5.2.3 Single-stranded DNA binding assays

As for the horseshoe interface mutants, the assays were done on RGL13, a 49 nt ssDNA (4.2.3).

Removing the hydrogen bond by R118A, and the R118C disulphide bond did not seem to affect RdgC's ssDNA binding (Figure 5-3). The mutations in F120 (F120T, F120S) severely damaged its ssDNA binding (Figure 5-3). Only at their high concentration did $RdgC_{F120T}$ and $RdgC_{F120S}$ show a sign of DNA binding, but also accompanied by massive dissociations (Figure 5-3 A).

However, this defect could probably be overcome by the additional R118C disulphide bond (Figure 5-3 A R118C·F120T). The protein with double mutations formed sharp band with the ssDNA on the gel, suggesting the protein:DNA complex was very stable. The binding isotherm also showed an obvious sign of ssDNA binding improvement (Figure 5-3 B). Nonetheless, its slightly weaker affinity indicates that the disulphide bond could not fully recover the loss of the gate interface hydrophobic interactions. To be specific, wild type RdgC, $RdgC_{R118C}$, $RdgC_{R118A}$ and $RdgC_{R118C·F120T}$ started to shift 0.2 nM RGL13 at 1 nM (~5 RdgC dimers per DNA molecule); $RdgC_{F120T}$ and $RdgC_{F120S}$ barely started at 20 nM (~100 RdgC dimers per DNA molecule) (Figure 5-3 B).

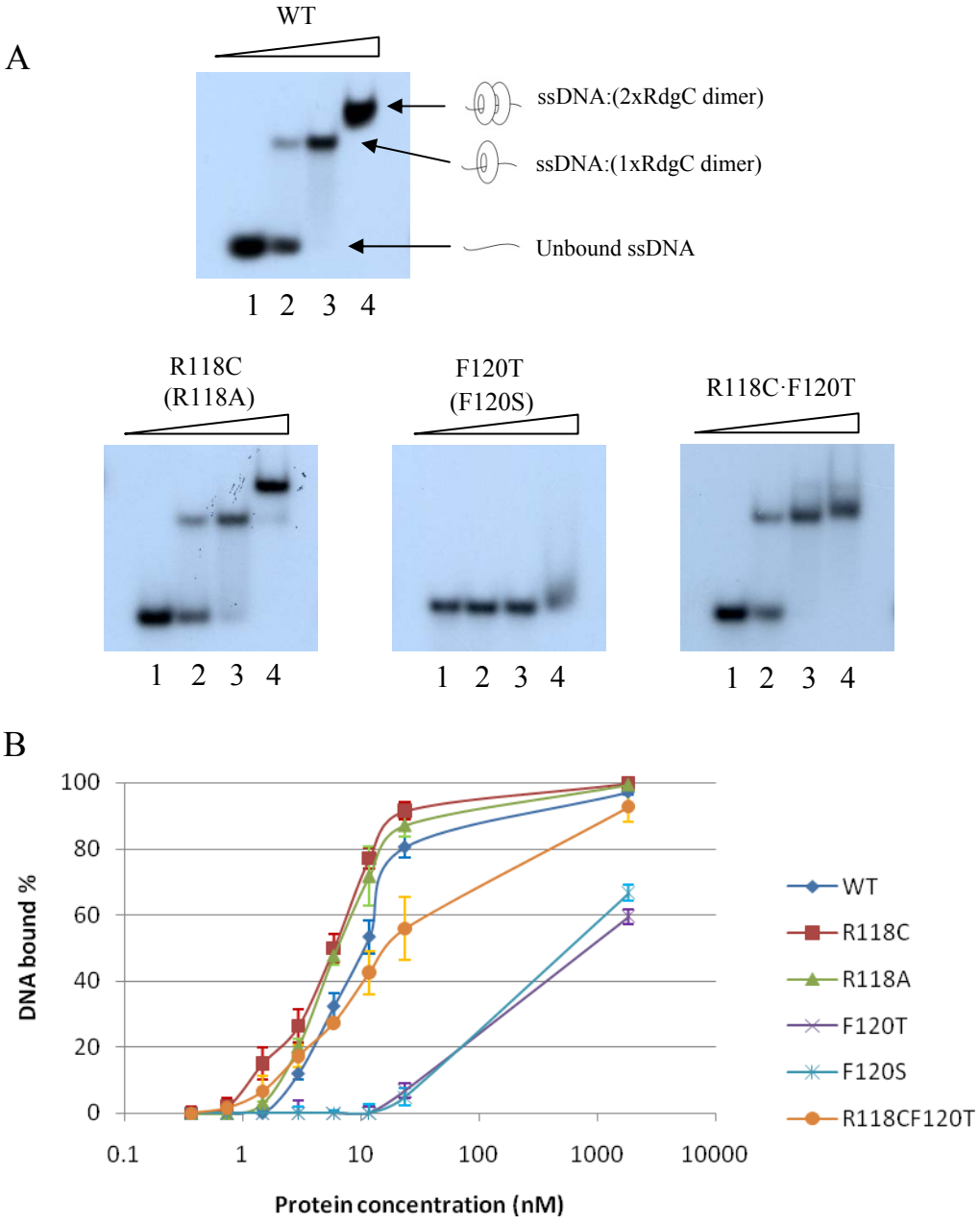


Figure 5-3 RdgC binds ssDNA

A. EMSAs showing ssDNA binding of wild type RdgC and the gate interface mutants. DNA substrate was RGL13 (49 nt), at 0.2 nM. Protein concentrations: Lane 1, 0 nM; Lane 2, 5.92 nM; Lane 3, 23.68 nM; Lane 4, 1850 nM. $RdgC_{R118A}$ had similar shift pattern to the one of $RdgC_{R118C}$; $RdgC_{F120S}$ had similar shift pattern to the one of $RdgC_{F120T}$.

B. Binding isotherm for ssDNA binding (RGL13) of the RdgC proteins, assembled from 3 EMSA's. For the reactions, 0.2 nM RGL13 was used and the protein titrations were at 0, 0.37, 0.74, 1.48, 2.96, 5.92, 11.84, 23.68 and 1850 nM.

5.2.4 Double-stranded DNA binding assays

As for the horseshoe mutants, the assays were done on RGL13/17 (4.2.4).

All the gate mutants showed similar preference for dsDNA as for ssDNA. Neither the elimination of the gate hydrogen bond ($RdgC_{R118A}$) nor its replacement with a disulphide bond ($RdgC_{R118C}$) showed any obvious effect on dsDNA binding. The mutations in F120 ($RdgC_{F120T}$ and $RdgC_{F120S}$) proved to be detrimental to dsDNA binding. As expected, the R118C disulphide bond helped improve $RdgC_{F120T}$ dsDNA binding, and $RdgC_{R118C\cdot F120T}$ only showed a moderately decreased affinity for dsDNA (Figure 5-4 and 5-5).

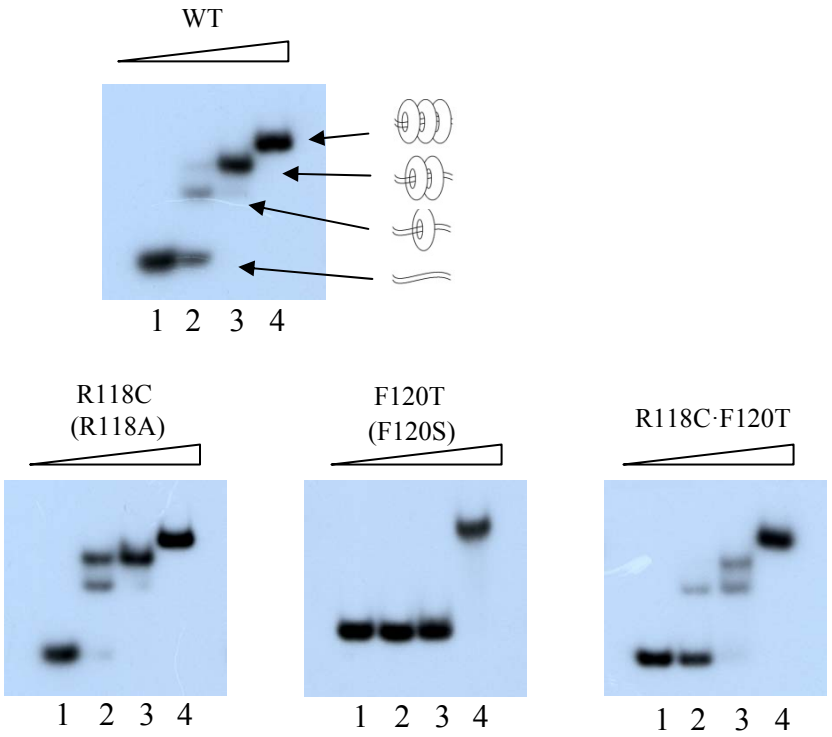


Figure 5-4 EMSA assays showing dsDNA binding of wild type RdgC and RdgC mutants with mutation at the gate interface.

DNA substrate was RGL13/17 (49 bp), at 0.2 nM. Protein concentrations: Lane 1, 0 nM; Lane 2, 5.92 nM; Lane 3, 23.68 nM; Lane 4, 1850 nM. $RdgC_{R118A}$ had similar shift pattern to the one of $RdgC_{R118C}$; $RdgC_{F120S}$ had similar shift pattern to the one of $RdgC_{F120T}$.

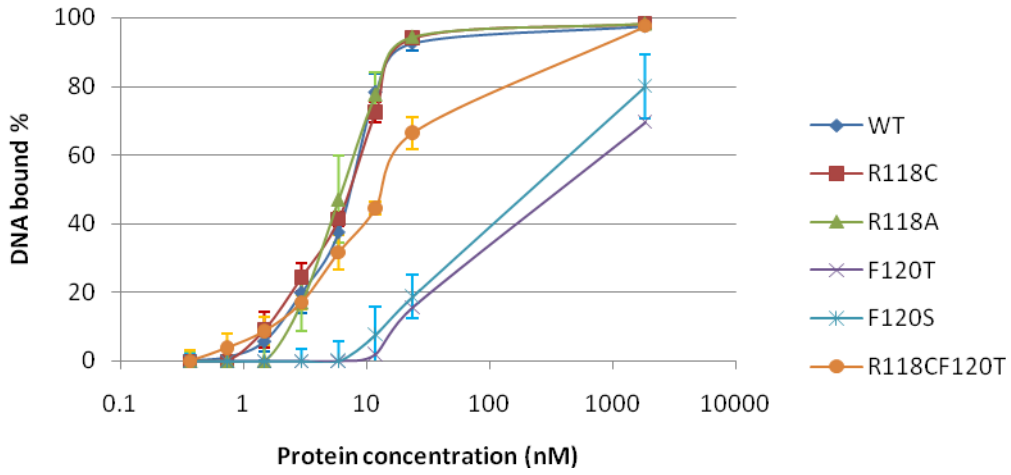


Figure 5-5 Binding isotherm for dsDNA binding (RGL13/17).

Data were assembled from 3 EMSAs. For the reactions, 0.2 nM RGL13/17 was used and the protein titrations were at 0, 0.37, 0.74, 1.48, 2.96, 5.92, 11.84, 23.68 and 1850 nM.

To be specific, wild type RdgC and $RdgC_{R118C\cdot F120T}$ started to shift 0.2 nM RGL13/17 at 0.3 nM (~2 RdgC dimers per DNA molecule); $RdgC_{R118C}$ started at 1 nM; $RdgC_{R118A}$ started at 1.5 nM; $RdgC_{F120T}$ and $RdgC_{F120S}$ barely started at 10 nM (~60 RdgC dimers per DNA molecule) (Figure 5-5 B).

5.2.5 Circular DNA binding assays

As in 4.2.5, the assays used the 3000 bp pGEM-7zf(+) as the DNA substrate at a fixed concentration value of 4.0 ng/µl. The plasmid DNA was mixed with RdgC protein at 0, 18.5, 74.0 and 1590 ng/µl (Figure 5-6).

Similar to the results in 4.2.5, the effects on circular DNA binding of the mutations at the gate interface were not obvious. Even with much impaired affinity for short linear oligos, $RdgC_{F120T}$ and $RdgC_{F120S}$ could still make stable contact with the plasmid DNA, making sharp shift bands on the gel (Figure 5-

6). The additional R118C disulphide bond did not alter the protein's affinity for the plasmid DNA to any obvious extent.

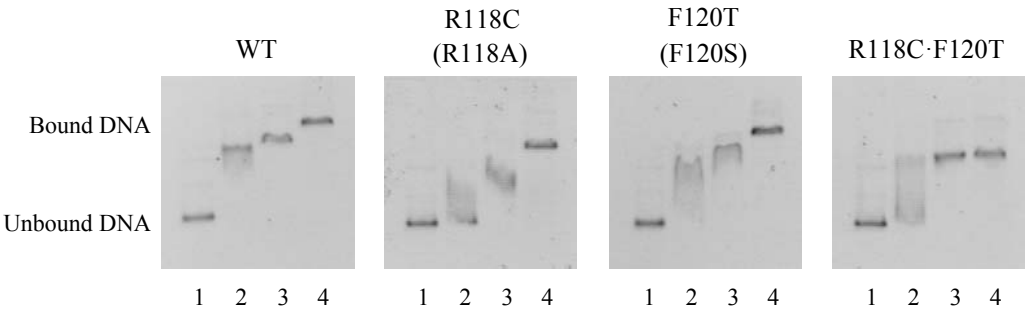


Figure 5-6 RdgC binds circular DNA

The DNA substrate was pGEM-7zf(+) at 4.0 ng/µl. Protein concentrations at Lane 1-Lane 4 were 0, 18.5, 74 and 1590 ng/µl, respectively.

5.3 *in vivo* synthetic lethality assays

As in 4.3, all the gate interface mutants were tested for their *in vivo* activity in strains using synthetic lethality constructs of the three genotypes.

5.3.1 RdgC mutants in *priA*$^-$ strains

None of the gate mutants were able to substitute for wild type RdgC in the strain background on LB indicator plates, since plasmid-free segregants were inviable (no white colonies; see Figure 5-7 A). This is almost also true on minimal agar, where most of the mutants did not segregate any white colonies (Figure 5-7 B). However, $RdgC_{R118C}$ and $RdgC_{R118A}$ were able to substitute for wild type RdgC in this strain background on minimal agar, reflected by white

colonies similar to the wild type on minimal indicator plates (Figure 5-7 B, AM1843 and AM1848). This could suggest that only under slow growth conditions can these two mutants function as wild type RdgC in this strain background; this function is lost if the cells are under stressful DNA repair pressure in fast growth conditions.

5.3.2 RdgC mutants in *priA⁻ dnaC810* strains

In this strain background, only $RdgC_{R118C}$ and $RdgC_{R118A}$ mutants segregated white colonies on LB indicator plates. Nonetheless, given the number of the white colonies and the white sectors circling the blue colonies, the mutants appeared to be fully functional in *priA⁻ dnaC810* strain background (Figure 5-8 A, AM1853 and AM1858).

Similar to 4.3.3, this strain background did not yield informative data on the *in vivo* function of the mutants under slow growth conditions, as all constructs segregated white colonies on minimal indicator plates (Figure 5-8 B).

5.3.3 RdgC mutants in *priA⁻ dnaC809, 820 priC⁻* strains

Very similar to the scenario seen in a *priA⁺*/*priA⁻* background, all the tested gate interface mutants ($RdgC_{R118C}$, $RdgC_{R118A}$, $RdgC_{F120T}$, $RdgC_{F120S}$ and $RdgC_{R118C\cdot F120T}$) failed to substitute for wild type RdgC in *priA⁻ dnaC809,820 priC⁻* background on LB indicator plates (compare Figure 5-7 A with Figure 5-9 A). Only the strains encoding $RdgC_{R118C}$ (YJ033) and $RdgC_{R118A}$ (AM2049) developed white colonies on minimal indicator plates (Figure 5-9 B).

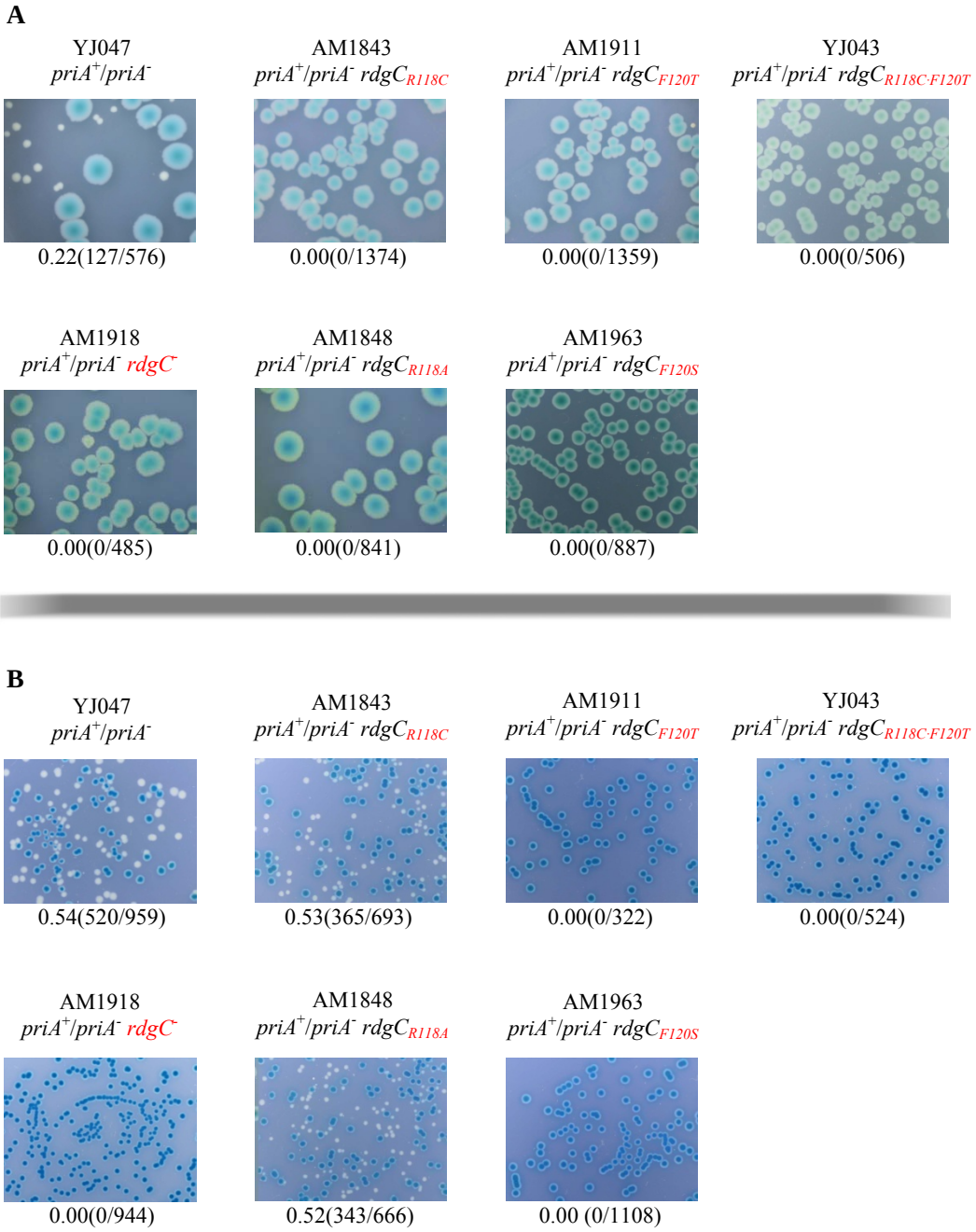


Figure 5-7 Synthetic lethality assays in *priA*$^{+}$/*priA*$^{-}$ for the gate interface mutants.

A. Synthetic lethality assays on LB media containing X-gal and IPTG
B. Synthetic lethality assays on minimal media containing X-gal and IPTG

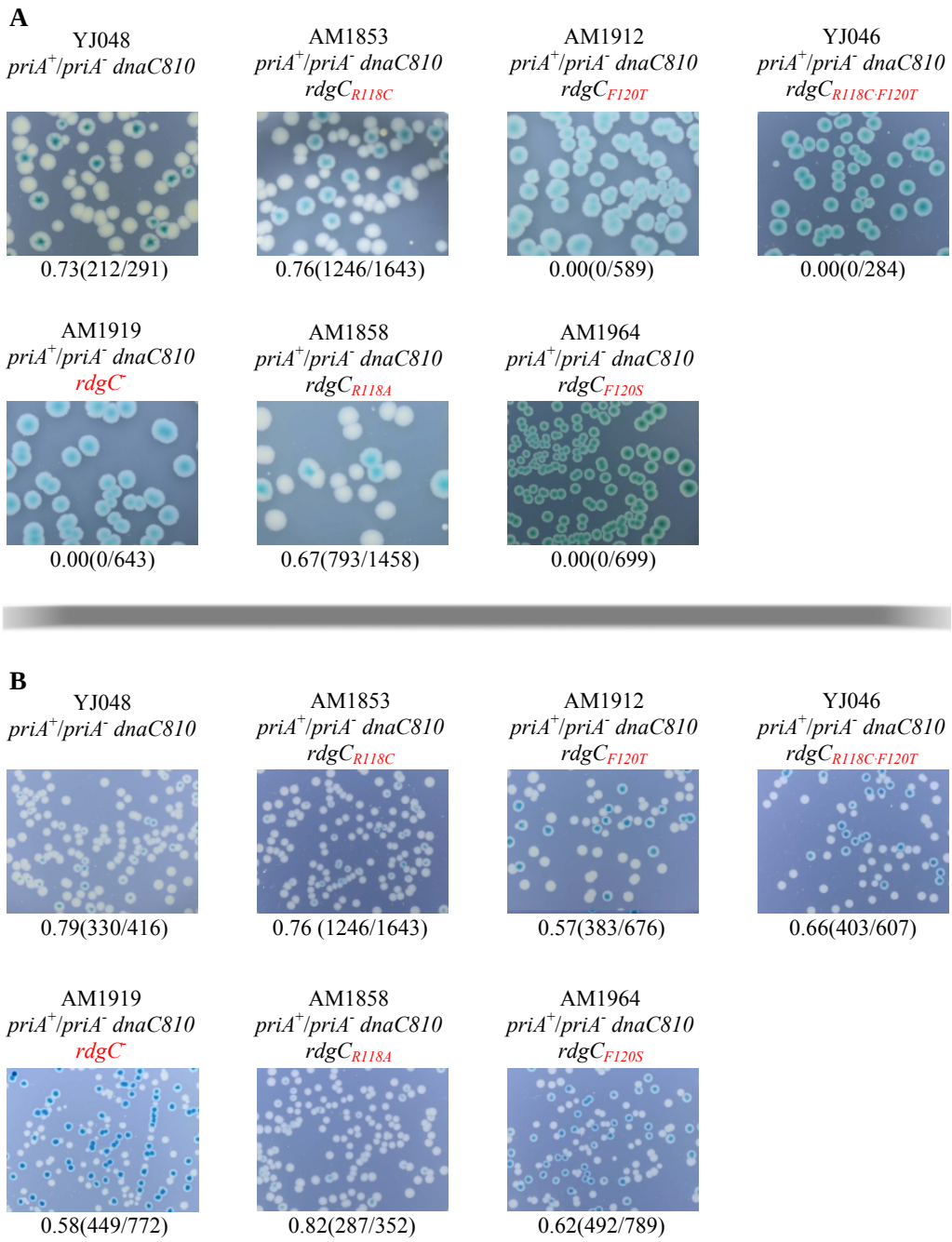


Figure 5-8 Synthetic lethality assays in *priA*$^{+}$/*priA*$^{-}$ *dnaC810* for the horseshoe interface mutants.

A. Synthetic lethality assays on LB media containing X-gal and IPTG
B. Synthetic lethality assays on minimal media containing X-gal and IPTG

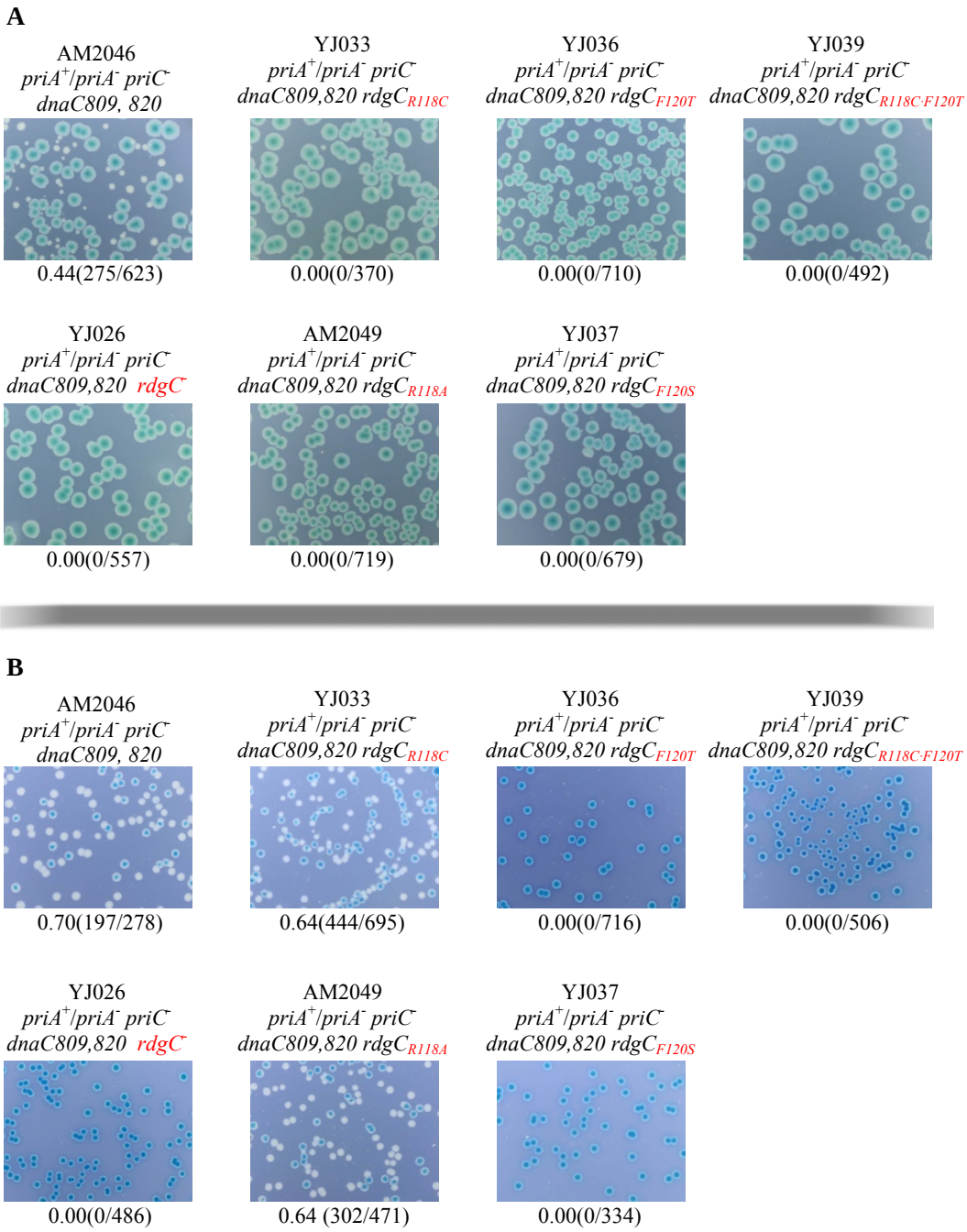


Figure 5-9 Synthetic lethality assays in *priA*$^{+}$/*priA*$^{-}$ *dnaC809,820* for the gate interface mutants.

A. Synthetic lethality assays on LB media containing X-gal and IPTG
B. Synthetic lethality assays on minimal media containing X-gal and IPTG

5.4 Discussion

5.4.1 The hydrophobic interactions are crucial for DNA binding

The DNA binding assays on the gate interface mutants revealed the key role of the hydrophobic interactions in RdgC's DNA binding activity. Both of the mutants ($RdgC_{F120T}$ and $RdgC_{F120S}$) lost stable contact with ssDNA and dsDNA. The other force that helps hold together the gate interface is the hydrogen bond formed between R118 from each side chains. Neither its removal ($RdgC_{R118A}$) nor replacement by a disulphide bond ($RdgC_{R118C}$) appeared to affect its DNA binding in any aspect. It is thus possible that the R118 hydrogen bond is not critical for holding the dimer interface shut.

An artificial gate interface was made by replacing the hydrophobic interactions with a disulphide bond ($RdgC_{R118C\cdot F120T}$). This mutant had an obvious but only moderate decrease of binding on short DNA oligos (5.2.3 and 5.2.4). Compared to $RdgC_{F120T}$, the added disulphide bond clearly improved its DNA binding; but its relatively weak binding indicates it could not fully substitute for the gate interface hydrophobic interactions. Another possibility is that the two C118 did not interact to form a disulphide bond as predicted. The enhanced affinity for some DNA structures could reflect an improvement in the gate strength; but this might attribute to a more hydrophobic environment brought by the R118C, rather than a disulphide bond.

5.4.2 The mutations at the gate interface interrupted RdgC's *in vivo* functions

The removal of the gate interface hydrophobic interaction (F120T and F120S) was clearly destructive to the protein's function, as the mutant proteins could not substitute for wild type RdgC in *in vivo* assays (5.3). The changes in the gate interface hydrogen bond, however, were less destructive to its *in vivo* function. When cells were under limited pressure (e.g. on minimal media, or in *priA*$^{-}$ *dnaC810*), the loss/stronger-replacement of the R118 hydrogen bond did not develop any obvious difference from the wild type. Only in cells under serious pressure did the defect of the changes in the R118 hydrogen bond emerge (e.g. in *priA*$^{-}$ or *priA*$^{-}$ *dnaC809*,*820 priC*$^{-}$). This suggests some special functionality lying in the hydrogen bond.

Although the R118C disulphide bond could help recover DNA binding of RdgC with interrupted gate hydrophobic interactions ($RdgC_{R118C\cdot F120T}$, see 5.4.1), its assistance *in vivo* was insubstantial. This again reinforces the idea that the gate interface hydrophobic interactions are vital for the protein's functionality.

Chapter 6
The finger domains

6.1 Background

The finger domains consist of two anti-parallel α helices that project perpendicular to the ring on either side of the dimer interface at the gate (Figure 6-1). They span from S77 to L115. These domains are special because: a) they have high B factors in the crystal structure; b) a high proportion of conserved lysines on the fingers cause a highly electropositive upper surface and a basic tip at the turn of the two anti-parallel helices. The questions to this point are: do the finger domains assist DNA binding, and do they possess any other functions?

This chapter attempts to find the answers by using the same *in vitro* and *in vivo* assays as used for the interfaces.

6.2 *In vitro* DNA binding assays

6.2.1 Mutant design

Since the finger domains are not part of the dimer ring structure, the complete deletion of them is not supposed to largely affect the protein folding. $RdgC_{\Delta F}$ was made by changing P76 to glycine and L116 to threonine, and the residues in between deleted. Because some of the altered residues are components of the hydrophobic pocket that helps hold together the gate interface (3.4), the R118C

disulphide bond was also created together with the finger deletion to achieve an artificial gate lock ($RdgC_{\Delta F \cdot R118C}$).

Given the flexibility and highly electropositive charge of the finger domains, they are hypothetically involved in RdgC's first touch on its binding substrates. The frontier is the basic fingertip, consisting of an arginine (R97) and three lysines (K98, 100, 101). I therefore hoped to reveal the functionality of the finger domains by $RdgC_{FTM}$ (FTM stands for **F**inger **T**ip **M**utations; they are R97S, K98Q, K100Q and K101E).

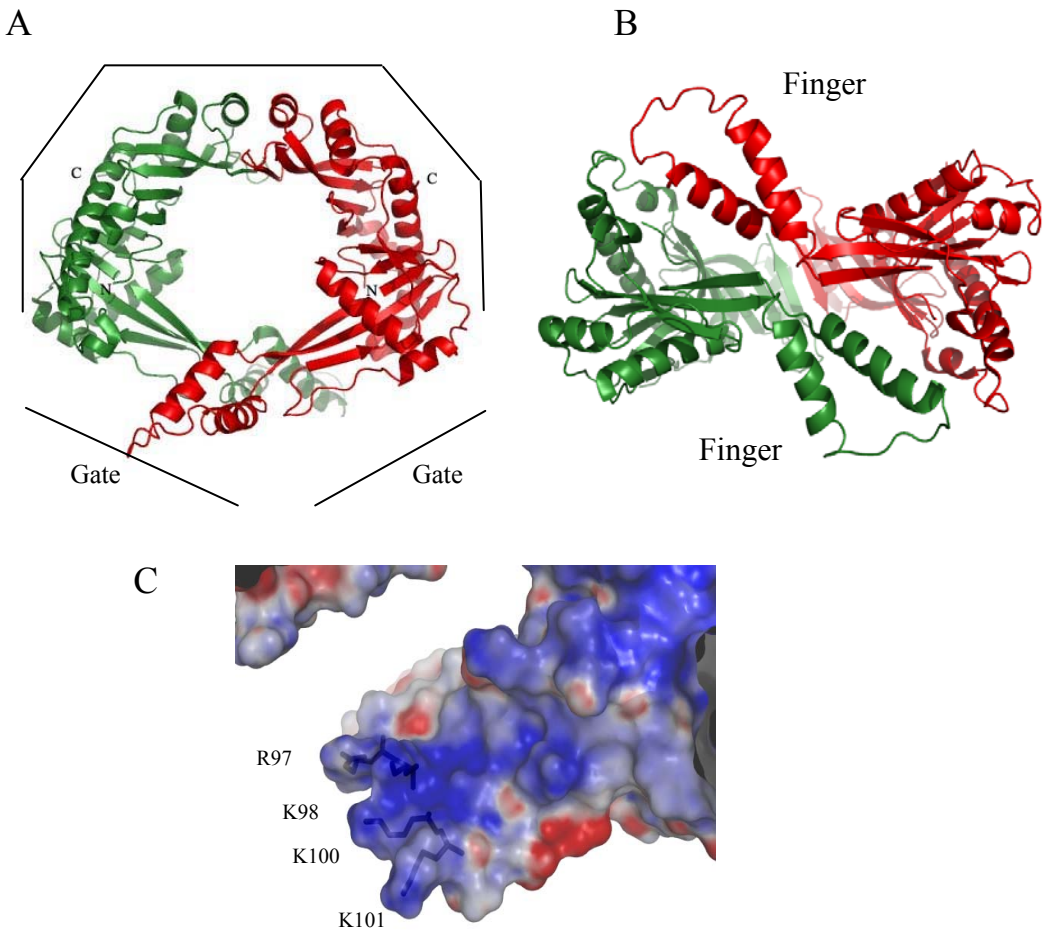


Figure 6-1 Overall structure of RdgC showing the three different domains and one of its finger tips

A. Ribbon diagram showing the ring structure of the dimer.
B. Rotated ribbon diagram showing the finger domains
C. An electrostatic surface representation showing the four basic residues involved in the mutant assays.

6.2.2 Cloning, overexpression and purification

The pYJ001 or pGS853 derivatives having the mutations of interests were designed and created as described in Section 2.2.7. Overexpression and purification of the proteins were performed as for the horseshoe interface mutants (4.2.2.2, 4.2.2.3) (Figure 6-2). $RdgC_{\Delta F}$ and $RdgC_{\Delta F \cdot R118C}$ showed a much decreased affinity for the heparin column (Appendix 1).

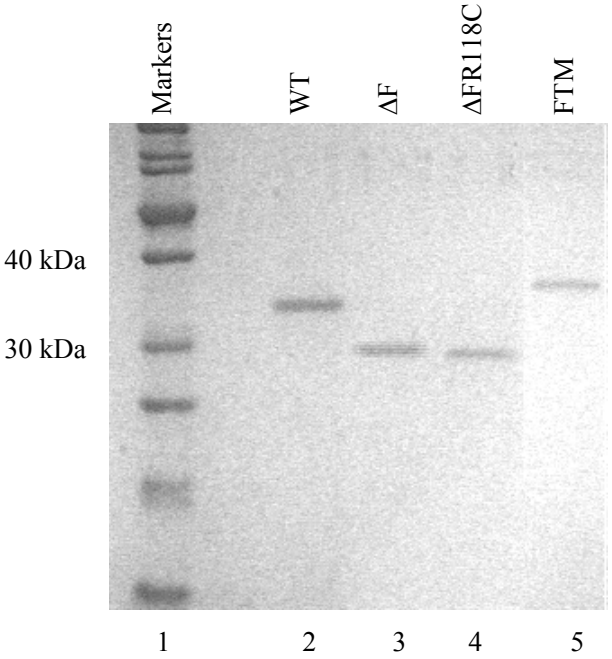


Figure 6-2 12.5% SDS-Page gels showing purified mutant RdgC proteins with *E. coli* native RdgC

Lane 1, Molecular weight markers; Lane 2, wild type RdgC; Lane 3-5, RdgC proteins with mutations as labelled. FTM=Finger Tip Mutation.

6.2.3 Single-stranded DNA binding assays

As for the horseshoe interface mutants, the assays were done on RGL13, a 49 nt linear ssDNA (4.2.3).

The finger domains are crucial for ssDNA binding, as the deletion of the whole region ($RdgC_{\Delta F}$) almost completely abolished the protein's ssDNA binding activity (Figure 6-3). Only at the highest concentration (RdgC dimers:DNA ≈10,000:1) did $RdgC_{\Delta F}$ barely blur the DNA band. The R118C disulphide bond was able to bring back some of the binding activity, the level of which was however quite limited (Figure 6-3, $RdgC_{\Delta F \cdot R118C}$). Altering the charge of the finger tips ($RdgC_{FTM}$) had a major effect on ssDNA binding too; but the mutant was able to shift the DNA as a band (Figure 6-3 A, FTM), although there was still smear tailing it. To be specific, wild type RdgC showed signs of shift at 1 nM (RdgC dimers:DNA≈4:1); $RdgC_{FTM}$ started at 20 nM (RdgC dimers:DNA≈100:1); $RdgC_{\Delta F}$ and $RdgC_{\Delta F \cdot R118C}$ only shifted the DNA at the highest concentration (Figure 6-3 B).

6.2.4 Double-stranded DNA binding assays

As for the horseshoe mutants, I used RGL13/17 as the substrate for the assays (4.2.4).

RdgC with its whole finger domains deleted ($RdgC_{\Delta F}$) did not show any obvious sign of dsDNA binding activity. Interestingly, albeit with very weak and unstable ssDNA binding activity, $RdgC_{\Delta F \cdot R118C}$ and $RdgC_{FTM}$ retained moderate level of dsDNA activity and both made sheer shifts at high concentrations (Figure 6-4 A). The isotherms of the mutants suggested an energy hurdle when binding to dsDNA; but once the hurdle was overcome, the mutants could make stable contact with the dsDNA.

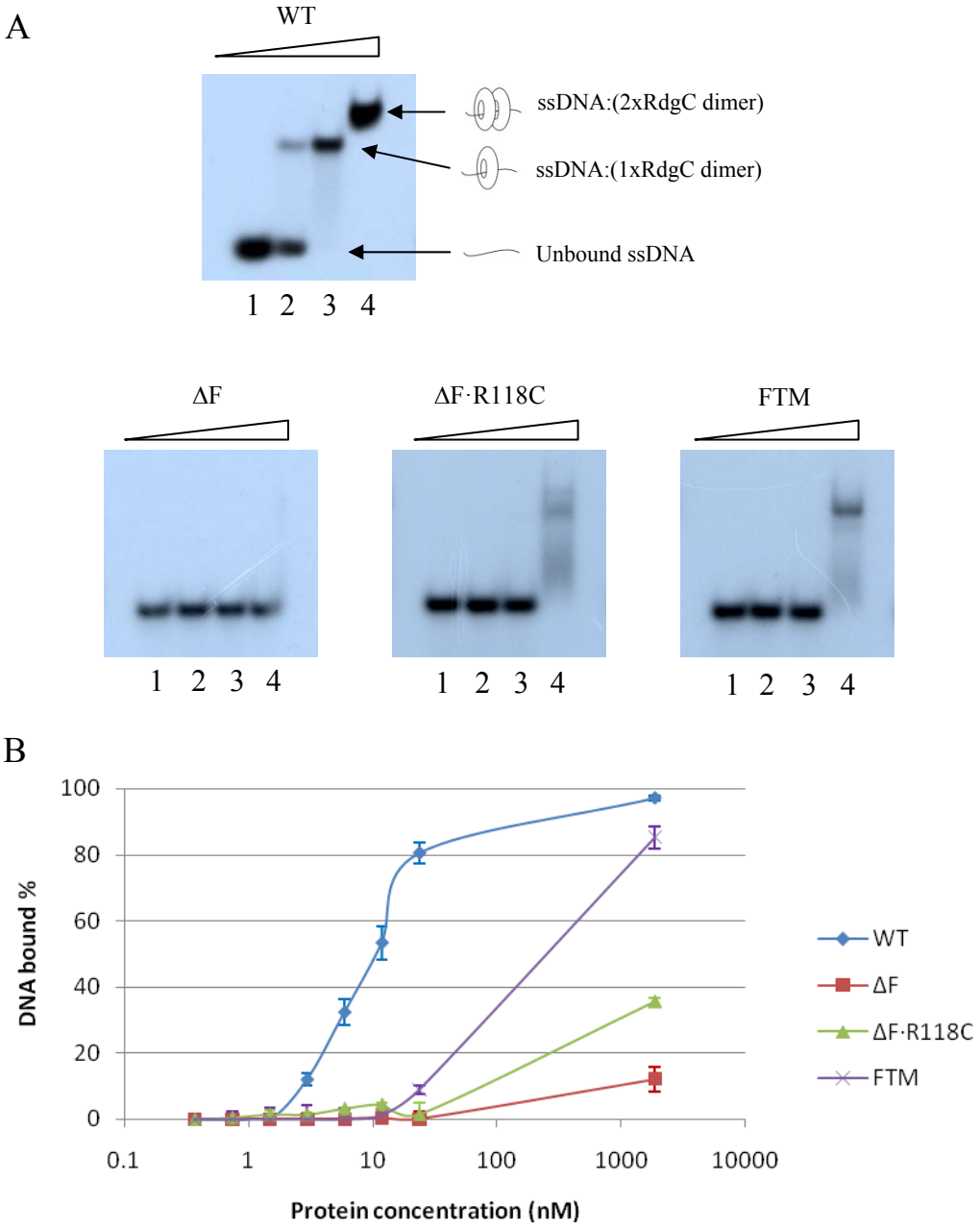


Figure 6-3 RdgC binds ssDNA

A. EMSA assays showing ssDNA binding of wild type RdgC and the finger mutants. DNA substrate was RGL13 (49 nt), at 0.2 nM. Protein concentrations: Lane 1, 0 nM; Lane 2, 5.92 nM; Lane 3, 23.68 nM; Lane 4, 1850 nM.

B. Binding isotherm for ssDNA binding (RGL13) of the RdgC proteins, assembled from at least 2 EMSAs. For the reactions, 0.2 nM RGL13 was used and the protein titrations were at 0, 0.37, 0.74, 1.48, 2.96, 5.92, 11.84, 23.68 and 1850 nM. The error bars stand for standard deviations of each value point.

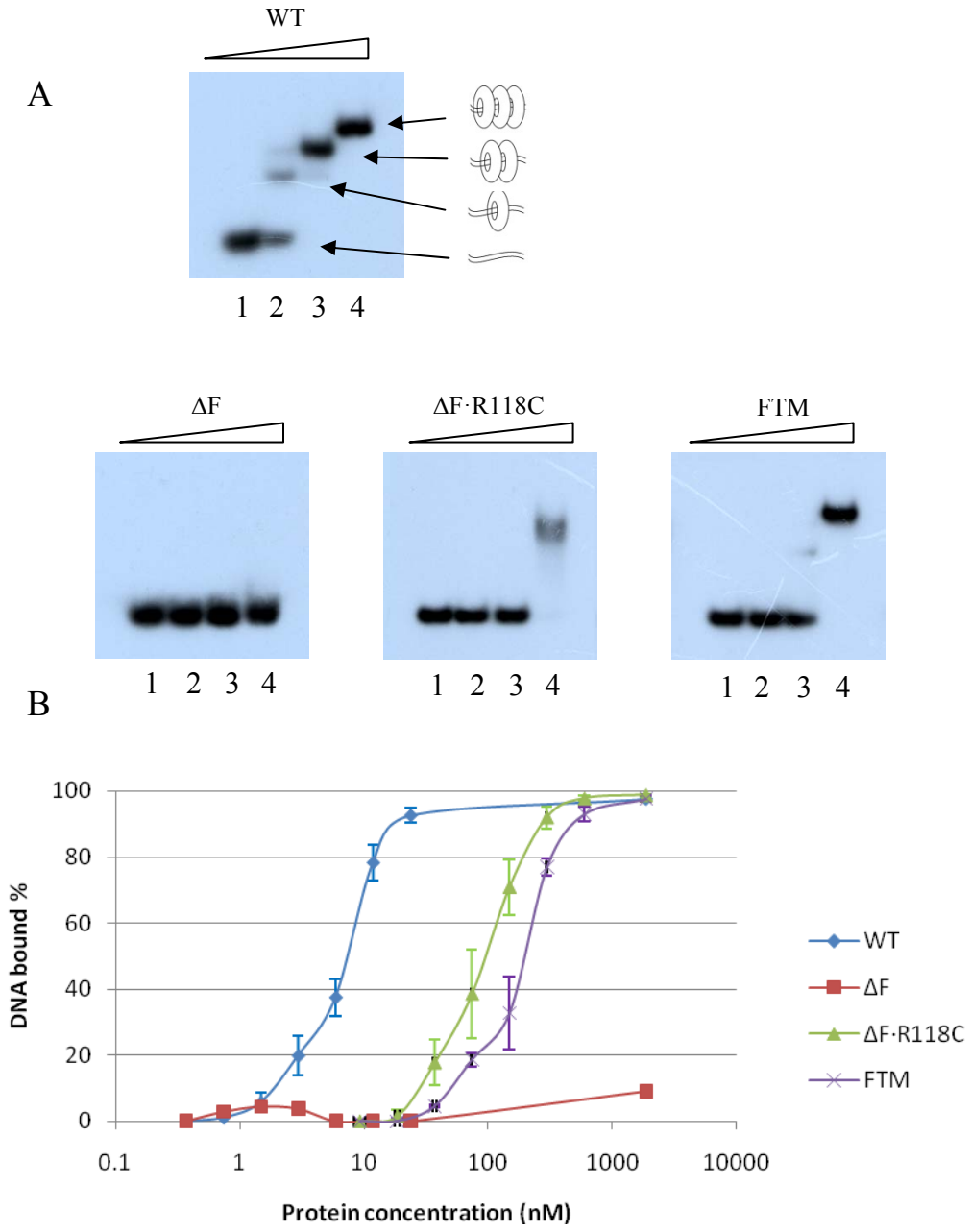


Figure 6-4 RdgC binds dsDNA

A. EMSA assays showing dsDNA binding of wild-type RdgC and RdgC mutants with mutation at the fingers. DNA substrate was RGL13/17 (49 bp), at 0.2 nM. Protein concentrations: Lane 1, 0 nM; Lane 2, 5.92 nM; Lane 3, 23.68 nM; Lane 4, 1850 nM.

B. Binding isotherm for dsDNA binding (RGL13/17) of the RdgC proteins, assembled from at least two EMSAs (except for ΔF). For the reactions, 0.2 nM RGL13/17 was used. The protein titrations of WT and ΔF were at 0, 0.37, 0.74, 1.48, 2.96, 5.92, 11.84, 23.68 and 1850 nM. The titrations of ΔF·R118C and FTM were at 0, 9.25, 18.5, 37, 74, 148, 296, 592 and 1850 nM. The error bars stand for the standard deviation of each value point (except for ΔF).

To be specific, wild type RdgC showed signs of shift at 0.5 nM (~2 RdgC dimers per DNA molecule); $RdgC_{FTM}$ and $RdgC_{\Delta F \cdot R118C}$ started at 40 nM (~200 RdgC dimers per DNA molecule); $RdgC_{\Delta F}$ only shifted the DNA at the highest concentration (Figure 6-4 B).

6.2.5 Circular DNA binding assays

As in 4.2.5, the assays used the 3000 bp pGEM-7zf(+) as the DNA substrate at a fixed concentration value of 4.0 ng/µl. The plasmid DNA was mixed with RdgC protein at 0, 18.5, 74.0 and 1590 ng/µl (Figure 6-5).

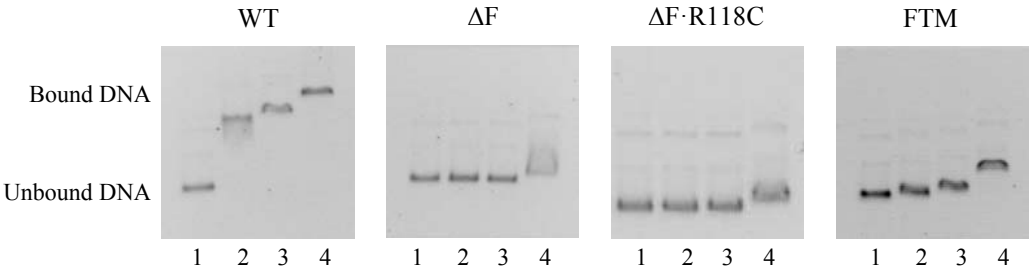


Figure 6-5 RdgC binds circular DNA

The DNA substrate was pGEM-7zf(+) at 4.0 ng/µl. Protein concentrations at Lane 1-Lane 4 were 0, 18.5, 74 and 1590 ng/µl, respectively.

In these assays, the changes introduced in the finger domains of the protein showed much reduced circular DNA binding activity. The R118C did not deliver any obvious improvement to $RdgC_{\Delta F}$'s circular DNA binding.

RdgC with altered fingertips charge ($RdgC_{FTM}$) also failed to show much active circular DNA binding. Note that the migration speed decreased gradually in a

very slow but seemingly constant rate as the protein concentration increased. This suggests that the amount of the mutant RdgC protein attached to DNA was much lower than the one of the wild type protein. However, the protein:DNA complex was still stable since the shifted bands were sharp, as seen in the wild type protein.

6.3 *in vivo* synthetic lethality assays

As in 4.3, all the finger domain mutants were tested for their *in vivo* activity in strains using synthetic lethality constructs of the three genotypes.

6.3.1 RdgC mutants in *priA*$^-$ strains

Using the same synthetic lethality assay described previously, all of the tested finger mutants ($RdgC_{\Delta F}$, $RdgC_{\Delta F \cdot R118C}$ and $RdgC_{FTM}$) failed to produce any white colonies on LB or minimal indicator plates (Figure 6-6), suggesting that none of them could substitute for wild type RdgC.

6.3.2 RdgC mutants in *priA*$^-$ *dnaC810* strains

On LB indicator plates, RdgC mutants with the finger domains deleted ($RdgC_{\Delta F}$ and $RdgC_{\Delta F \cdot R118C}$) did not appear to substitute for wild type RdgC, since plasmid-free segregants were inviable (no white colonies on LB indicator plates) (Figure 6-7 A, AM1854 and AM1855). The strain encoding $RdgC_{FTM}$ managed to produce a few tiny white colonies, suggesting some limited activity retained on the mutant (Figure 6-7 A, AM1949).

Similar to 4.3.3, the assay on minimal indicator plates is non-informative (Figure 6-7 B).

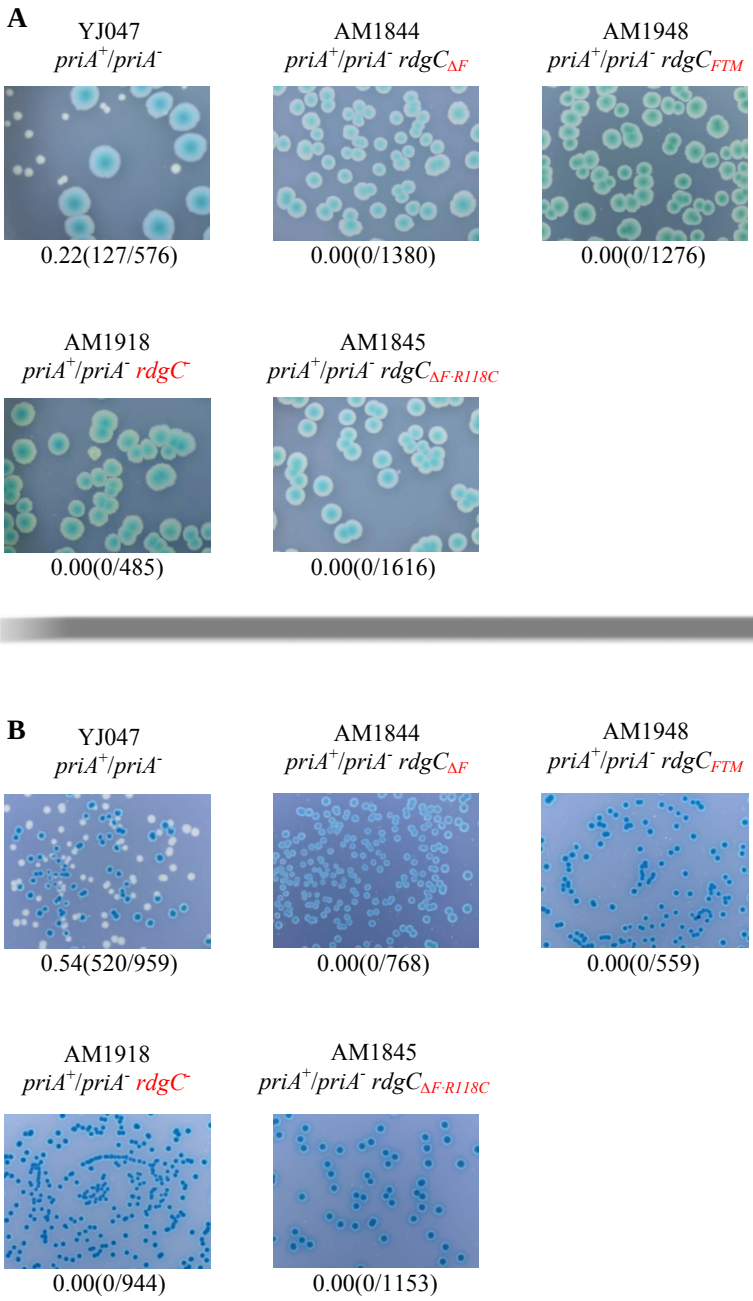


Figure 6-6 Synthetic lethality assays in *priA*$^+$/*priA*$^-$ for the finger mutants.

A. Synthetic lethality assays on LB media containing X-gal and IPTG
B. Synthetic lethality assays on minimal media containing X-gal and IPTG

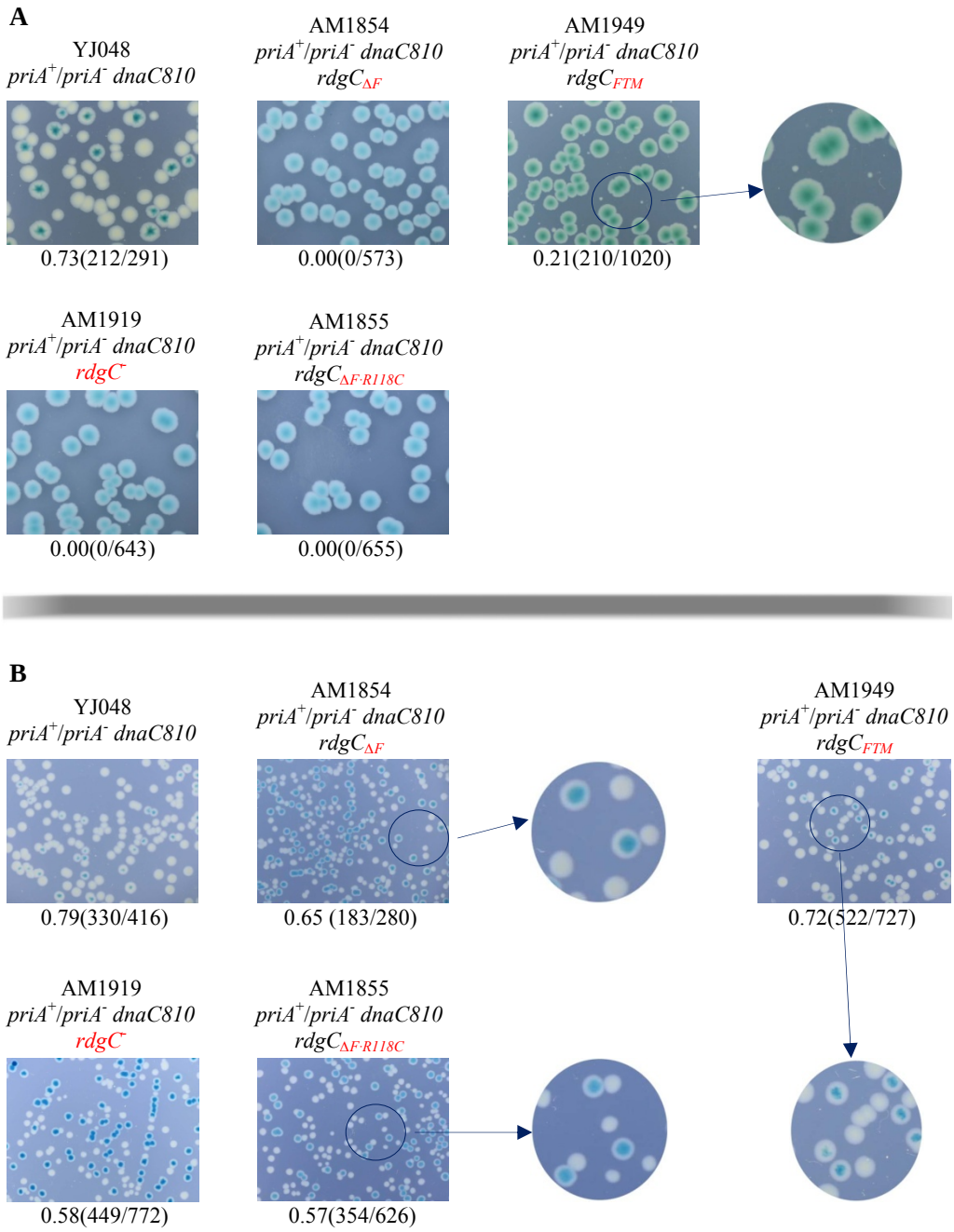


Figure 6-7 Synthetic lethality assays in *priA*$^+$/*priA*$^-$ *dnaC810* for the finger mutants.

A. Synthetic lethality assays on LB media containing X-gal and IPTG
B. Synthetic lethality assays on minimal media containing X-gal and IPTG

The circular pictures are magnifications of sections as indicated by gray arrows.

6.3.3 RdgC mutants in *priA*$^{-}$ *dnaC809, 820 priC*$^{-}$ strains

Similar to the results of the assay in the *priA*$^{+}$/*priA*$^{-}$ background, all the finger mutants ($RdgC_{\Delta F}$, $RdgC_{\Delta F \cdot R118C}$ and $RdgC_{FTM}$) failed to substitute for wild type RdgC in *priA*$^{-}$ *dnaC809,820 priC*$^{-}$ background on LB indicator plates (compare Figure 6-6 A with Figure 6-8 A).

Again, as with the constructs in the *priA*$^{+}$/*priA*$^{-}$ background (Figure 6-6 B) $RdgC_{\Delta F}$ and $RdgC_{\Delta F \cdot R118C}$ still could not deliver any activity in the strain background on minimal agar (Figure 6-8 A, YJ034 and YJ035). However, the strain encoding $RdgC_{FTM}$ behaved differently, as it produced a few small white colonies (Figure 6-8 B, YJ038, compared to Figure 6-6 B, AM1948). Although the small size and number of the white colonies suggest poor viability of the plasmid-free segregants, it proved once again $RdgC_{FTM}$ did retain some level of functionality.

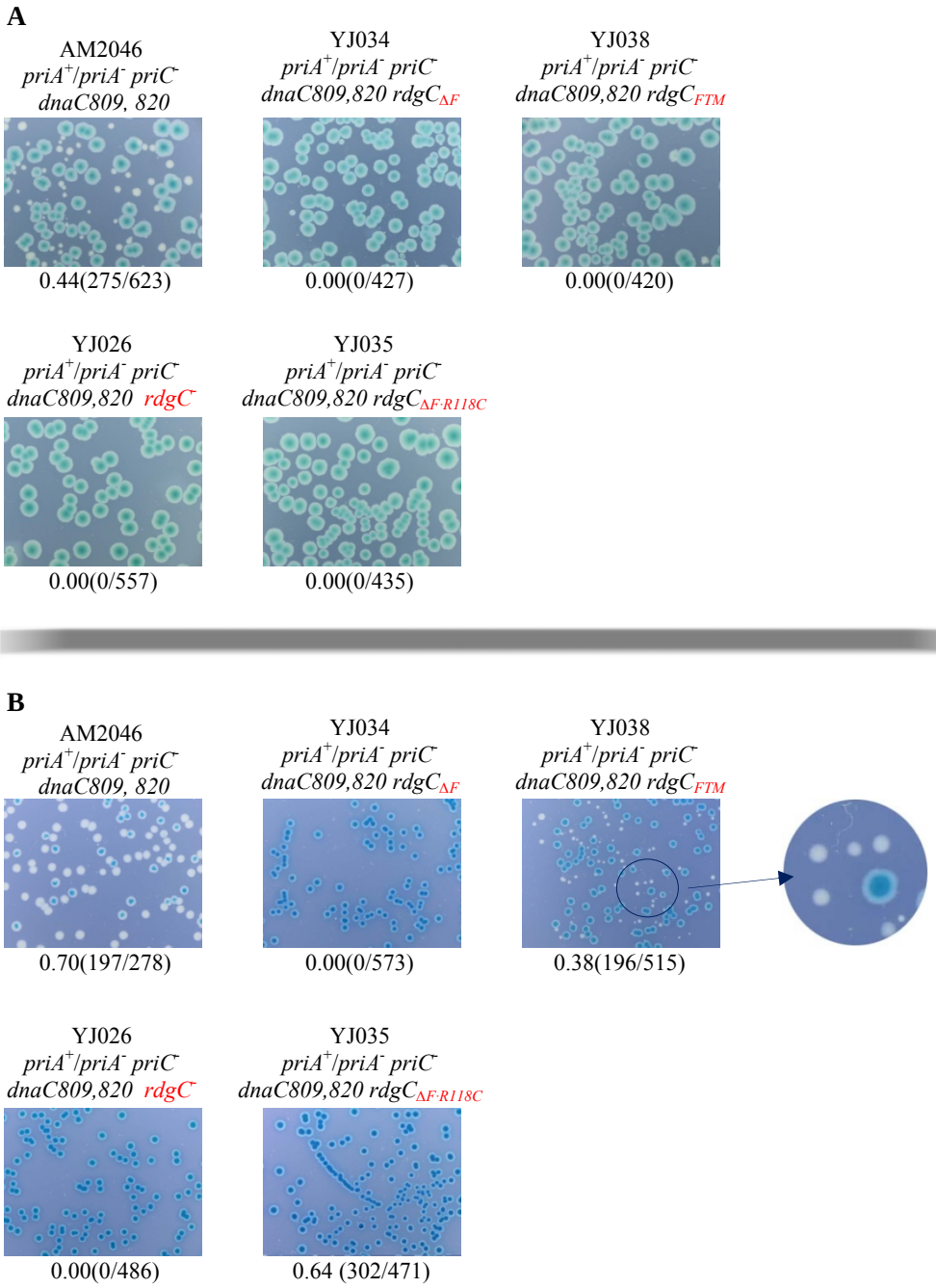


Figure 6-8 Synthetic lethality assays in *priA*$^{+}$/*priA*$^{-}$ *dnaC809,820* for the gate interface mutants.

A. Synthetic lethality assays on LB media containing X-gal and IPTG
B. Synthetic lethality assays on minimal media containing X-gal and IPTG

6.4 Discussion

6.4.1 The finger domains are crucial for DNA binding

RdgC with its whole finger domains deleted ($RdgC_{\Delta F}$) lost most of its DNA binding activities. Especially in short linear oligo binding assays, there was hardly a sign of DNA shifting (6.2.3 and 6.2.4). Only at very high concentration did the mutant manage to make yet very unstable contact with the circular DNA, which caused a smear featured shift (6.2.5).

The additional disulphide bond provided by the R118C ($RdgC_{\Delta F \cdot R118C}$) improved short linear oligo binding of $RdgC_{\Delta F}$, especially on dsDNA, suggesting that its inability to bind dsDNA is partially due to the weak link between the $RdgC_{\Delta F}$ gates (6.2.3 and 6.2.4). The improvement on ssDNA was limited. Interestingly, there was no sign of improvement regarding to circular DNA binding (6.2.5). This suggests a crucial role of the finger domains in ssDNA and circular DNA binding. It is also worth noting that the disulphide bond might not exist, as mentioned in the previous chapter (Section 5.4); the improved gate strength could attribute to a more hydrophobic environment brought by the R118C substitution.

The neutralised finger tips ($RdgC_{FTM}$) had a moderate negative effect upon short linear oligo binding (6.2.3 and 6.2.4). The mutant was still able to make stable contact with the duplex, suggesting that the central hole alone is enough for binding tightly to dsDNA. However, there appeared to be an obstacle in the initial DNA binding; but once the DNA is bound, they form a stable complex. This is also probably the case when $RdgC_{FTM}$ binds circular DNA (6.2.5). The

only difference was that the obstacle for the mutant to bind circular DNA appeared to be more difficult to overcome than when it bound ssDNA or dsDNA.

6.4.2 RdgC's *in vivo* functionality was greatly impaired by the finger mutations.

RdgC with finger domain deleted ($RdgC_{\Delta F}$) was found to retain little activity in all strain backgrounds tested. Its *in vivo* function was not recovered even with improved DNA binding affinity, attributing to the introduced R118C disulphide bound ($RdgC_{\Delta F \cdot R118C}$).

A less dramatic change by neutralising the electropositive charge of the fingertips ($RdgC_{FTM}$) also showed a major effect upon its *in vivo* functionality. However, it did retain some activity, which was only detectable in a few conditions tested. $RdgC_{FTM}$ was probably the weakest mutant protein so far that could substitute for wild type RdgC *in vivo*, even weaker than $RdgC_{H222A}$ (see Chapter 4 for details).

Chapter 7
General discussion

The work described in this thesis attempted to unfold the role of RdgC with structural, biochemical and genetic approaches. Previous studies had pointed towards RdgC's involvement with DNA recombinational repair, specifically to prevent illegitimate loading of RecA onto dsDNA, thus limiting unnecessary exchanges. This idea is supported by three pieces of evidence: 1) RdgC binds DNA, with a greater affinity for dsDNA (Moore et al., 2003); 2) in cells with defect in DNA replication restart, RdgC was found to prevent RecA's over-activity (Moore et al., 2003); 3) RdgC has been reported to limit RecA loading by substrate competition *in vitro* (Drees et al., 2006).

In this work, the RdgC protein of *E. coli* was overexpressed and purified to high quality. It was then crystallised either with or without DNA. It was clear that the addition of DNA molecules enhanced the crystal growth. Crystallisation in the absence of DNA produced smaller and non-isomorphous crystals, and diffracted to only 2.8 Å, which could be enhanced up to 2.4 Å when co-crystallised with DNA. The structure of RdgC dimer was identified to mimic a ring structure, with two additional fingers attached to the gate interface. Biochemical and genetic studies proved that intact dimer interfaces and finger domains are necessary for both of its DNA binding activity and its functionality *in vivo*.

7.1 Does RdgC bind DNA by encircling it?

Although the structure of the RdgC dimer was resolved from the RdgC:DNA co-crystallisation, the DNA molecule could not be located. This is not surprising since RdgC is known for its non-specific preference on structure or sequence, with which RdgC is unable to hold DNA in unique position or conformation. However, given the size and the positive charge of its inner hole (30 Å diameter), it is very likely where RdgC interacts with DNA when they form a stable complex. This may also explain why crystallisation without DNA did not yield high quality of crystals. RdgC exists in a monomer-dimer equilibrium, with a lower percentage of higher-order oligomers also present (Tessmer et al., 2005; Drees et al., 2006). The presence of DNA might help mediate dimer formation and drive the equilibrium towards dimerisation, a possible reason why no monomers were observed in crystals. In fact, the RdgC protein appeared to exist only as dimers on the gel-filtration column during purification, eluting as a single peak at size corresponding to a dimer. These suggest that RdgC exists primarily as a dimer, and monomers rarely occur and are transient in the presence of DNA. This also rejects the possibilities proposed by Drees, et al., 2006, who concluded RdgC bound DNA with higher affinity as a monomer, compared with a dimer, and the protein existed almost exclusively as monomers at low concentrations. If they were true, then $RdgC_{FTM}$ would not have had trouble to bind linear or circular DNA at low protein concentrations, which was disproved by this work. This also makes clear that the lower shifts seen in ssDNA and dsDNA binding assays actually reflected one dimer bound per one DNA molecule, rather than one monomer

bound per one DNA molecule (Figure 4-3), i.e. RdgC binds DNA primarily as a dimer.

Although RdgC has a unique protein fold (Dali server search (Holm and Sander, 1993)), there are DNA binding proteins that also adopt a ring structure with a positively charged central DNA binding channel. Sliding clamp proteins such as PCNA (Krishna et al., 1994) and β clamp (Kong et al., 1992) are such examples. Sliding clamps are processivity factors that encircle DNA at replication forks. The β clamp of *E. coli* is a dimer, with each monomer containing three domains of identical topology (Figure 7-1 A). The tightly closed ring has a 35 Å hole, with the ability to maintain a topological interaction with DNA, whilst not directly interacting with it (Kelman and O'Donnell, 1995). This allows the β clamp to slide along the replicating DNA without limiting the speed of the replisome. In comparison, another toroidal protein, the SARS-CoV nsp7-nsp8 hexadecamer (Figure 7-1 B) has similar dimensions of the hole (30 Å) to that of RdgC and is supposed to interact directly with the backbone of RNA (Zhai et al., 2005). SARS-CoV (Severe Acute Respiratory Syndrome Coronavirus) is a positive-stranded RNA virus with complex RNA-processing machinery (Gorbalenya, 2001; Snijder et al., 2003; Thiel et al., 2003), and the SARS protein is thought to act as a processivity factor and encircles dsRNA. It is therefore reasonable to propose that RdgC encircles and interacts directly with the backbone of dsDNA (25 Å diameter) with its positively charged sidechains lining the hole.

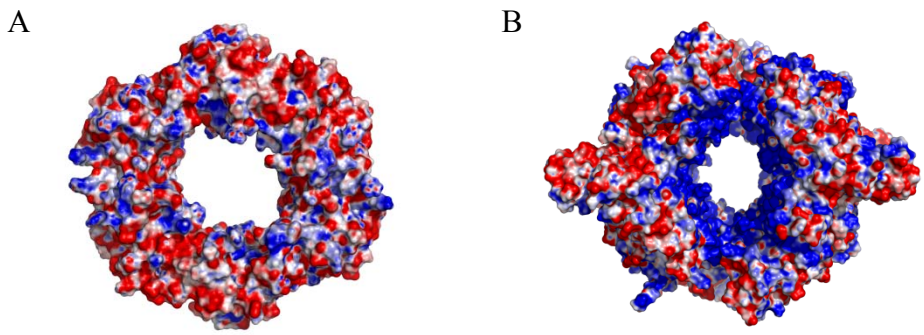


Figure 7-1 Electrostatic surface representation of a sliding clamp protein and the SARS-CoV nsp7-nsp8 hexadecamer (pictures made by Geoff Briggs)
A. β clamp
B. SARS-CoV nsp7-nsp8

The proposed DNA binding channel of RdgC is 40 Å from the front to the back of the ring and 55 Å from fingertip to fingertip. This area could interact with approximately 18 bp of DNA, agreeing well with the 15-20 bp required to make a stable complex *in vitro* (Moore et al., 2003; Moore et al., 2004). However, the fact that the RdgC protein with neutralised or deleted finger tips ($RdgC_{FTM}$ and $RdgC_{\Delta F \cdot R118C}$) could still make stable contact with dsDNA implies that the central hole alone is enough to confer sufficient binding to dsDNA (6.2.4). In addition, this model can also answer why RdgC has no specific preference on any structure or sequence, and its slightly better affinity for dsDNA than ssDNA. It is very likely that the protein is primarily interacting with the negatively charged backbone if DNA is bound in the central hole. Therefore, as long as the DNA molecule has a region of the correct dimensions (a duplex of at least 20 bp), the binding of different structures or sequences will be indistinguishable. On the other hand, the slightly weaker ssDNA binding might reflect a poor positioning due to a less stable contact between the protein's positively charged binding channel and the

ssDNA molecule's backbone phosphates. It is also worth noting that $RdgC_{\Delta F \cdot R118C}$ and $RdgC_{FTM}$ showed a less stable interaction with ssDNA (6.2.3). This might imply that the positively charged fingertips are important to ssDNA binding, but not to dsDNA binding.

7.2 How does RdgC load onto DNA?

After obtaining the ring structure of RdgC and its proposed DNA binding channel, another purpose of this project naturally emerged as to solve how RdgC loads onto DNA.

As proposed in Chapter 4, the RdgC ring may either slide onto DNA ends without the need to open either dimer interfaces, or open one or both of them to 'embrace' the DNA. However, if RdgC binds DNA primarily as a dimer, it then must open at least one of the dimer interfaces to bind to circular DNA.

The horseshoe interface is held together by hydrogen bonds between the β strands and also between the 2 anti-parallel helices on the outer surface. The gate interface is held primarily by sitting F120 from one chain into a hydrophobic pocket formed by residues from the other chain. A hydrogen bond between R118 from each chain completes the interface. The gate interface is also winged by two finger domains. This extreme asymmetry between the two interfaces outlines potential different functions for the protein, implying that only one interface is unlocked for DNA binding.

In vitro DNA binding assays revealed that the mutations on the conserved residues at the horseshoe interface did not seem to affect the protein's circular

DNA binding (Chapter 4). This suggests that the protein with impaired horseshoe interface was still able to open and close the ring for DNA binding. However, some mutants ($RdgC_{E218R}$ and $RdgC_{K227A}$) had obvious reduced affinity for linear oligos. It appears that disrupted horseshoe interface leads to a change of the central hole dimension, causing a weakened protein-DNA backbone contact and therefore allowing the RdgC ring to slide along DNA molecule more easily. This did not affect circular DNA binding presumably because they fall off the end of small linear oligos, whereas they move freely along circular DNA but stay attached, like beads on a necklace.

The scenario on the other interface was similar, except that interfering with the hydrophobic interactions was even more destructive to the protein's activity on short linear oligos and circular DNA binding. This is not surprising as the main force to stabilise the gate interface is the hydrophobic interactions, whereas the residues substituted at the horseshoe interface are less influential. The R118 hydrogen bond was proved to be a minor force in maintaining the gate interface.

The question to this point is how RdgC opens which interface(s) for DNA binding. Unlike the β clamp protein that relies on a clamp loader protein for DNA loading, RdgC has a mechanism of opening the ring by itself. The finger domains are a possible candidate for such a role because of their high flexibility (reflected by high B factors) and their electropositive potential (recognising DNA). This also raises the possibility that RdgC opens the gate interface for access to DNA via the assistance of the finger domains in vicinity.

Indeed the finger mutants exhibited very different behaviours compared to their interfaces counterparts, which would lead the fingers to such a role as the RdgC ring loader. The protein with the whole finger domain deleted ($RdgC_{\Delta F}$) was unable to form stable complex with either short linear oligos or circular plasmid molecules. This could be partially because the disrupted hydrophobic interactions by the deletion cannot hold together the gate interface, thus losing the conformational constraint on the DNA binding. This again proved the remaining R118 hydrogen bond alone is not able to hold the gate closed. The results from a rather milder mutant with neutralised fingertips ($RdgC_{FTM}$, see Chapter 6) were more interesting. With intact dimer interfaces the mutant could still make stable contact with dsDNA similar to the wild type protein. The only problem appeared to be a restricted access to dsDNA molecules, as represented by a delayed-in-concentration binding isotherm seen in Figure 6-4 B. This delay can be explained by the results from circular plasmid DNA binding assays, in which the finger tip mutation almost completely abolished the protein's binding activity on circular DNA. $RdgC_{\Delta F \cdot R118C}$ is the only other mutant that bound well to dsDNA but poorly to circular DNA. A plausible explanation is that both have problems of recognising DNA and opening the ring for DNA binding, especially for circular DNA molecules; while short linear oligos can find their way to the binding channel by sliding in from either DNA end. This would certainly be more difficult than the usual way of loading, and results in a shift of the binding curve to higher protein concentrations (Chapter 6). It is therefore clear that the electropositive finger domains are probably responsible for recognising DNA and opening the gate interface for binding DNA, especially circular DNA. Given the fact that disrupted gate

interface (F120T and F120S) caused the RdgC:linear DNA complex constantly to dissociate, the gate interface is possibly required to close after the protein ring accommodating DNA in the central hole for a usual binding.

Therefore, I propose that RdgC loading onto DNA is normally finger domains mediated, which consists of four stages: 1) the finger domains recognise and make first contact with target DNA molecules (Figure 7-2 A); 2) then the finger domains rotate on binding, which disrupts the F120-mediated interaction at the gate interface (Figure 7-2 B); 3) disruption of the interface causes the gate to open, and the RdgC can slide over the DNA (Figure 7-2 C); 4) once the DNA is properly located in the central binding channel, the finger domain rotates back into its original position, the gate closes, and the ring structure is reformed (Figure 7-2 D). In this model, RdgC forms a dimer before it encounters a DNA molecule, and the horseshoe interface remains linked during binding.

Another possible mechanism for DNA loading may be achieved by DNA mediated RdgC dimerisation. Several lines of evidence suggest that RdgC is present as monomer, dimer and higher-order oligomers in solution (Tessmer et al., 2005; Moore et al., 2003; Drees et al., 2006). It is therefore possible for RdgC to attach to a DNA molecule as a monomer on one side, then another monomer lands on the DNA properly to complete the binding and protein dimerisation. This binding mode can be used to explain the binding pattern of $RdgC_{FTM}$ and $RdgC_{\Delta F \cdot R118C}$ on the circular plasmid DNA (Figure 6-5 FTM and ΔF·R118C). The gradual minor shifts seen on the gel might be a result from the rare occurring DNA mediated RdgC dimerisation. The relatively later shift in

$RdgC_{\Delta F \cdot R118C}$ could be due to its disulphide bond (or improved hydrophobic environment, see section 5.4) driving the monomer-dimer equilibrium towards the dimer side, thus less monomers are available for DNA mediated dimerisation. The weak shifts again proved the very low proportion of the monomer species in solution, as proposed in section 7.1. In contrast, $RdgC_{\Delta F}$ could only make a smear shift at the highest protein concentration, a possible result from its weak gate interface constantly opening to release bound DNA (Figure 6-5 ΔF).

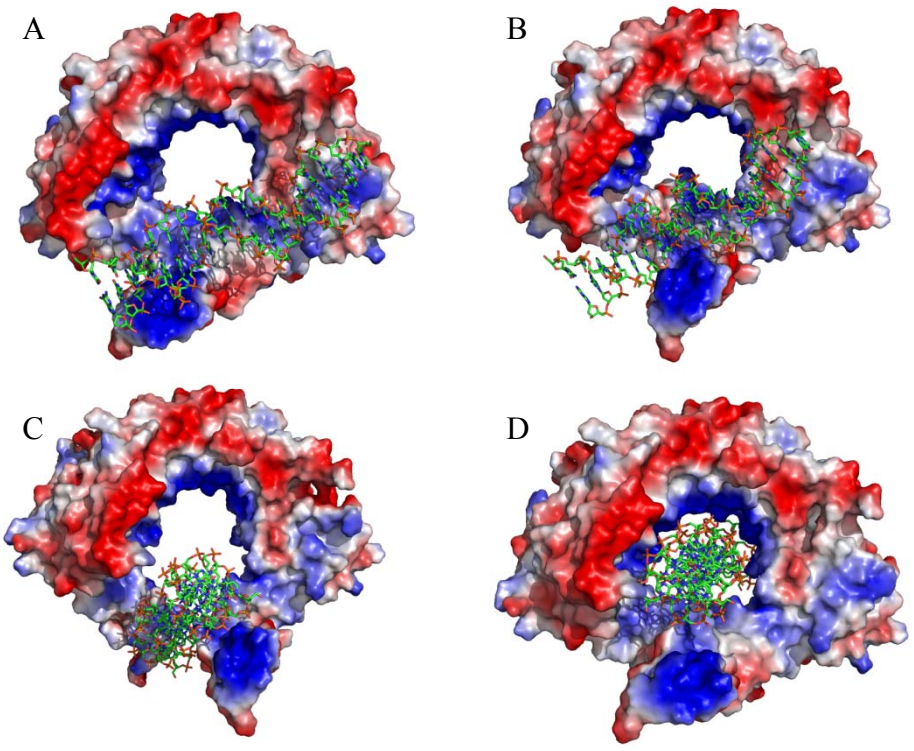


Figure 7-2 finger-domain-mediated DNA loading model (pictures made by Geoff Briggs)

A. The finger domains recognise and make first contact with the target DNA molecule
B. The finger domains rotate on binding, which disrupts the F120-mediated interaction at the gate interface
C. Disruption of the interface causes the gate to open, and the RdgC can slide over the DNA
D. Once the DNA is properly located in the central binding channel, the finger domain rotates back into its original position, the gate closes, and the ring structure is reformed

In summary, RdgC binds DNA presumably through finger domains-mediated DNA loading model. Nonetheless, RdgC could still slide onto its substrates without opening the ring when it encounters open DNA ends; or form dimers on the target DNA (DNA mediated dimerisation). However, they might rarely occur for the wild type protein.

7.3 What does RdgC do *in vivo*?

As is already known, RdgC is hardly found to do anything else other than bind to DNA. This special property leads to the idea that RdgC functions *in vivo* mainly as a DNA binding protein. Therefore, it had been expected that weak DNA binding RdgC mutants would function poorly *in vivo*, and *vice versa*. This is generally true as weak DNA binding mutants did not substitute for wild type RdgC in all conditions tested, such as $RdgC_{E218R}$, $RdgC_{K227A}$, $RdgC_{\Delta F}$, $RdgC_{F120T}$ and $RdgC_{F120S}$. However, further analysis of my data reveals that different RdgC mutants exhibited preference for different DNA substrates. In general, the interfaces mutants favoured circular plasmid DNA, whereas some finger mutants ($RdgC_{\Delta F \cdot R118C}$ and $RdgC_{FTM}$) had a preference towards short linear oligos. The question is therefore affinity for which DNA species is linked to RdgC's activity *in vivo*, and why. Furthermore, strong evidence supports that the RdgC ring opens when loading onto DNA, and closes afterwards (refer to the finger domains-mediated DNA loading model, see the previous section for details); but it did not exclude the sliding-on and DNA-mediated dimerisation mode. A question could be asked as to which occurs for RdgC's *in vivo* functionality. Last but not the least, if RdgC acts *in vivo* purely

as a DNA binding protein, then mutants that retain good DNA binding affinity (such as $RdgC_{H222A}$, $RdgC_{Q212A}$, $RdgC_{R118A}$ and $RdgC_{R118C}$) would be expected to also retain full functionality *in vivo*. However, none of the mutants could fully substitute for wild type RdgC in the three genetic backgrounds tested.

In general, RdgC mutants with poor binding activity on short linear oligos retained very weak or no functionality *in vivo*, disregarding their affinities for circular DNA, such as $RdgC_{E218R}$ and $RdgC_{K227A}$. As discussed in the previous section, the "affinity" of a mutant for 'circular' DNA does not necessarily reflect how well it binds DNA. This could suggest linking RdgC's *in vivo* functionality directly to the tightness it binds DNA. For instance, $RdgC_{E218R}$ and $RdgC_{K227A}$ were poor short linear DNA binding proteins. Despite that their interfaces are closed tightly (good affinity for circular DNA), their weak interaction with DNA failed to deliver any valid wild type functionality *in vivo* (see Chapter 4 for details).

As the data of the *in vitro* assays show, RdgC binds DNA probably by opening the ring, sliding onto its substrate and/or DNA-mediated dimerisation, but with a clear preference for the first. This exactly fits the data from *in vivo* assays, in which mutants with the ability to open the gate interface (e.g. $RdgC_{H222A}$, $RdgC_{K211A}$, $RdgC_{R118A}$ and $RdgC_{R118C}$, see Chapter 4 and 5) generally retained more functionality than the ones without (e.g. $RdgC_{FTM}$ and $RdgC_{\Delta F \cdot R118C}$, see Chapter 6). $RdgC_{R118A}$, $RdgC_{R118C}$ and especially $RdgC_{K211A}$ could substitute for wild type RdgC under many conditions. On the other hand, $RdgC_{FTM}$ was only poorly functional in *priA⁻ dnaC810* on LB and *priA⁻ pirC⁻ dnaC809,820* on MA; $RdgC_{\Delta F \cdot R118C}$ was inactive in all conditions tested.

Another puzzle was raised by the mutants which showed wild-type like DNA binding activities *in vitro*, but had very poor or even no functionality *in vivo*, such as $RdgC_{H222A}$ and $RdgC_{Q212A}$. Q212 was expected to form a triad with E218 and K227 (section 3.4), and the two histidines at 222 to interact with each other. The *in vitro* assays proved the roles of E218 and K227 in maintaining the horseshoe interface, as the substitutions greatly weakened the protein's linear DNA binding. The *in vitro* data suggest that Q212 does not interact with E218 and K227, nor do H222 with each other; or they do interact, but the interactions are less crucial in maintaining the configuration of the horseshoe interface. The crystal structure clearly supports the latter (Figure 4-1). However, the *in vivo* data revealed that Q212 and H222 are actually important for the protein's functionality. This could suggest an involvement of other important biological processes of Q212 and H222, or altered DNA binding activities undetected with the mutations. On the other hand, $RdgC_{K211A}$ has intact dimer interfaces, but its affinity for DNA is slightly reduced. In contrast to $RdgC_{H222A}$ and $RdgC_{Q212A}$, $RdgC_{K211A}$ could substitute for wild type RdgC in most genetic background tested. The integrity of the horseshoe interface is therefore vital for RdgC *in vivo* functionality.

7.4 How does RdgC regulate RecA's activity?

RecA has been found to be toxic in cells having defects in DNA replication restart system (Moore et al., 2003). Cells lacking PriA and RdgC and while carrying the *dnaC810* suppressor of *priA* are barely viable on LB. This defect can be alleviated by removing *recF, recO, recR* or *recA* (R. G. Lloyd, unpublished data). It is therefore likely the RecFOR mediated RecA loading

pathway that contributes to the growth defect. The fact that *priA⁻ dnaC810 rdgC⁻* cells grew on minimal medium indicate that toxic RecA loading is primarily a feature of cells growing rapidly in rich media.

RecFOR are known to load RecA on ssDNA gaps. They first identify platforms for loading RecA, and then dislodge bound SSB proteins. With assistance from RecFOR, RecA nucleates on the ssDNA and the RecA:ssDNA filament extends within the region of ssDNA from 5'-3' (Morimatsu and Kowalczykowski, 2003; Sakai and Cox, 2009). Thereafter, the RecA filament starts searching dsDNA homologous to the bound ssDNA. When suitable dsDNA molecule is found, the RecA filament aligns the bound ssDNA with it, and promotes a strand exchange between them. Considering RdgC as a DNA binding protein, it possibly regulates RecA activity either by binding to ssDNA gaps to limit RecA's loading, or by covering homologous dsDNA to prevent recombination.

7.4.1 Does RdgC bind dsDNA to limit RecA activities?

The central hole of the RdgC ring is well designed for dsDNA binding. The tight binding of RdgC to dsDNA would hinder the access to the dsDNA from RecA:ssDNA filaments. Its high affinity for dsDNA can even dislodge any RecA filaments already bound to dsDNA (Drees et al., 2006). It is generally true because most of the mutations that did not obviously affect the protein's dsDNA binding could substitute for wild type RdgC under certain conditions, such as $RdgC_{K211A}$, $RdgC_{R118C}$, $RdgC_{R118A}$ and $RdgC_{H222A}$; whilst mutants with poor affinity for DNA were mostly inactive *in vivo*, such as $RdgC_{E218R}$,

$RdgC_{K227A}$, $RdgC_{\Delta F}$ and $RdgC_{F120T}$. However, this model needs further evidence. Firstly, RdgC does not have any sequence or structural preference, and it only reaches 1000 copies per cell at its prime in the exponential phase, compared to 15,000 RecA molecules per cell (Moreau, 1987; Sassanfar and Roberts, 1990; Stohl et al., 2003). Both features will make the protein difficult to build barriers to RecA activities at hot spots. To overcome this, some yet unknown proteins may be required to recruit RdgC from the limited resources to sites where RecA activities need to be limited. Secondly, $RdgC_{Q212A}$ had dsDNA binding affinity comparable to the one of the wild type protein, but was inactive in all the genetic backgrounds tested. Although there could be undetected dsDNA binding defect of the mutant protein, the substitution might also interfere with possible protein-protein interactions, which may be also important to the wild type protein's function. Indeed, *Escherichia coli* DNA ligase, the β subunit of the histone-like protein Hu and endonuclease V have been reported to interact with RdgC (Butland et al., 2005). DNA ligase catalyses the sealing of 5' phosphate and 3' hydroxyl termini at nicks in duplex DNA (Gottesman et al., 1973; Konrad et al., 1973; Sriskanda and Shuman, 2001). Hu is one of the most abundant DNA-binding proteins in *Escherichia coli*, and it contributes to the compaction of the genome into tight nucleosome-like structures (Rouvière-Yaniv et al., 1979). Neither of the proteins appears to be featured as RdgC recruiter, as they do not seem to bind specifically to dsDNA molecules that are potentially involved in illegitimate recombination. Endonuclease V, the gene product of *nfi*, is a DNA repair enzyme that repairs deaminated base damage by hydrolysing the second phosphodiester bond 3' from a base lesion (Yao et al., 1994; Huang et al., 2001). Nonetheless, One of

its biochemical properties revealed by binding and kinetic analyses is the ability to remain bound to nicked deaminated lesions after strand cleavage (Yao and Kow, 1995; Huang et al., 2001), suggesting that it might recruit other proteins to assist the DNA repair. RdgC could be among the recruited proteins for limiting RecA access to the region, for recombining with a damaged duplex as a repair template is no good and sometimes can even be disastrous. However, more evidence needs to strengthen the idea, as the reported interactions between RdgC and other proteins could be indirect and mediated by DNA.

The idea that RdgC has specific binding partners in *E. coli* might explain the failure of *Neisseria* RdgC protein to substitute for the native protein in *E. coli*. I made the relevant synthetic lethality constructs to test this and found that the heterogeneous protein did not allow recovery of plasmid-free cells (Appendix 2). Given that *Neisseria* RdgC binds DNA much like the *E. coli* protein (Moore et al., 2004), the failure may reflect the fact that the RdgC binding partners are species specific.

7.4.2 Does RdgC bind ssDNA to limit RecA activities?

This possibility cannot be ruled out just because RdgC has relatively weaker affinity for ssDNA compared with dsDNA. RdgC has been reported to exhibit higher affinity for ssDNA with secondary structures, which resulted in a potent inhibition to RecA activities (Moore, 2002; Drees et al., 2006). The ability to bind ssDNA also does not conflict with the role of SSB. SSB is only inhibitory to RecA nucleation. Once nucleation completes (normally assisted by other regulatory proteins, such as RecFOR), the RecA:ssDNA filament readily

removes SSB at its 3' proximal end, and SSB facilitates the filament extension by removing secondary structures on ssDNA. RdgC offers complementary functions. Unlike SSB, RdgC does not bind ssDNA tightly, which leaves room for RecA loading. However, RdgC is able to stabilise secondary structures in ssDNA regions, and cannot be 'pushed' off from ssDNA by RecA filament extension if the ssDNA region is flanked by dsDNA or secondary structures. Both features could limit RecA filament extension, but not RecA nucleation. Interestingly, ssDNA flanked by duplexes is a favourable substrate for RecFOR.

This hypothesis is supported by the *in vitro* and *in vivo* mutant assays, where mutants with weak ssDNA binding also functioned very poorly *in vivo*. However, the data cannot discriminate between this model and the one proposed earlier in the previous section, since the weak ssDNA binding mutants generally also bound poorly to dsDNA. Inevitably, this model requires further evidence, too. The RdgC ring is designed well for dsDNA binding, but not for ssDNA. The main duty of the protein is probably achieved by dsDNA binding. The ssDNA binding is therefore at most a bonus to the protein if it is required in certain situations. Nonetheless, RdgC's slightly weak ssDNA binding necessitates its presence on ssDNA prior to RecA loading for any inhibitory effect to occur (Drees et al., 2006). Even under suitable conditions, such inhibitory effect requires large number of the RdgC proteins, exceeding its physiological sense. As the model proposed in 7.4.2, it therefore also needs a recruiter to guide the proteins to sites of interest.

7.4.3 Does RdgC need to bind both dsDNA and ssDNA to limit RecA activities?

This is a tempting idea as RdgC can bind to both dsDNA and ssDNA. One interesting point is that the finger domains are evidently dispensable for a stable contact with dsDNA, but are crucial for interaction with ssDNA, as suggested by the *in vitro* investigation of $RdgC_{FTM}$ and $RdgC_{\Delta F \cdot R118C}$. In effect, RdgC will have optimum binding at dsDNA-ssDNA junctions, with full contact between the central hole and the dsDNA backbone, and with interaction between the ssDNA in vicinity and one of the fingers. Such junctions are exactly where the boundaries of ssDNA gaps are defined, a preferable target for RecFOR to load RecA. This model would allow limited number of RdgC proteins to be distributed specifically, and require only a few proteins per site for required activities. In fact, a recent study reported that RecFR binds randomly on dsDNA (Sakai and Cox, 2009). An unknown factor has therefore been proposed to guide RecFR complex to gap junctions, which in turn facilitates RecA loading with assistance from RecO (Sakai and Cox, 2009). RdgC can also help restrict RecA filament extension within the region of ssDNA. This study indeed revealed RdgC with altered finger domains (responsible for ssDNA contact in this model) or weakened dsDNA binding had severely reduced activities *in vivo*. However, cells having defects in the RecFOR pathway are susceptible to DNA damage, a result of impaired recombinational DNA repair system, whereas deleting *rdgC* does not confer any obvious defects to the cells (Moore et al., 2003). It appears that the RecFR

complex is still capable of initiating RecA loading at appropriate sites in the absence of RdgC.

To sum up, RdgC regulates RecA activities mainly by binding to DNA. Although I have proposed three models on how RdgC might achieve the effect, they are compatible with each other and might co-exist *in vivo*.

7.5 RdgC in other species

The *rdgC* gene is strictly distributed within β and γ subdivisions of Proteobacteria, including many of the human pathogens such as *Neisseria meningitidis*, *Neisseria gonorrhoeae* and *Pseudomonas aeruginosa* (Moore, 2002).

7.5.1 RdgC from *Neisseria meningitidis* and *Neisseria gonorrhoeae*

RdgC from *N. meningitidis* shares 35% identity with the one from *E. coli*. Extensive studies performed by Moore and colleagues revealed its biochemical features very similar to *E. coli* RdgC, in that *N. meningitidis* RdgC also binds DNA with little specificity for sequence or structure (Moore et al., 2004). However, unlike its *E. coli* counterpart, *N. meningitidis* RdgC binds ssDNA almost as well as dsDNA (Moore et al., 2004). The similarity of the proteins *in vitro* activities suggests that RdgC performs a similar function in both *E. coli* and *N. meningitidis*. Indeed, this is probably true in *N. gonorrhoeae*, a very close relative of *N. meningitidis*. Under normal conditions, *N. gonorrhoeae*

undergoes frequent site-specific RecA-mediated recombinations to constantly change its pilin conformation, a process very important for the cells to evade the host immune response. This process is known as pilin antigenic variation, which has been found to require the RecFOR pathway to load RecA for recombination (Mehr and Seifert, 1998; Skaar et al., 2002). RdgC is also required because cells lacking it have 2-4 fold decreased pilin antigenic variation (Mehr et al., 2000). It is therefore obvious that RdgC is involved in the RecFOR-RecA mediated DNA recombination in *N. gonorrhoeae*. However, it is still not clear how RdgC contributes to the process. Given that another negative regulator of RecA—RecX, is also required in pilin antigenic variation, it is very likely that both RdgC and RecX act in concert to limit RecA activities to an appropriate level in *Neisseria gonorrhoeae*.

In this work, I tested the RdgC protein from *N. meningitidis* with the synthetic lethality assays. Quite unexpectedly, the *N. meningitidis* protein did not substitute for *E. coli* wild type RdgC in any conditions tested (Appendix 2). As previously mentioned in 7.4.1, this is probably due to the species specific RdgC binding partners. Nonetheless, it is also possible that the relatively better ssDNA binding disables *N. meningitidis* RdgC in *E. coli*.

7.5.2 RdgC from *Pseudomonas aeruginosa*

The structure of RdgC from *P. aeruginosa* has recently been revealed as very similar to the one from *E. coli* (Figure 7-3) (Ha et al., 2007), and they share 48% identity in amino acid sequence.

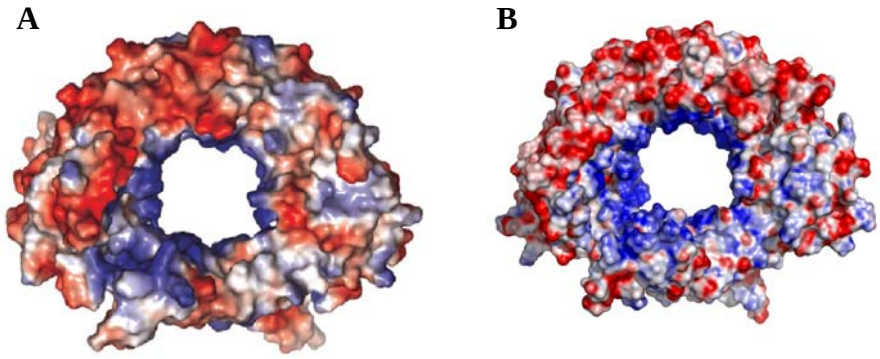


Figure 7-3 RdgC structures of *Pseudomonas aeruginosa* and *Escherichia coli*

A. RdgC structure of *P. aeruginosa* (Ha et al., 2007)
B. RdgC structure of *E. coli* (Briggs et al., 2007)

The two proteins share many of the conserved residues at both interfaces, and both have positively charged fingers (Figure 3-6). Although the group also failed to locate the bound dsDNA in the structure experimentally, they proposed the positively charged central hole as the binding channel (Ha et al., 2007). The mutant studies by the group also established the importance of both of the interfaces to dsDNA binding. Interestingly, every tested mutation on conserved positively charged residues along the DNA binding channel had a considerable effect on dsDNA binding, even including the replacements to a neutral residue (alanine). Every basic residue in the central hole is therefore indispensable for a tight dsDNA binding. Substitutions at either interfaces had a similar effect on dsDNA binding as I observed on *E. coli* RdgC: F120A (counterpart of F120T and F120S in this study) was detrimental, and R211D (counterpart of K211A) was comparable to the wild type. The only difference rose from Q212A (same as in this study), which moderately weakened the *P. aeruginosa* protein's affinity for a 414bp dsDNA, but which did not effect much to the *E. coli* protein. The Q212 residue was also proposed to interact

with R211 and K227 in *P. aeruginosa* RdgC (Figure7-4) (Ha et al., 2007). The discrepancy might be due to the different experimental conditions: the group performed the binding assay on agarose gel in TAE buffer (45 mM Tris-acetate at pH 8.3, 1 mM EDTA), whereas the one performed in this work was on poly-acrylamide gel in LIS buffer. Of course it is also possible that both results were genuine and robust, and the discrepancy simply reflects different functions the Q212 residue delivers for the two proteins. It is also worth noting the fact that the Q212A mutation might weaken the protein's DNA binding affinity which correlates well with the results of the *in vivo* assays, where *E. coli* $rdgC_{Q212A}$ could not substitute for wild type *rdgC* in any condition tested.

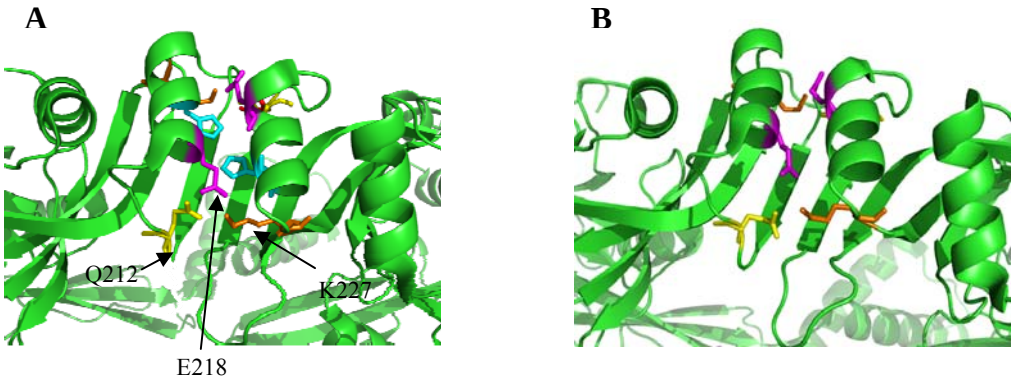


Figure 7-4 Horseshoe interface comparison between *E. coli* RdgC and *P. aeruginosa* RdgC

A. Horseshoe interface of *E. coli* RdgC
B. Horseshoe interface of *P. aeruginosa* RdgC

Corresponding residues are shaded in the same colour.

Given that *P. aeruginosa* RdgC has a structure very similar to the one of *E. coli*, it is fairly reasonable to suppose that *Neisseria* RdgC also adopts a similar protein fold. As the previous section discussed, although both proteins might

play a similar role in their native species, *N. meningitidis* RdgC could not substitute for the *E. coli* protein at all. In addition, it has a relatively better affinity for ssDNA than *E. coli* RdgC. Unfortunately, neither ssDNA binding of *P. aeruginosa* RdgC nor its *in vivo* substitution in *E. coli* has been tested. The two questions are therefore unable to be answered. I attempted to crystallise *N. meningitidis* RdgC, but the quality was too poor to be solved.

7.6 Future directions

This project offered a further understanding of RdgC as a DNA binding protein, with special reference to its structure and to how it loads onto DNA. However, how the RdgC protein is dislodged from DNA has not been studied. This is important in understanding why RdgC's DNA binding does not inhibit normal DNA processing, e.g. DNA replication and transcription. Competition experiments using various DNA molecules to compete for the RdgC already bound to DNA in conjugation with mutant studies might give insight into the mechanism of dissociation.

Another perspective focuses on protein-protein interactions, as brought by several implications. Some RdgC mutants with disrupted horseshoe interface did not show a defect in DNA binding activities, whereas they lost most or all functionality *in vivo* ($RdgC_{H222A}$ and $RdgC_{Q212A}$, see Chapter 4 for details). This leaves the horseshoe interface a potential as a protein interaction domain, although it is also possible that the interface mediates the formation of functional higher order oligomers. In the models proposed in 7.4.1, 7.4.2 and 7.4.3, RdgC are all required to interact with proteins, either to be guided (7.4.1

and 7.4.2), or to guide certain functional proteins to sites of interest (7.4.3). As mentioned, several proteins have been reported to interact with RdgC, but this needs further validation. Other proteins with potential to interact with RdgC include RecF and RecR. Nonetheless, RdgC has been found to antagonise RecA activities by competing for its substrates (Drees et al., 2006). If this is the main activity of RdgC *in vivo*, then the same experiment design will expect to yield different results on $RdgC_{Q212A}$ and $RdgC_{H222A}$, despite of their integral affinity for DNA.

As mentioned in section 7.5.2, if we are lucky to have the *rdgC* gene from *P. aeruginosa*, we can therefore perform ssDNA and circular DNA binding assays, and make relevant synthetic lethality constructs to investigate whether it is able to substitute the *E. coli rdgC*. With the results we can probably get more insight into the protein's function on the structural level.

7.7 Other ideas

This project and many of the previous studies proved a strong link between RdgC and the RecFOR mediated RecA loading. However, in strains lacking RecBC and SbcBC, deletion of *rdgC* makes the strains growth dependent on the remaining RecA activities, mostly RecFOR mediated. This indicates that RdgC is also involved in some DNA repair process parallel to the one mediated by RecA. Indeed, the inviability seen in *priA⁻ rdgC⁻* and *priA⁻ dnaC809,820 priC⁻ rdgC⁻* cells cannot be overcome by inactivating RecFOR or RecA (R. G. Lloyd, unpublished data), suggesting that the regulation of RecA is only part of RdgC's function *in vivo*.

The only conditions where the toxicity of RecA is revealed include *priA*$^-$ (R. G. Lloyd, unpublished data), *priA*$^-$ *rdgC*$^-$ *dnaC810* (R. G. Lloyd, unpublished data) and *priA*$^-$ *rdgC*$^-$ *dnaC212* (Moore et al., 2003). In these conditions, Replication restart relies on Rep and PriC to load DnaB. It is therefore a putative target of toxic RecA activities. Rep together with PriC act in a way similar to PriA, where Rep's 3'-5' helicase activity unwinds the lagging strand to provide a landing pad for PriC, which in turn mediates loading of DnaB (Sandler, 2000). The only vulnerable moment for RecA's attack is after a segment of lagging strand is unwound by Rep and before PriC is in position. This ssDNA gap can be used by RecFOR to load RecA to initiate illegitimate recombination.

If RdgC protects the process from being disrupted by RecA, then how is it achieved? The answer is not directly clear, as the ssDNA exposed by Rep is readily bound by SSB, and the opposite homologous leading strand does not appear to have any specific features that could attract enough RdgC for limiting recombination. A plausible answer is that RdgC could be directed by Rep to the unwound ssDNA and bind before SSB acts on it. In this case, Rep could be the putative guider as previously described in 7.4.2.

The regulation of RecA by RdgC might be more than just by competing common substrates. A direct prediction of the competing model is that overexpressing RdgC in *E. coli* should confer phenotypes similar to those of cells lacking RecFOR, which have greatly reduced resistance to UV damage. However, cells had overexpressed RdgC did not show any defects in growth or UV resistance (unpublished data), which could mean that RdgC regulating RecA activities is directed.

Given the profile of RdgC we know so far, it is more like a dsDNA binding protein equivalent to SSB. SSB is known to exert a variety of activities by binding to ssDNA, mainly by protecting covered ssDNA from being abused and removing its secondary structures, a usual waylayer for normal ssDNA processing. In reflection, most or all functions of RdgC reported so far under certain conditions could simply use its dsDNA binding property. More interestingly, both proteins have similar expression levels in cells.

References

Abramoff, M. D., Magelhaes, P. J. and Ram, S. J. (2004). "Image Processing with ImageJ." Biophotonics International **11**: 36-42.

Anderson, D. G. and Kowalczykowski, S. C. (1997). "The recombination hot spot chi is a regulatory element that switches the polarity of DNA degradation by the RecBCD enzyme." Genes Dev **11**: 571-81.

Anderson, D. G. and Kowalczykowski, S. C. (1997). "The translocating RecBCD enzyme stimulates recombination by directing RecA protein onto ssDNA in a chi-regulated manner." Cell **90**: 77-86.

Bachmann, B. J. (1996). Derivations and genotypes of some mutant derivatives of Escherichia coli K-12. Washington, D.C., ASM Press.

Beernink, H. and Morrical, S. (1999). "RMPs: recombination/replication mediator proteins." Trends Biochem Sci **24**: 385-9.

Bernhardt, T. G. and de Boer, P. A. (2004). "Screening for synthetic lethal mutants in Escherichia coli and identification of EnvC (YibP) as a periplasmic septal ring factor with murein hydrolase activity." Mol Microbiol **52**: 1255-69.

Bianco, P. R. and Kowalczykowski, S. C. (1997). "The recombination hotspot Chi is recognized by the translocating RecBCD enzyme as the single strand of DNA containing the sequence 5'-GCTGGTGG-3'." Proc. Natl. Acad. Sci. USA **94**: 6706-11.

Briggs, G. S., McEwan, P. A., Yu, J., Moore, T., Emsley, J. and Lloyd, R. G. (2007). "Ring structure of the Escherichia coli DNA-binding protein RdgC associated with recombination and replication fork repair." J Biol Chem **282**: 12353-7.

Butland, G., Peregrín-Alvarez, J. M., Li, J., Yang, W., Yang, X., Canadien, V., Starostine, A., Richards, D., Beattie, B., Krogan, N., Davey, M., Parkinson, J., Greenblatt, J. and Emili, A. (2005). "Interaction network containing conserved and essential protein complexes in *Escherichia coli*." Nature **433**: 531-7.

Cadman, C. J. and McGlynn, P. (2004). "PriA helicase and SSB interact physically and functionally." Nucleic Acids Res **32**: 6378-87.

Connelly, J. and Leach, D. (1996). "The sbcC and sbcD genes of Escherichia coli encode a nuclease involved in palindrome inviability and genetic recombination." Genes Cells **1**: 285-91.

Datsenko, K. A. and Wanner, B. L. (2000). "One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products." Proc Natl Acad Sci U S A **97**: 6640-5.

Dillingham, M. S. and Kowalczykowski, S. C. (2008). "RecBCD enzyme and the repair of double-stranded DNA breaks." Microbiol Mol Biol Rev **72**: 642-71.

Dixon, D. A. and Kowalczykowski, S. C. (1993). "The recombination hotspot χ is a regulatory sequence that acts by attenuating the nuclease activity of the E. coli RecBCD enzyme." Cell **73**: 87-96.

Doublie, S. (1997). "Preparation of Selenomethionyl Proteins for Phase Determination." Methods Enzymol **276**: 523-30§.

Drees, J. C., Chitteni-Pattu, S., McCaslin, D. R., Inman, R. B. and Cox, M. M. (2006). "Inhibition of RecA protein function by the RdgC protein from Escherichia coli." J. Biol. Chem **281**: 4708-17.

Drees, J. C., Lusetti, S. L., Chitteni-Pattu, S., Inman, R. B. and Cox, M. M. (2004). "A RecA filament capping mechanism for RecX protein." Mol Cell **15**: 789-98.

Gorbalenya, A. E. (2001). "Big nidovirus genome. When count and order of domains matter." Adv Exp Med Biol **494**: 1-17.

Gottesman, M. M., Hicks, M. L. and Gellert, M. (1973). "Genetics and function of DNA ligase in Escherichia coli." J. Mol. Biol **77**: 531-47.

Gregg, A., McGlynn, P., Jaktaji, R. and Lloyd, R. (2002). "Direct rescue of stalled DNA replication forks via the combined action of PriA and RecG helicase activities." Mol Cell **9**: 241-51.

Gupta, R., Golub, E., Bi, B. and Radding, C. (2001). "The synaptic activity of HsDmc1, a human recombination protein specific to meiosis." Proc. Natl. Acad. Sci. USA **98**: 8433-9.

Ha, J. Y., Kim, H. K., Kim, D. J., Kim, K. H., Oh, S. J., Lee, H. H., Yoon, H. J., Song, H. K. and Suh, S. W. (2007). "The recombination-associated protein RdgC adopts a novel toroidal architecture for DNA binding." Nucleic Acids Res **35**: 2671-81.

Heller, R. C. and Marians, K. J. (2007). "Non-replicative helicases at the replication fork." DNA Repair (Amst) **6**: 945-52.

Hobbs, M. D., Sakai, A. and Cox, M. M. (2007). "SSB limits RecOR binding onto single strand DNA." J. Biol. Chem **282**: 11058-67.

Holm, L. and Sander, C. (1993). "Protein structure comparison by alignment of distance matrices." J. Mol. Biol **233**: 123-38.

Huang, J., Lu, J., Barany, F. and W, C. (2001). "Multiple cleavage activities of endonuclease V from Thermotoga maritima: recognition and strand nicking mechanism." Biochemistry **40**: 8738-48.

Kantake, N., Madiraju, M. V., Sugiyama, T. and Kowalczykowski, S. C. (2002). "Escherichia coli RecO protein anneals ssDNA complexed with its cognate ssDNA-binding protein: A common step in genetic recombination." Proc Natl Acad Sci U S A **99**: 15327-32.

Kelman, Z. and O'Donnell, M. (1995). "Structural and functional similarities of prokaryotic and eukaryotic DNA polymerase sliding clamps." Nucleic Acids Res **23**: 3613-20.

Kodadek, T., Wong, M. L. and Albertss, B. M. (1988). "The mechanism of homologous DNA strand exchange catalyzed by the bacteriophage T4 uvsX and gene 32 proteins." J Biol Chem **263**: 9427-36.

Kogoma, T., Cadwell, G., Barnard, K. and Asai, T. (1996). "The DNA replication priming protein, PriA, is required for homologous recombination and double-strand break repair." J. Bacteriol **178**: 1258-64.

Kolodner, R. D. (1995). "Mismatch repair: mechanisms and relationship to cancer susceptibility." Trends Biochem Sci **20**: 397-401.

Kong, X., Onrust, R., O'Donnell, M. and Kuriyan, J. (1992). "Three-dimensional structure of the beta subunit of *E. coli* DNA polymerase III holoenzyme: a sliding DNA clamp." Cell **69**: 425-37.

Konrad, E. B., Modrich, P. and Lehman, I. R. (1973). "Genetic and enzymatic characterization of a conditional lethal mutant of Escherichia coli K12 with a temperature-sensitive DNA ligase." J. Mol. Biol **77**: 519-29.

Krishna, T., Kong, X., Gary, S., Burgers, P. and Kuriyan, J. (1994). "Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA." Cell **79**: 1233-43.

Kuzminov, A. (1999). "Recombinational repair of DNA damage in *Escherichia coli* and bacteriophage λ." Microbiol Mol Biol Rev **63**: 751-813.

Lane, H. and Denhardt, D. (1975). "The rep mutation. IV. Slower movement of replication forks in Escherichia coli rep strains." J. Mol. Biol **97**: 99-112.

Lestini, R. and Michel, B. (2007). "UvrD controls the access of recombination proteins to blocked replication forks." EMBO J **26**: 3804-14.

Liu, J. and Marians, K. J. (1999). "PriA-directed assembly of a primosome on D loop DNA." J Biol Chem **274**: 25033-41.

Lloyd, R. G. and Buckman, C. (1985). "Identification and genetic analysis of sbcC mutations in commonly used recBC sbcB strains of Escherichia coli K-12." J. Bacteriol **164**: 836-44.

Luisi-DeLuca, C. and Kolodner, R. (1994). "Purification and characterization of the Escherichia coli RecO protein. Renaturation of complementary single-stranded DNA molecules catalyzed by the RecO protein." J Mol Biol **236**: 124-38.

Lusetti, S., Voloshin, O., Inman, R., Camerini-Otero, R. and Cox, M. (2004b). "The DinI protein stabilizes RecA protein filaments." J Biol Chem **279**: 30037-46.

Lusetti, S. L., Drees, J. C., Stohl, E. A., Seifert, H. S. and Cox, M. M. (2004a). "The DinI and RecX proteins are competing modulators of RecA function." J Biol Chem **279**: 55073-9.

Magner, D. B., Blankschien, M. D., Lee, J. A., Pennington, J. M., Lupski, J. R. and Rosenberg, S. M. (2007). "RecQ Promotes Toxic Recombination in Cells Lacking Recombination Intermediate-Removal Proteins " Cell **26**: 273-86.

Mahdi, A. A., Buckman, C., Harris, L. and Lloyd, R. G. (2006). "Rep and PriA helicase activities prevent RecA from provoking unnecessary recombination during replication fork repair." Genes Dev **20**: 2135-47.

Makharashvil, N., Mi, T., Koroleva, O. and Korolev, S. (2009). "RecR-mediated Modulation of RecF Dimer Specificity for Single- and Double-stranded DNA." J Biol Chem **284**: 1425-34.

Matson, S. W. and Robertson, A. B. (2006). "The UvrD helicase and its modulation by the mismatch repair protein MutL." Nucleic Acids Res **34**: 4089-97.

McCullough, A. K., Dodson, M. L. and Lloyd, R. S. (1999). "Initiation of base excision repair: glycosylase mechanisms and structures." Annu Rev Biochem **68**: 255-85.

Mehr, I. J., Long, C. D., Serkin, C. D. and Seifert, H. S. (2000). "A Homologue of the Recombination-Dependent Growth Gene, rdgC, Is Involved in Gonococcal Pilin Antigenic Variation." Genetics **154**: 523-32.

Mehr, I. J. and Seifert, H. S. (1998). "Differential roles of homologous recombination pathways in Neisseria gonorrhoeae pilin antigenic variation, DNA transformation and DNA repair." Mol Microbiol **30**: 697-710.

Mertens, N., Remaut, E. and Fiers, W. (1995). "Tight transcriptional control mechanism ensures stable high-level expression from T7 promoter-based expression plasmids." Biotechnology (N Y) **13**: 175-9.

Miller, J. H. (1972). Experiments in molecular genetics. N.Y., Cold Spring Harbor Laboratory Press,U.S.

Moore, T. (2002). The RdgC Protein of *Escherishia coli*. The Institute of Genetics. Nottingham, University of Nottingham. **Doctoral**.

Moore, T., McGlynn, P., Ngo, H. P., Sharples, G. J. and Lloyd, R. G. (2003). "The RdgC protein of Escherichia coli binds DNA and counters a toxic effect of RecFOR in strains lacking the replication restart protein PriA." Embo Journal **22**: 735-45.

Moore, T., Sharples, G. J. and Lloyd, R. G. (2004). "DNA binding by the meningococcal RdgC protein, associated with pilin antigenic variation." J. Bacteriol **186**: 870-4.

Moreau, P. (1987). "Effects of overproduction of single-stranded DNA-binding protein on RecA protein-dependent processes in *Escherichia coli*." J Mol Biol **194**: 621-34.

Morel, P., Hejna, J. A., Ehrlich, S. D. and Cassuto, E. (1993). "Antipairing and strand transferase activities of *E. coli* helicase-II (UvrD)." Nucleic Acids Res **21**: 3205-9.

Morimatsu, K. and Kowalczykowski, S. C. (2003). "RecFOR Proteins Load RecA Protein onto Gapped DNA to Accelerate DNA Strand Exchange-A Universal Step of Recombinational Repair " Mol Cell **11**: 1337-47.

Murphy, A. K., Tammaro, M., Cortazar, F., Gindt, Y. M. and Schelvis, J. P. (2008). "Effect of the cyclobutane cytidine dimer on the properties of Escherichia coli DNA photolyase." J Phys Chem B **112**: 15217-26.

Murphy, L. D., Rosner, J. L., Zimmerman, S. B. and Esposito, D. (1999). "Identification of Two New Proteins in Spermidine Nucleoids Isolated from *Escherichia coli*." J. Bacteriol **181**: 3842-4.

Nurse, P., Zavitz, K. H. and Marians, K. J. (1991). "Inactivation of the Escherichia coli priA DNA replication protein induces the SOS response." J Bacteriol **173**: 6686-93.

Ogawa, T., Yu, X., Shinohara, A. and Egelman, E. (1993). "Similarity of the yeast RAD51 filament to the bacterial RecA filament." Science **259**: 1896-9.

Rocha, E. P. C., Cornet, E. and Michel, B. (2005). "Comparative and evolutionary analysis of the bacterial homologous recombination systems." PLoS Genet **1**: e15.

Rouvière-Yaniv, J., Yaniv, M. and Germond, J.-E. (1979). "E. coli DNA binding protein HU forms nucleosome-like structure with circular double-stranded DNA." Cell **17**: 265-74.

Ryder, L., Sharples, G. J. and Lloyd, R. G. (1996). "Recombination-dependent growth in exonuclease-depleted recBC sbcBC strains of Escherichia coli K-12." Genetics **143**: 1101-14.

Sakai, A. and Cox, M. M. (2009). "RecFOR and RecOR as Distinct RecA Loading Pathways." J Biol Chem **284**: 3264-72.

Sancar, A. (1996). "DNA excision repair." Annu Rev Biochem **65**: 43-81.

Sandler, S. J. (1996). "Overlapping functions for recF and priA in cell viability and UV-inducible SOS expression are distinguished by dnaC809 in Escherichia coli K-12." Mol Microbiol **19**: 871-80.

Sandler, S. J. (2000). "Multiple Genetic Pathways for Restarting DNA Replication Forks in Escherichia coli K-12." Genetics **155**: 487-97.

Sassanfar, M. and Roberts, J. (1990). "Nature of the SOS-inducing signal in *Escherichia coli*. The involvement of DNA replication." J. Mol. Biol **212**: 79-96.

Seitz, E. M., Brockman, J. P., Sandler, S. J., Clark, A. J. and Kowalczykowski, S. C. (1998). "RadA protein is an archaeal RecA protein homolog that catalyzes DNA strand exchange." Genes Dev **12**: 1248-53.

Skaar, E. P., Lazio, M. P. and Seifert, H. S. (2002). "Roles of the recJ and recN Genes in Homologous Recombination and DNA Repair Pathways of Neisseria gonorrhoeae " J. Bacteriol **184**: 919-27.

Snijder, E. J., Bredenbeek, P. J., Dobbe, J. C., Thiel, V., Ziebuhr, J., Poon, L. L., Guan, Y., Rozanov, M., Spaan, W. J. and Gorbalenya, A. E. (2003). "Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage." J Mol Biol **331**: 991-1004.

Sprague, K. U., Faulds, D. H. and Smith, G. R. (1978). "A single base-pair change creates a Chi recombinational hotspot in bacteriophage lambda." Proc. Natl. Acad. Sci. USA **75**: 6182-6.

Sriskanda, V. and Shuman, S. (2001). "A second NAD+-dependent DNA ligase (LigB) in Escherichia coli." Nucleic Acids Research **29**: 4930-4.

Stohl, E. A., Brockman, J. P., Burkle, K. L., Morimatsu, K., Kowalczykowski, S. C. and Seifert, H. S. (2003). "Escherichia coli RecX inhibits RecA recombinase and coprotease activities in vitro and in vivo." J Biol Chem **278**: 2278-85.

Studier, F. W. and Moffatt, B. A. (1986). "Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes." J. Mol. Biol **189**: 113-30.

Sung, P. (1994). "Catalysis of ATP-dependent homologous DNA pairing and strand exchange by yeast RAD51 protein." Science **265**: 1241-3.

Tabor, S. and Richardson, C. C. (1985). "A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes." Proc. Nati. Acad. Sci. USA **82**: 1074-8.

Taylor, A. F. and Smith, G. R. (1985). "Substrate specificity of the DNA unwinding activity of the RecBC enzyme of *Escherichia coli*."

Tessmer, I., Moore, T., Lloyd, R. G., Wilson, A., Erie, D. A., Allen, S. and Tendler, S. J. B. (2005). "AFM studies on the role of the protein RdgC in bacterial DNA recombination." J. Mol. Biol **350**: 254-62.

Thiel, V., Ivanov, K. A., Putics, A., Hertzig, T., Schelle, B., Bayer, S., Weissbrich, B., Snijder, E. J., Rabenau, H., Doerr, H. W., Gorbalenya, A. E. and Ziebuhr, J. (2003). "Mechanisms and enzymes involved in SARS coronavirus genome expression." J Gen Virol **84**: 2305-15.

Umezu, K. and Kolodner, R. (1994). "Protein interactions in genetic recombination in *Escherichia coli*. Interactions involving RecO and RecR overcome the inhibition of RecA by single-stranded DNA-binding protein." J. Biol. Chem **269**: 30005-13.

Veaute, X., Delmas, S., Selva, M., Jeusset, J., Cam, E. L., I. Matic, F. F. and Petit., M. A. (2005). "UvrD helicase, unlike Rep helicase, dismantles RecA nucleoprotein filaments in *Escherichia coli*." EMBO J **24**: 180-9.

Viswanathan, M. and Lovett, S. T. (1998). "Single-strand DNA-specific exonucleases in Escherichia coli. Roles in repair and mutation avoidance." Genetics **149**: 7-16.

Webb, B., Cox, M. and Inman, R. (1997). "Recombinational DNA repair: the RecF and RecR proteins limit the extension of RecA filaments beyond single-strand DNA gaps." Cell **91**: 347-56.

Webb, B. L., Cox, M. M. and Inman, R. B. (1995). "An interaction between the Escherichia coli RecF and RecR proteins dependent on ATP and double-stranded DNA." J Biol Chem **270**: 31397-404.

Willetts, N. S., Clark, A. J. and Low, B. (1969). "Genetic Location of Certain Mutations Conferring Recombination Deficiency in *Escherichia coli*." J. Bacteriol **97**: 244-9.

Yancey-Wrona, J. and Matson, S. (1992). "Bound Lac repressor protein differentially inhibits the unwinding reactions catalyzed by DNA helicases." Nucleic Acids Res **20**: 6713-21.

Yao, M., Hatahet, Z., Melamede, R. J. and Kow, Y. W. (1994). "Purification and characterization of a novel deoxyinosine-specific enzyme, deoxyinosine 3' endonuclease, from *Escherichia coli*." J. Biol. Chem **269**: 16260-8.

Yao, M. and Kow, Y. (1995). "Interaction of deoxyinosine 3'-endonuclease from Escherichia coli with DNA containing deoxyinosine." J Biol Chem **270**: 28609-16.

Zhai, Y., Sun, F., Li, X., Pang, H., Xu, X., Bartlam, M. and Rao, Z. (2005). "Insights into SARS-CoV transcription and replication from the structure of the nsp7-nsp8 hexadecamer." Nat Struct Mol Biol **12**: 980-6.

Appendix 1

Proteins affinity for Heparin column

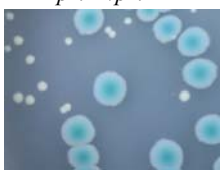
Protein	*NaCl concentration range at elution peak (M) (Heparin column)*
E. coli native RdgC	0.70-0.90
$RdgC_{K211A}$	0.70-0.90
$RdgC_{Q212A}$	0.70-0.90
$RdgC_{E218R}$	0.70-0.90
$RdgC_{H222A}$	0.70-0.90
$RdgC_{K227A}$	0.50-0.70
$RdgC_{R118A}$	0.70-0.90
$RdgC_{R118C}$	0.70-0.90
$RdgC_{\Delta F}$	0.40-0.60
$RdgC_{\Delta FR118C}$	0.40-0.60
$RdgC_{F120T}$	0.70-0.90
$RdgC_{F120S}$	0.70-0.90
$RdgC_{R118CF120T}$	0.70-0.90
$RdgC_{FTM}$	0.50-0.70
N. Meningitidis RdgC	0.70-0.90
Se-Met RdgC	0.70-0.90

Appendix 2

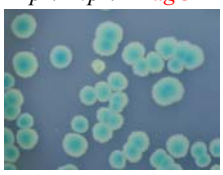
Synthetic lethality assay for *N. meningitidis rdgC* (*NrdgC*)

On LB

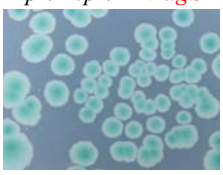
YJ047
priA^{+}/*priA*$^{-}$



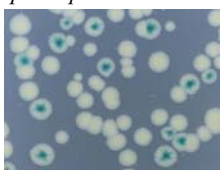
AM1918
priA^{+}/*priA*$^{-}$ *rdgC*$^{-}$



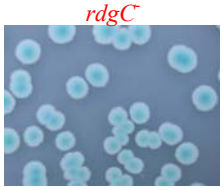
AM1913
priA^{+}/*priA*$^{-}$ *NrdgC*



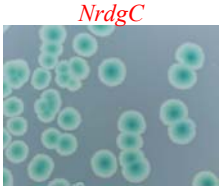
YJ048
priA^{+}/*priA*$^{-}$ *dnaC810*



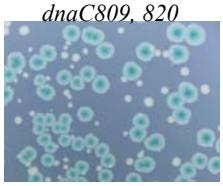
AM1919
priA^{+}/*priA*$^{-}$ *dnaC810*
rdgC^{-}



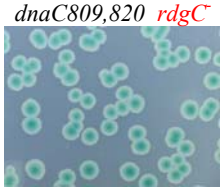
AM1914
priA^{+}/*priA*$^{-}$ *dnaC810*
NrdgC



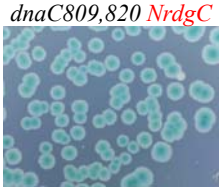
AM2046
priA^{+}/*priA*$^{-}$ *priC*$^{-}$
dnaC809, 820



YJ026
priA^{+}/*priA*$^{-}$ *priC*$^{-}$
dnaC809,820 *rdgC*$^{-}$

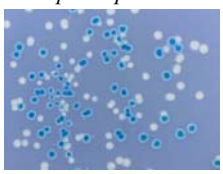


YJ040
priA^{+}/*priA*$^{-}$ *priC*$^{-}$
dnaC809,820 NrdgC

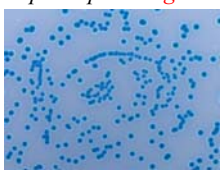


On MA

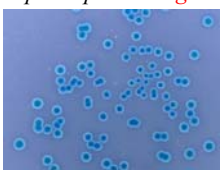
YJ047
priA^{+}/*priA*$^{-}$



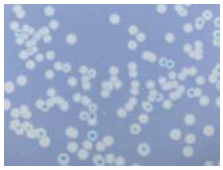
AM1918
priA^{+}/*priA*$^{-}$ *rdgC*$^{-}$



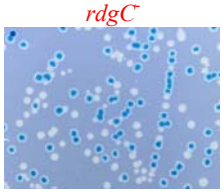
AM1913
priA^{+}/*priA*$^{-}$ *NrdgC*



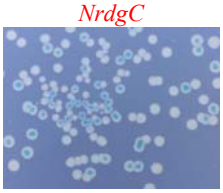
YJ048
priA^{+}/*priA*$^{-}$ *dnaC810*



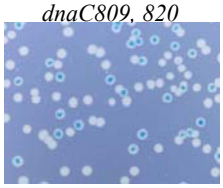
AM1919
priA^{+}/*priA*$^{-}$ *dnaC810*
rdgC^{-}



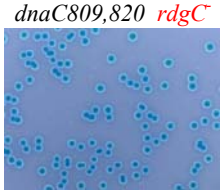
AM1914
priA^{+}/*priA*$^{-}$ *dnaC810*
NrdgC



AM2046
priA^{+}/*priA*$^{-}$ *priC*$^{-}$
dnaC809, 820



YJ026
priA^{+}/*priA*$^{-}$ *priC*$^{-}$
dnaC809,820 *rdgC*$^{-}$



YJ040
priA^{+}/*priA*$^{-}$ *priC*$^{-}$
dnaC809,820 NrdgC