Identification of pre-synaptic processing proteins from *Bacteroides fragilis*

Frances Parry

Thesis presented for the degree of Doctor of Philosophy

The University of Edinburgh

2010

Declaration

I declare that this is my own work and that any contribution made by other parties is clearly acknowledged

Frances Parry

July 2010

Acknowledgements

I would like to thank my supervisor, Garry Blakely, for his support throughout my project. I will always be grateful for all he has taught me. I would like to thank all staff and students of Darwin building. I would like to have special mention for Robyn Emmins, Jen Bell and Lorna Holding; thank you for all the fun. I would also like to thank the BBSRC for funding of my PhD.

This thesis is dedicated to my family.

Abstract

The repair of DNA double-strand breaks (DSBs) is required for the survival of all organisms. In bacteria, DNA DSBs can occur during normal housekeeping processes such as DNA replication or by exogenous damage due to chemicals or radiation. DSBs will compromise the integrity of the genome if left un-repaired, and can be fatal to an organism. Repair of DSBs by homologous recombination (HR) replicates missing chromosomal regions before joining of the separated DNA ends. In *Escherichia coli* the HR repair steps are; pre-synapsis, synapsis and post-synapsis. In the pre-synaptic stage a DSB is processed into a 3′ single-strand overhang, the substrate required for strand invasion in the synapsis stage and the eventual repair of the DSB. At present there are three identified pre-synapsis systems involved in recombination in bacteria; represented by the AdnAB, AddAB and the RecBCD protein complexes. Each system functions in a similar manner but differ in the physical composition of the machinery.

This project investigated the pre-synaptic system of *Bacteroides fragilis* NCTC9343. Genes encoding putative pre-synapsis proteins were initially identified through analysis of the NCTC9343 genome. The function of these proteins was investigated in vivo by rescue of a repair-deficient strain of *E. coli*. This demonstrated that *Bacteroides fragilis* encodes a two component system, where both genes products are required to work in concert for pre-synaptic processing of DSBs. The identified genes were BF2192 and BF2191, and have been renamed *addA* and *addB*, respectively. To further examine the role of the AddAB proteins in DSB repair, a *Bacteroides fragilis* strain with a deletion of *addAB* was constructed and shown to be extremely sensitive to DNA damaging agents. The AddAB complex was purified and found to be an ATP-dependant helicase and exonuclease that acted on double-stranded DNA ends. In conclusion, this project has identified the proteins involved in pre-synaptic processing of DSBs in *B. fragilis* NCTC9343, consisting of AddAB homologues, and shown their protective role in repair of DNA damage.

Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1 Double Strand Break Repair

The repair of DNA double-strand breaks (DSBs) is required for the survival of all organisms. In bacteria, DNA DSBs can occur during normal housekeeping processes such as DNA replication (regression of stalled forks) or by exogenous damage due to chemicals or radiation. DSBs can compromise the integrity of the genome if left un-repaired, and can be fatal to an organism; DNA may be degraded by endogenous single or double-strand exonucleases, chromosomes cannot replicate properly and illegitimate and/or mutagenic rearrangements could occur. Incorrect repair can cause mutations, such as deletions, therefore accurate and efficient repair of DNA DSBs without loss of genetic information is important for maintaining genome integrity and survival of the organism (Reviewed by Nowosielska, 2007).

There are two known mechanisms for repairing DNA DSBs, the homology independent mechanism, termed non-homologous end joining (NHEJ) and the homology dependant mechanism, termed homologous recombination (HR). In NHEJ DSB repair, double-stranded (ds) DNA ends are joined together irrespective of homology; this causes a high rate of mutations. HR DSB repair requires an intact copy of the missing chromosomal DNA regions to act as template, so that the absent regions can be replicated before joining of the separated ends. The homology dependant method is more efficient at repairing without errors, since any missing DNA is replaced, and is the preferred method of repair in bacteria (reviewed by Cromie *et al*., 2001)

1.1.1 DSB repair by Non-Homologous End Joining

NHEJ has been predominately studied in Eukaryotes. In the G1 phase of the eukaryotic cell cycle there is only one copy of each chromosome present, therefore NHEJ is the pathway used for DSB repair. A number of accessory proteins are required to process the DNA template, but the essential proteins involved in joining DNA ends are Ku, a DNA end binding protein, and Lig4, a DNA ligase. The Ku

protein binds DSB DNA ends, protects them from degradation and brings the two free ends into close proximity. Ku stimulates ligase activity at the DNA ends by recruiting a complex of proteins that contains Lig4 (reviewed by Hefferin & Tomkinson, 2005).

Until recently it was believed that DSB repair in bacteria was only mediated via homologous recombination, but bioinformatics analysis identified Ku-like and Lig-like proteins in several bacterial genomes. Studies have shown NHEJ to be present in the genera *Mycobacterium* and *Bacillus* (Della *et al*., 2004; Weller *et al*., 2002). The proteins are thought to act in a similar manner to the eukaryote proteins, but in contrast to eukaryote NHEJ, bacterial NHEJ does not require any accessory proteins, perhaps reflecting the less complex architecture of chromosomes in prokaryotes (reviewed by Pitcher *et al.*, 2007).

NHEJ can be useful to bacteria when there is only a single copy of the chromosome present and there is no homologous segment available for DSB repair by HR. This could happen during stationary phase, or sporulation/spore stage (of spore forming bacteria). Since DSB HR cannot operate in these situations it would seem appropriate to have a system for repair that does not require a homologous template. *B. subtilis* forms spores and employs NHEJ during this stage of its life cycle. The *B. subtilis* Ku and LigD proteins are up-regulated when sporulation begins; GFP tagged Ku and LigD are associated with the nucleoid and then in the fore-spore, reflecting the increased need for NHEJ during this stage of the life cycle (Wang *et al*., 2006). *B. subtilis Ku* and *LigD* mutants displayed increased sensitivity to DNA damaging agents, UV, X-ray irradiation and hydrogen peroxide during the spore stage implying a role in DNA repair (Moeller *et al*., 2007).

1.1.2 DSB repair by Homologous Recombination

The disadvantage of the NHEJ pathway is the possible loss of DNA during joining of unmatched ends. The advantage of HR is that it allows the genome to be maintained more accurately; since it employs homology or near homologous sequences for repair, thus minimising the loss of DNA from the genome. In bacteria there are two distinctive structural classes of enzymes involved in pre-synaptic processing; the RecBCD and the AddAB/RexAB protein complexes. The RecBCD system is found in *E. coli* and is seen as the paradigm for HR DSB repair, but it is becoming apparent that the AddAB/RexAB system is more widespread than previously thought (Cromie, 2009). The process of DSB repair by HR is highly conserved, though proteins involved vary in different organisms. Since DSB repair has been most extensively studied in *E. coli*, this will be the model used for the following description of DSB repair (Reviewed by Kuzminov, 1999). For HR to proceed a second intact copy of the chromosome is required as a template for subsequent DNA synthesis; in bacteria this can be a newly replicated or partially replicated chromosome. HR repair of DNA DSBs consists of three stages; pre-synapsis, synapsis and post-synapsis (Figure 1.1.2.1). In *E. coli*, pre-synapsis is mainly mediated by the RecBCD complex; a DNA DSB is processed by RecBCD to generate 3′ single-stranded (ss) DNA coated with the strand-exchange protein RecA (Figure 1.1.2.1 a, b, c and d); this is the required substrate for the next stage, synapsis. During synapsis the RecA coated 3′ ss DNA invades the homologous duplex to form a D-loop structure (Figure 1.1.2.1 e). RecA facilitates pairing of the 3′ DNA strand with regions of homology on the intact DNA duplex. To initiate replication PriA enters the D-loop and recruits DNA Polymerase III to the 3′ strand (Figure 1.1.2.1). Once the strand has been replicated the final stage of recombination can proceed. In post-synapsis, the D-loop is converted to a holiday junction structure with four ds DNA arms. These are processed by the branch migration proteins RuvAB and the branch resolution protein RuvC to produce two separate intact DNA duplexes (Figure 1.1.2.1 g and h).

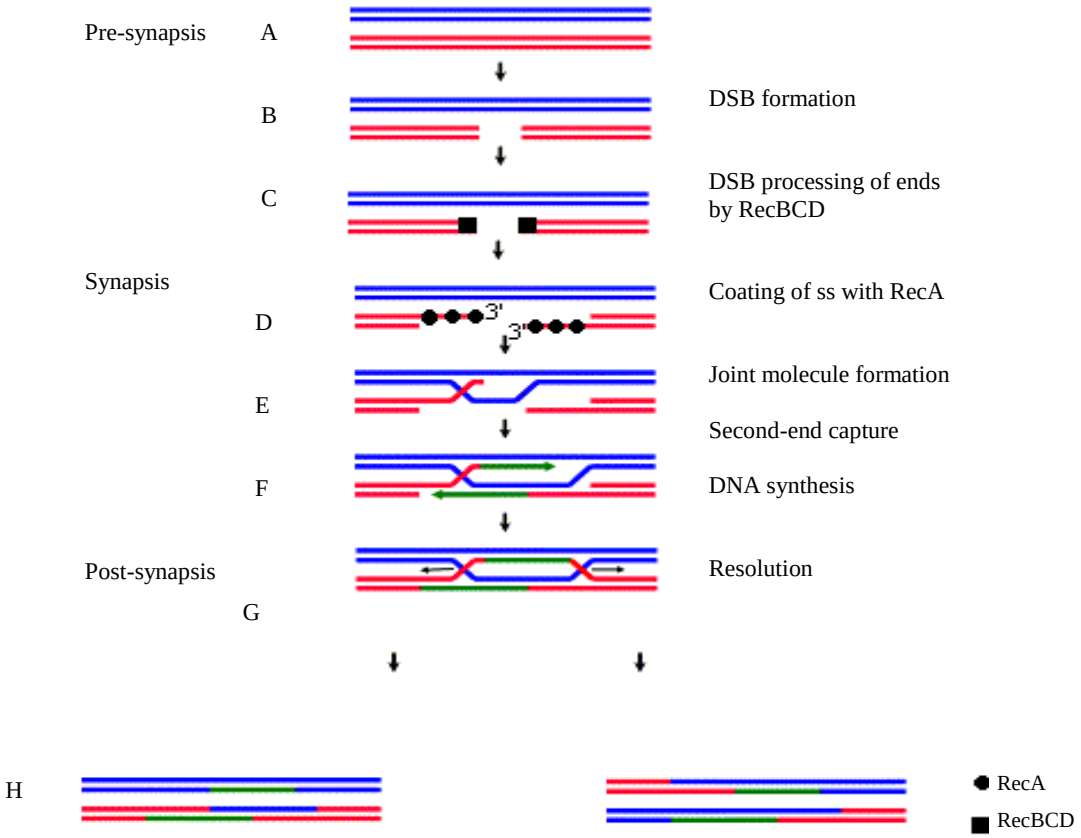


Figure 1.1.2.1: HR DNA DSB Repair

Representation of DSB repair via HR. The three stages of DSB repair are shown: pre-synapsis (B-D) synapsis (E-F) post-synapsis (F-H). In *E. coli* pre-synapsis involves DSB processing by the enzyme RecBCD to produce a 3′ overhang coated in RecA. This is the substrate required for synapsis. Strand invasion can occur followed by DNA strand exchange generating a joint molecule between damaged and undamaged duplexes. Sequence information that is missing at the DSB site is restored by DNA synthesis. The interlinked molecules are then processed by branch migration (indicated by right and left arrows), Holliday junction resolution and DNA ligation. Red and Blue represent homologous sequences. Green represents newly synthesised DNA. Black small circles represent RecA protein. RecBCD is represented by a large black square. Adapted from Dudáš & Chovanec (2004).

1.2 Pre-synaptic processing in *E. coli*

To date, *E. coli* is the most studied organism for the pre-synaptic processing step of HR DSB repair. In *E. coli* a least one DSB occurs per round of replication, and 99% of DSBs in *E. coli* are processed by the RecBCD complex (Cox *et al*., 2000; Reviewed by Kowalczykowski, 2000). RecBCD is large complex (330kDa) consisting of three proteins RecB, RecC, and RecD. RecB is 134kDa, RecC is 129kDa and RecD is 67kDa (Finch *et al.*, 1986a; Finch *et al.*, 1986b; Finch *et al.*, 1986c). The complex has ATPase, helicase, nuclease and RecA loading activity, allowing processing of DSBs into the required substrate for synapsis, which is a 3′ ss overhang coated with RecA (Figure 1.2.1). RecBCD loads onto ds ends and translocates along the double helix unwinding the two strands. As it unwinds it degrades both strands (with a preference for the 3′ strand), until it meets an eight base pair sequence (chi site) on the DNA that initiates a change in its activity. Unwinding continues at a reduced rate while degradation of the 3′ strand stops and RecA is actively loaded onto the 3′ strand. RecBCD then disassociates from the DNA leaving the required substrate ready for synapsis.

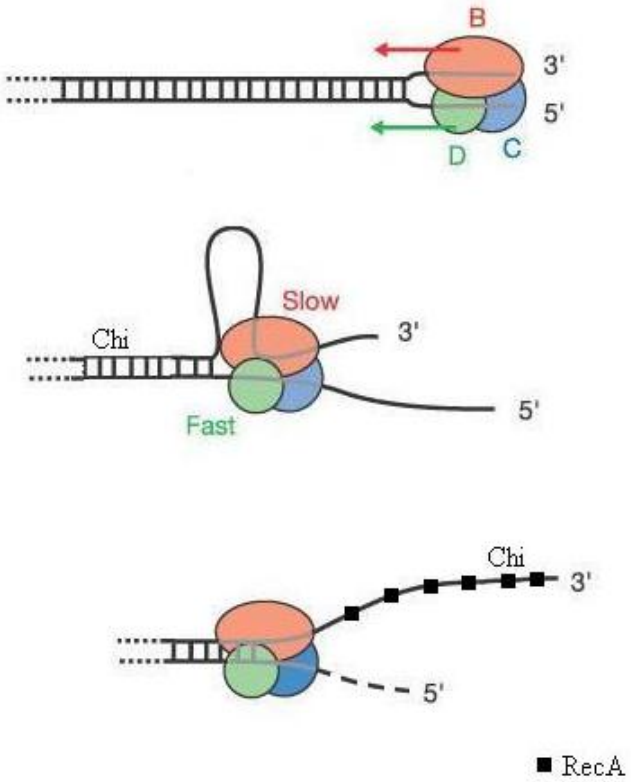


Figure 1.2.1: Pre-synaptic processing by RecBCD

RecBCD attaches to double-stranded DNA ends. RecB translocates along the 3′ strand while RecD translocates along the 5′ strand. The RecB and RecD helicase activities unwind the DNA duplex while the exonuclease activity of RecB degrades the strands. On encountering a chi site, exonuclease activity is attenuated and the 5′ strand is preferentially degraded while the 3′ strand is coated in RecA. Strand invasion can now proceed. Adapted from Spies *et al*., (2003).

1.2.1 Loading

Dimeric and monomeric forms of the complex have been purified but only the monomeric form of RecBCD is able to bind DNA ends (Taylor and Smith 1995). Wong *et al*., (2005) showed that ATP is not required for binding, which concurs with the binding shown by the crystal structure of the RecBCD initiation complex formed in the absence of ATP (Singleton *et al.,* 2004). Mg^{2+} is not essential but it does enhance binding (Taylor and Smith, 1995).

RecBCD can load onto blunt or nearly blunt ds DNA ends. Experiments showed that RecBCD does not load onto circular DNA; competition experiments with two possible RecBCD substrates, ss or ds DNA, showed that RecBCD has a preference for ds blunt ends (Roman and Kowlaczykowski, 1989) The DNA DSB ends do not have to be perfectly blunt as RecBCD can bind ends with small overhangs. Taylor and Smith (1995) demonstrated that a 4bp 5′ overhang was preferable to a completely blunt substrate.

The preference for blunt or near blunt ends can be explained by the presence of the two helicase subunits in the complex; the RecD and RecB helicases load onto and unwind strands of opposite polarities. The crystal structure of the RecBCD initiation complex bound to DNA showed the 3′ strand passed across RecB, close to the RecB binding cavity, where it was expected the 3′ strand would enter (Figure 1.2.1.1; Singleton *et al.,* 2004). No interaction of DNA with RecD was seen, probably due the proteins being in the initiation complex stage, but the 5′ DNA tail passed through a tunnel in RecC that opens onto the RecD motor domain (Figure 1.2.1.1; Singleton *et al.,* 2004).

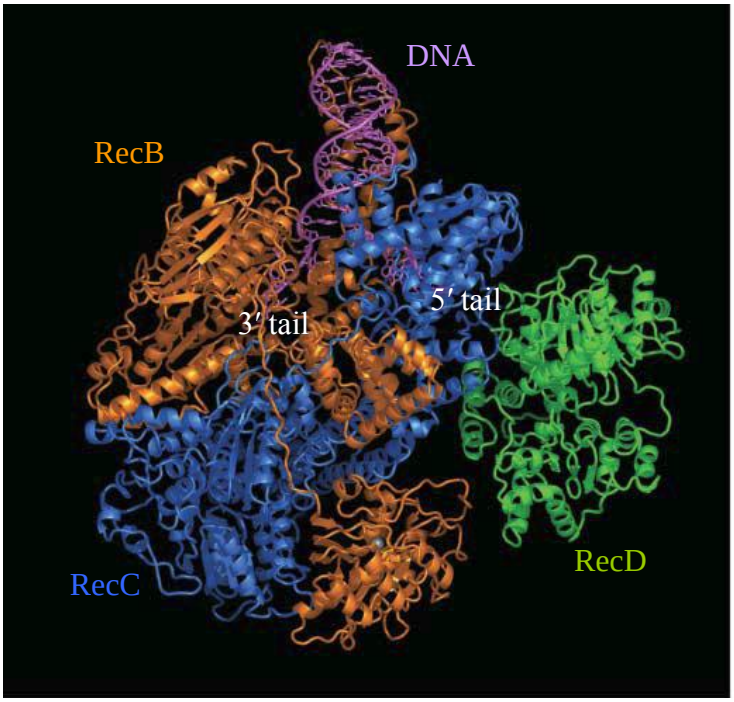


Figure 1.2.1.1: RecBCD initiation complex associated with ds DNA

RecBCD associates to form a heterotrimeric complex that binds to ds ends. The double helix splits across the RecC pin-like structure. The 5′ terminating strand is funnelled through a tunnel in RecC which opens on to the motor domain of RecD. The 3′ terminating strand passes through RecC to RecB. Adapted from Singleton *et al*., (2004).

The isolated RecB subunit has a preference for binding ss DNA (Phillips *et al.*, 1997). Complexes that were either RecB helicase deficient or RecD helicase deficient were used by Dillingham *et al*., (2005) to test binding of each motor. RecBCD belongs to the Super Family 1 helicases, indicative of this family is the presence of 7 helicase motifs in the amino acid sequence, particularly the Walker A motif which binds ATP (reviewed by Singleton and Wigley, 2002). Lysine to glutamine substitutions in Walker A motifs of the helicases allow binding of ATP but prevent hydrolysis, and therefore abolishe helicase activity. Both defective complexes could still bind ds DNA, but with a lower affinity compared to wild type. Preferences for binding substrates were observed for the helicase-defective complexes, the RecD motor deficient complex showed a slight preference for small

3′ overhangs, while the RecB motor deficient complex could only bind and unwind if there was a short 5′ overhang. This showed that RecD loads onto the 5′ end and RecB loads onto the 3′ end (Dillingham *et al.,* 2005). This was in keeping with a UV cross-linking study that showed RecB cross-linked to the 3′ end of DNA and RecD cross-linked to the 5′ end (Ganesan and Smith, 1993 cited by Singleton *et al.,* 2004). It would seem that the physical proximity of the two motors in the complex limits the binding substrate to ds DNA blunt/near blunt ends.

1.2.2 Helicase (pre-chi)

The binding preference of the RecD and RecB subunits for strands of opposite polarity is indicative of the helicase activity of each subunit; RecD has 5′-3′ activity and RecB has 3′-5′ activity. Purified RecB subunit can bind and translocate along ssDNA in a 3′-5′ direction (Phillips *et al.,* 1997). The RecD ATPase activity is also stimulated in the presence of ss DNA (Chen *et al.,* 1997). Dillingham *et al.,* (2003) observed that the RecD subunit can unwind a 40bp duplex if it had a 5′ terminated tail. There was no helicase activity on a blunt substrate and 5′-3′ unwinding was reduced 10 fold on a ds DNA substrate with a 3′ overhang (Dillingham *et al.,* 2003). These data indicated that RecD and RecB are ss DNA helicases with preference for strands of opposite polarity.

The translocation rates of RecBCD, along DNA before chi, are extremely rapid, up to 1350bps^{-1} (Dillingham *et al.,* 2005). The reasons why this is required and the mechanisms used to attain this speed will be discussed in section 1.2.8. Many studies have looked at the helicase properties of the RecBCD complex by mutating the ATPase motif in either of the motor units of the complex. Taylor and Smith (2003) used electron microscopy to observe the polarity and relative speeds that each subunit contributes. Complexes were made that had alterations in the Walker A motif of either subunit; RecB^{K29Q}CD and RecBCD^{K177Q}. Complexes with a motor-deficient RecD, unwound DNA at about 20% of wild type rates (73 nucleotides per second compared to 370 nucleotides per second). The RecB-deficient complex translocated

at nearly wild type velocity (Taylor and Smith, 2003). These results established that RecD made a greater contribution to the speed of the complex. A short DNA loop structure was observed emanating from the complex; it was postulated that the loop was formed by the fast RecD motor unwinding the duplex while the slow RecB motor would move the ss of the loop structure through to produce a short tail (Taylor and Smith, 2003). Dillingham *et al*., (2003) looked at motor deficient complexes and came to similar conclusions; motor deficient complexes, whether in RecB or RecD still displayed helicase activity, but complexes containing both RecB and RecD defective motors showed no helicase activity. RecB motor deficient complexes translocation rates were 30% of wild type conversely, RecD motor deficient enzymes were 50% of wild type translocation speeds (Dillingham *et al.,* 2005). These data indicate that RecD is the faster of the two motors.

1.2.3 Nuclease (pre chi)

Before the RecBCD complex encounters chi it degrades both strands of the DNA helix. In conditions resembling the cellular environment where there is an excess of Mg^{2+}, RecBCD displays high exonuclease activity, preferentially degrading the 3′ strand (Dixon and Kowalczykowski, 1993) The RecB subunit contains the nuclease domain (Singleton *et al.,* 2004). Yu *et al*., (1998) confirmed that a truncated RecB protein, consisting of only the N-terminal domain, had no nuclease activity but had helicase activity. Conversely, Sun *et al.,* (2006) showed that a truncated RecB protein consisting of the 30 kDa C-terminal domain, had nuclease activity; interestingly the truncated protein could also act as an endonuclease. While the RecB nuclease domain has endonucleolytic properties, the RecBCD complex essentially acts as an exonuclease due to the physical restrictions provided by the loading constraints of the two linked motors of RecBCD. Thus the conformational restraints of the complex will affect the physical actions of the complex, restricting DNA access to RecB so that the complex acts as an exonuclease.

1.2.4 Chi Recognition

Chi (chi- crossover hotspot instigator (Lam *et al.*, 1974) is a DNA sequence that controls DNA degradation and induces RecA loading. The *E. coli* chi site is an 8 base pair sequence rich in G residues. It occurs 1008 times in the genome, averaging every 4.5kb; this is more than would be expected by chance. 75% of chi sites are orientated in the direction of replication with 98% located within coding regions (Blattner *et al.*, 1997; Arawkawa *et al*., 2007). It would seem probable that maintaining chi in the genome at this high frequency is advantageous.

The crystal structure of RecBCD showed a putative ssDNA binding site in RecC that might provide recognition of chi (Singleton *et al.*, 2004). Complexes that have mutations in RecC allow recognition of sequences that are similar to chi (Handa *et al*., 1997; Arnold *et al*., 2000). These mutations map to an area of RecC where the 3′ strand passes through the complex (Singelton *et al*., 2004).

Single molecule studies, following the displacement of the fluorescent dye YOYO-1 from DNA, allowed the translocation of RecBCD to be visualised (Bianco *et al*., 2001; Spies *et al.*, 2003). Upon encountering a chi site in the appropriate orientation (5′GCTGGTGG3′) on the 3′ strand, the complex temporarily pauses on average for 4.6 seconds before resuming translocation (Spies *et al.*, 2003). Tagging of RecD with a 40nm streptavidin-coated fluorescent bead has also allowed visualisation of the translocating complex, and has shown the complex pausing at chi for an average of 4.9 seconds (Handa *et al*., 2005). These data indicate that RecC recognises the chi sequence, and then signals to RecD, to initiate a pause (Singleton *et al*., 2004). The pause would allow alterations in activity to be instigated, so that RecB cuts at chi, then changes exonuclease activity to the 5′ strand and loads RecA onto the 3′ strand.

1.2.5 Chi modified helicase

After interacting with chi the translocation velocity of the RecBCD complex is reduced. Handa *et al*., (2005) fluorescently tagged the RecD subunit of the RecBCD complex, and quantified the translocation velocity pre and post chi. Prior to chi, the average speed was 960 ± 470bp/s, after chi the average speed was reduced 2 fold to 450 ± 230bp/s (Handa *et al.*, 2005). This was in keeping with the single molecule fluorescent dye displacement studies of Spies *et al.*, (2003) which showed that after the chi instigated pause RecBCD translocation resumed but at a reduced speed of roughly half that of its value pre-chi. The reduction in speed was initially hypothesised to be caused by ejection of the faster RecD motor (Spies *et al.*, 2003). The Handa *et al*., (2005) study, however, used fluorescently tagged RecD within the RecBCD complex, disproved the RecD ejection hypothesis; after chi the RecD tagged complex could still be seen translocating along DNA. Spies *et al.,* (2007) further studied the contribution of RecD to chi interaction; a RecBCD complex was made which contained a RecD subunit with a lysine to glutamine substitution in the Walker A motif. This complex paused at chi but did not show alteration of translocation rate after chi. The translocation rate before and after chi was comparable to RecBCD wild type translocation rate after chi modification, implying that RecB is the only active motor after chi recognition in wild type complexes (Spies *et al.*, 2007).

1.2.6 Chi modified Nuclease

The presence of a chi sequence confers protective properties to ds DNA by preventing degradation by RecBCD. In wild type cells, rolling circle plasmids are degraded as they replicate because they produce a free end that is open to degradation by RecBCD. Rolling circle plasmids replicate to produce high molecular weight multimers (HMW) when RecBCD is absent, or if chi is present in the correct orientation on the plasmid. The more chi sites that there are present on DNA the

greater the protection (Dabert *et al.,* 1992). Nuclease activity on the 3′ strand is 500 fold lower after chi (Dixon and Kowalczykowski, 1993) while activity on the 5′ strand is upregulated (Anderson & Kowalczykowski, 1997a). This produces the 3′ overhang that terminates at chi.

Single amino acid changes in the conserved helicase motif VI of RecB resulted in *recB* mutants that were shown to have exonuclease activity and resistance to infecting phage comparable to wild type cells, but the mutants did not cut at chi and instead cut at a novel position that was relative to the length of the DNA substrate. From this data it was proposed that as the mutant RecB motor moved more slowly than the RecD motor, that DNA cuts were initiated when RecD slowed upon reaching the end of the DNA substrate. When altered enzymes entered either end of one DNA duplex, the DNA was always cut in the middle. These results indicated that cleavage occurs when the complex is prevented from translocating. A mutation in the RecD motor that produced a slower translocation velocity for RecBCD showed that RecB would reach the end of the DNA first and consequently no novel cuts would be made, the complex also regained chi recognition but did not load RecA. This led to the proposal that RecC recognises chi, RecC then signals to RecD to attenuate its activity and RecD signals RecB to cut DNA and load RecA (Amundsen *et al.,* 2007).

1.2.7 RecA loading

RecBCD is active in the loading of RecA protein onto the 3′ strand generated by the exonuclease activity (Anderson & Kowalczykowski, 1997b). RecBC acts as a DNA helicase with no nuclease activity, but it constitutively loads RecA (Churchill *et al.,* 1999). RecBC can unwind duplex and facilitate the homologous pairing seen in strand invasion (coupled pairing reactions) by loading RecA (Churchill and Kowalczykowski, 2000). To determine where the RecA loading function was contained in RecBC, complexes containing only the C-terminal nuclease domain of RecB (RecBcterC) were made for use in coupled pairing assays (pairing of RecA coated 3′ strand with homologous DNA during synapsis). RecBcterC did not load RecA and therefore no coupled pairing reactions were observed. Compared to

RecBC, the RecBcterC complex had over a 7 fold reduction in the production of coupled pairing reactions, such that there was little or no activity (Churchill and Kowalczykowski, 2000). This showed that the C-terminal domain of RecB was important for RecA loading. In keeping with these results, other experiments showed that RecA protein co-eluted with a RecB truncated protein that consisted of only the C-terminal nuclease domain (Spies and Kowalczykowski, 2006).

1.2.8 Processivity and Velocity

Optical trapping of a fluorescently labelled DNA duplex allowed the measurement of the velocity and distance transversed by translocating RecBCD enzymes. The maximum velocity observed was 972 ± 172 base pairs per second at 37^{o}C, (Bianco *et al.,* 2001). The high translocation speeds can be attributed to the combined action of the two motors. RecD and RecB both provide helicase activity, but it is the combination of opposite polarities of the proteins (RecD having a preference for the 5′ strand and RecB having preference for the 3′ strand) that allows the protein complex to translocate at great speed (Dillingham *et al.,* 2003; Dillingham *et al.,* 2005; Stukalin *et al.,* 2005). Two motors can increase the forward velocity more than their combined separate speeds. The physical linking of the motors enhances the forward velocity of the complex by preventing any residual backward motion on the DNA, thus ensuring maximum forward speed (Stukalin *et al.,* 2005). RecC contributes to forward motion by enhancing the shearing of ds end; the crystal structure of RecBCD bound to DNA showed that RecC facilitates unwinding of the DNA duplex into ss DNA by physically splitting DNA over a pin-like structure (Figure 1.2.8.1; Singleton *et al.,* 2004). The structure showed that 4bp of the DNA strand had been melted with no ATP required for the unwinding of ds DNA (Singleton *et al.,* 2004).

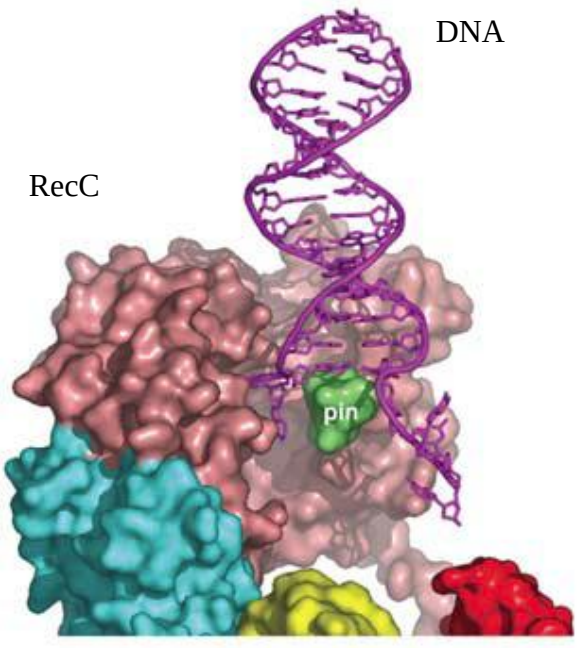


Figure 1.2.8.1: Splitting of DNA helix by RecC

Representation of the crystal structure of the RecC subunit complexed with ds DNA. The helix is split across the pin-like structure and funnelled into two channels in RecC. Taken from Singleton *et al*., 2004.

The high processivity ensures that RecBCD does not disassociate before a chi site is encountered. Phillips *et al*., (1997) compared the amount of DNA unwound by RecB, RecBC and RecBCD; RecBCD unwound the largest percentage, RecB the least. Interestingly RecBC showed an intermediate value; possibly because RecC could contribute to processivity. The crystal structure of RecBCD showed RecC funnelling the two DNA strands through its structure acting as a clamp and preventing dissociation and therefore contributing to the processivity. The structure also showed that RecBCD covers 16 bases of the 5′ strand and 13 bases of the 3′ strand allowing it to pass over ss gaps of 23 bases (Singleton *et al.*, 2004). For the complex to disassociate from the DNA, both motors would have to disassociate at the same time. Processivity is also reduced by mutations in the ATP binding motifs of either RecD (6 fold) or RecB (25fold) (Dillingham *et al.*, 2005). This may be because disassociation may be more likely if one of the helicase units is not translocating on the DNA. It is also interesting to note that RecB appears to have a more important role in processivity; RecB is the slower motor and contains an arm like structure that

protrudes and makes contact with the DNA contributing to association (Singleton *et al.,* 2004). The high velocity and ability to maintain contact with DNA would also allow RecBCD to remove any proteins bound to DNA that may impede movement along the duplex. This is demonstrated by the ability of RecBCD to displace nucleosomes from DNA (Eggleston *et al*., 1995). RecBCD requires high speeds and processivity to protect the cell by quickly degrading any incoming foreign DNA, such as phage, and it must process DNA DSBs quickly to facilitate DNA replication.

1.3 Pre-synaptic processing –AddAB/RexAB

The *E. coli* system has been accepted as the paradigm for recombinational repair, but as more bacterial genomes are sequenced, it is apparent there are different types of recombinational machinery. Heterodimeric protein complexes containing one helicase and two nuclease domains have been found to mediate DSB repair in many other bacteria; the prototypes are AddAB (ATP-dependant deoxyribonuclease) in *B. subtilis* and RexAB (recombination exonuclease) in *Lactoccocus lactis* (Kooistra *et al*., 1997; Quiberoni *et al.,* 2001). More recently another type of recombinational machinery has been found; the AdnAB (ATP-dependant nuclease) in Mycobacteria (*M. smegmatis* and *M. tuberculosis*). The mycobacteria have RecBCD but also AdnAB for repair of DSBs. AdnAB appears to be unique compared to the other recombinational machinery; it is a two component system but differs from AddAB/RexAB as each subunit appears to contain an active helicase and nuclease (Sinha *et al.,* 2009). The AddAB/RexAB systems have been more fully studied than the AdnAB proteins which has a limited distribution, therefore AddAB/RexAB will be the focus of discussion in reference to two protein systems.

The genes encoding AddAB/RexAB appear to form an operon; usually with *addB/rexB* followed by *addA/rexA* (Cromie, 2009). The AddAB/RexAB systems show little homology to RecBCD, yet they provide the same pre-synaptic processing functions (Figure 1.3.1). Though the systems have the same function they differ in physical characteristics. RecBCD is a heterotrimeric complex whereas the

AddAB/RexAB/AdnAB is a heterodimeric complex. RecBCD has two helicases whereas AddAB/RexAB has one. RecBCD has one nuclease but AddAB/RexAB has two. The only sequence similarity can be found in the helicase (SF1) and nuclease motifs (nuclease motif III). AddAB/RexAB preferentially binds double-stranded DNA ends and progressively unwinds double-strand DNA to produce single-strands. Both complexes react to a chi sequence (though each chi sequence is specific to each organism and their respective recombination system), whereby degradation of one strand (the 3′-5′) is attenuated. Thus the final product is a ss 3′ overhang; the substrate required for RecA-mediated strand invasion.

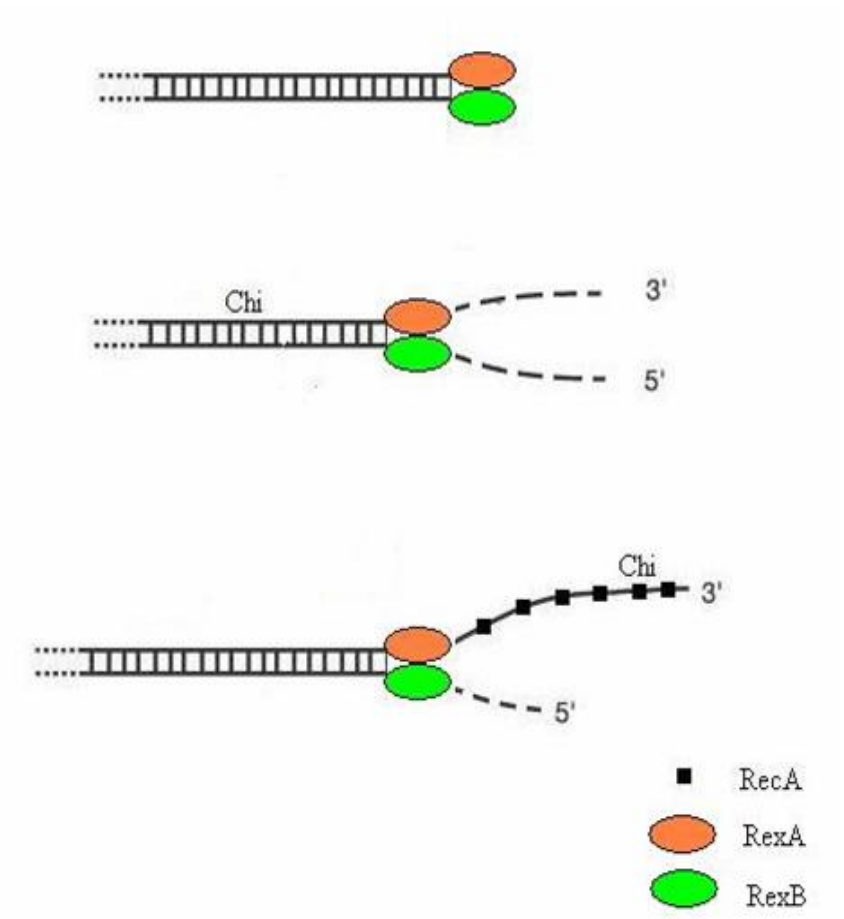


Figure 1.3.1: Pre-synaptic processing by RexAB

RexAB attaches to double-strand ends. The helicase activity unwinds the DNA duplex while the exonuclease activities degrade the strands. On encountering a chi site, activity is attenuated and the 3′ strand is no longer degraded. The strand is coated in RecA and strand invasion can now proceed with resolution of HR following. Adapted from Spies *et al.*, 2003.

1.3.1 Pre-synaptic processing –AddAB of *B. subtilis*

B. subtilis is a non-pathogenic, Gram-positive aerobe of the phylum Firmicutes, and is usually found in soil. The *B. subtilis* pre-synaptic processing proteins consist of AddA and AddB (141 kDa and 135 kDa) (Kooistra and Venema., 1991). Eight small regions (between 7 and 17 amino acids) in the AddA protein show similarity to *E. coli* RecB: though overall identities between the proteins are low, identities within these regions reached up to 70% (Kooistra and Venema., 1991). AddA has homology to the *E. coli* RecB SF1 helicase motifs and the nuclease motif III. *B. subtilis* AddB only shares homology to the *E. coli* RecB nuclease motif III. Haejima *et al.,* (1996a) highlighted 9 regions in the *B. subtilis* AddA protein with similarity to *E. coli* RecB; regions 1-2 were considered to be an ATP binding motif, regions 3-7 were proposed to be the helicase regions, regions 8-9 had similarity to nuclease domains.

Complementation studies were used to confirm the involvement of AddAB in DSB repair. The *B. subtilis addAB* genes were expressed in an *E. coli* Δ*recBCD* strain; the cell supernatants were found to have higher exonuclease activity in the presence of ATP compared to supernatant from an *E. coli* strain that lacked a recombination system (Kooistra and Venema., 1991). Expression of the *B. subtilis* AddAB system in *E. coli recBCD* null strains also alleviated the UV sensitivity and rescued the cell viability defect.

Haejema *et al.,* (1996a) investigated the proposed ATP binding sites of AddAB by introducing point mutations to produce two altered complexes, $AddA^{K36G}B$ and $AddAB^{K14G}$. The $AddA^{K36G}B$ complex was shown to be deficient for helicase and exonuclease activity, whereas the $AddAB^{K14G}$ complex retained helicase and exonuclease activity (Haejema *et al.,* 1996a). This showed that AddA is the helicase of the complex, and while both proteins have exonuclease activity, this demonstrated that helicase activity is required for nuclease function. Ca^{2+} was also shown to inhibit exonuclease function but not helicase function (Kooistra *et al.,* 1997). Deletion of the

C-terminal region (regions 8-9) of the AddA subunit produced the $AddA^{T}B$ complex that had greatly reduced nuclease activity but retained helicase activity. Interestingly even though there is reduced nuclease activity, the $AddA^{T}B$ complex can complement a recombination deficient *E. coli* strain and a recombination deficient *B. subtilis* strain (Haejema *et al.,* 1996b). This may be due to higher expression levels of $AddA^{T}B$ compensating for reduced activity, or perhaps the nuclease activity of the AddB subunit is sufficient.

The roles of each subunit within the complex were further elucidated by inactivating the proposed nuclease and helicase functions. Amino acid substitutions in the helicase (ATP binding motif) or the nuclease (mutation of the aspartate that coordinates with magnesium) did not affect binding to blunt DNA ends; binding was still strong at K_d = 0.2nM (Yeeles and Dillingham, 2007). AddAB requires ATP for nuclease and helicase activities; A helicase mutant containing a K36A mutation in the conserved lysine of the Walker A motif of the AddA subunit could not unwind DNA and therefore could not degrade DNA, implying that the helicase function is within the AddA subunit (Yeeles and Dillingham, 2007).

Wild type AddAB required high levels of Mg^{2+} (1mM-2mM) for nuclease activity but not for helicase activity (Yeeles and Dillingham, 2007). An aspartate residue has been shown to allow the Mg^{2+} co-factor to coordinate at the nuclease active site of *E. coli* RecB (Singleton *et al.,* 2004). Altered complexes that contained a mutation in the LDYK sequence of the nuclease motif III (aspartate to alanine) were constructed; $AddA^{N}B$, $AddAB^{N}$ and $AddA^{N}B^{N}$. All the altered complexes displayed wild type helicase activity, but nuclease activity was severely reduced in the $AddA^{N}B^{N}$ complex. Complexes with either altered AddA or AddB nuclease ($AddA^{N}B$, $AddAB^{N}$) showed an intermediate activity compared to the $AddA^{N}B^{N}$ and wild type complexes (Yeeles and Dillingham, 2007).

An iron sulphur cluster has also been identified within the AddB nuclease domain, consisting of 4 cysteines. The first cysteine is 300aa upstream of the N terminal of the nuclease domain and the other 3 cysteines are close to the C terminal of the

nuclease domain. Mutation of any of the cysteines within the iron sulphur cluster causes loss of AddAB binding to DNA ends, possibly by disrupting the structure of the protein (Yeeles *et al.,* 2009).

B. subtilis AddAB reacts to the chi sequence 5′ AGCGG 3′, which is over represented in the genome, occurring every 884bp (Chedin *et al.,* 1998). Chedin *et al.,* (2006) showed that DNA without chi sequences are nearly completely degraded (75%) by AddAB, but if they contain chi sequences there is protection from degradation. If one chi site is present about 50% of DNA is degraded if there are three chi sites then only 25% of DNA is degraded, showing that chi recognition is not 100% efficient and a single site has only a 25% chance of being recognised. Chedin *et al.,* (2006) showed that AddAB binds non-covalently to the chi site, this led them to propose that a ssDNA loop structure, similar to that seen with RecBCD, may be formed. Nuclease mutants have been used to determine the roles of the subunits in producing chi specific fragments (Yeeles and Dillingham, 2007). Wild type AddAB and the $AddAB^{N}$ complex produced chi specific fragments, whereas the $AddA^{N}B$ complex did not produce chi fragments; this showed that the AddA nuclease is required for producing chi specific cuts. AddB degrades 5′ - 3′ and AddA degrades 3′ - 5′, each nuclease is specific to strand orientation; at chi there is no switch in polarity or activity of the AddA nuclease indicating that the AddB subunit is down regulated (Yeeles and Dillingham, 2007). It was shown that the nuclease domains are not involved in the recognition of chi as all mutants were still able to recognise and form tight complexes with chi sequences (Yeeles and Dillingham, 2007).

1.3.2 Pre-synaptic processing – RexAB of *Lactococcus lactis*

Lactococcus lactis is a non-pathogenic, fermenting, Gram-positive, low GC bacterium with variable natural habitats. The *rexAB* genes were identified by screening for recombination deficient mutants; out of 30,000 only 3 were found that were defective in recombination. The three mutations mapped to two ORFs, the

chromosomal architecture of which was similar to that seen with *addAB* of *B. subtilis*; the overlapping genes were organised together, with *rexB* downstream and *rexA* upstream, indicating they may be co-transcribed. The genes produced proteins named RexA and RexB of 1,073 and 1,099 amino acids that share 23 and 34% identity with *B. subtilis* AddA and AddB, respectively. When compared to *E. coli* RecBCD, RexA has less than 20% homology to *E. coli* RecB and RexB has no significant global homology to any of the subunits of *E. coli* RecBCD (El Karoui *et al.,* 1998).

To further study the *L. lactis rexAB* genes, they were introduced to an *E. coli* strain lacking RecBCD to see if the recombination deficient phenotype could be rescued. The UV sensitive phenotype of the strain was partially alleviated and the conjugal recombination efficiency was markedly improved compared to the parental *E. coli* Δ*recBCD* strain (El Karoui *et al.,* 1998; Quiberoni *et al.,* 2001). The nuclease motif was identified in RexA as consensus G-i-i-D-x(12)-D-Y-K-t-d, (amino acids in lower case show some degeneracy) and RexB as G-r-i-D-R-i-D-x(9–12)-v-D-Y-K-S-s, both have homology to the nuclease motif of *E. coli* RecB; this implied that RexAB had two active nucleases. Mutations were made in both nuclease sites ($RexAB^{Nuc}$ and $RexA^{Nuc}B$) that resulted in reduced nuclease activities of the complexes, and when expressed in an *E. coli* Δ*recBCD* background resistance to UV was extremely low (Quiberoni *et al.,* 2001). The *L. lactis* chi sequence was identified as 5′-GCGCGTG-3′ by the presence of a sequence which prevented degradation of HMW products produced by rolling circle replication (Biswas *et al.,* 1995; El Karoui *et al.,* 1998)

1.3.3 AddAB in Gram negatives

Initially it was stated that the AddAB/RexAB systems were confined to Gram-positive bacteria (Chedin and Kowalczykowski, 2002). Genome sequencing and bioinformatics analysis however have shown that the AddAB system may be more widely distributed. (Amundsen *et al.,* 2008; Rocha *et al.,* 2005; Zuniga-Castillo *et al.,* 2004) Of the many species identified as containing AddAB homolouges, to date

only four genetic studies have been published (Amundsen *et al.*, 2008; Zuniga-Castillo *et al.*, 2004; Mertons *et al.*, 2008; Reuter *et al*., 2010).

Helicobacter pylori is a Gram-negative bacterium that causes infection resulting in inflammation of the human stomach. The *H. pylori addAB* genes were introduced into an *E. coli ΔrecBCD* strain and were shown to confer nuclease and helicase activities. Deletion mutants lacking *addA* or *addB* were sensitive to DNA damaging agents, such as UV and mitomycin C. *H. pylori addA* or *addB* mutants that had no ATP-dependant nuclease or helicase activity could be complemented by introduction of the genes on a plasmid (Amundesen *et al.*, 2008). *Coxiella burnetii* is a Gram-negative, obligate pathogenic bacterium that causes Query fever. The *addAB* genes in this bacterium show the same organisation as other organisms, with *addB* upstream of *addA*. The *C. burnetii* AddA sequence appears to contain a nuclease domain and a helicase domain. The *addAB* genes were introduced into an *E. coli recBCD* deletion strain and shown to restore resistance to treatment with DNA damaging agents such as UV light and mitomycin C (Mertons *et al.*, 2008). Consistant with other bacteria, *Rhizobium etli addA* mutants also showed increased sensitivity to mitomycin C, UV radiation and nalidixic acid, while *addAB* mutants were extremely sensitive to nalidixic acid (Zuniga-Castillo *et al.*, 2004). Strains lacking *addAB* could be rescued, by the introduction of plasmid borne copies of *addAB*. (Zuniga-Castillo *et al.*, 2004)

1.4 *Bacteroides fragilis*

1.4.1 Background

The *B. fragilis* strain (NCTC9343) used for this study is a clinical isolate from the 1950s. The NCTC9343 genome (5,241,700 bp) consists of the chromosome (5,205,140 bp) and the pBF9343 plasmid (36,560 bp). The genome contains 4,322 CDSs; 48 CDSs are on the plasmid and 4274 CDSs are chromosomal, of which 70

are pseudogenes. The chromosome has an average GC content of 43.19% whereas the pBF9343 plasmid is 32.24% (Cerdeno-Tarraga *et al*., 2005). The GC content places NCTC9343 within the low GC group of bacteria. The genus *Bacteroides* belongs to the phylum *Bacteroidetes* and is part of the *Bacteroidetes/Chlorobi* group that contains *Cytophagia, Flavobacteria, Sphingobacteria, Bacteroidia, Chlorobia* and *Bacteroidetes*. The phylum diverged early in the evolutionary lineage of bacteria and though *Bacteroides* are Gram-negative organisms, they are not closely related to the enteric Gram-negatives such as *E. coli* (Figure 1.4.1.1).

A. B.

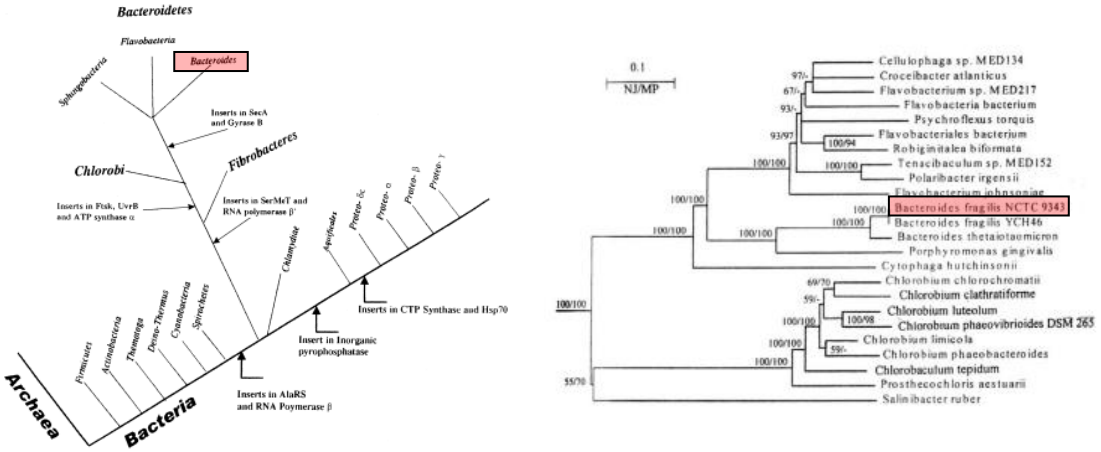


Figure 1.4.1.1: Branching order of main bacterial groups and evolutionary comparison of *Bacteroidetes/Chlorobi* group

A. Branching order of main bacterial groups based on signature protein sequence comparisons. To the right and above, arrows indicate where proteins signatures were introduced. B. Neighbour joining tree based on 12 signature proteins. The figures demonstrate that there is a large evolutionary distance between the Bacteroides group and the Gram-negative enteric group which includes *E. coli*. It is also interesting to note that there is large phylogenetic divergence within the Bacteroides/Flavobacterium group. Taken from Gupta, 2004; Gupta & Lorenzini, 2007.

The *Bacteroides* are rod shaped cells, saccharolytic, bile tolerant, non-sporing, non-motile, Gram-negative, obligate anaerobes that inhabit the gastrointestinal (GI) tracts of mammals. *Bacteroides* are obligate anaerobes but can survive exposure to air for 24 - 72 hours. They typically produce neuraminidase, DNase, hyaluronidase, gelatinase, fibrinoylsin, superoxide dismutase and heparinase. The major

carbohydrate fermentation end products are acetate, succinate, and propionate, and they have a large number of genes involved in the breakdown of plant polysaccharides that are indigestible to the human host (Reviewed by Salyers, 1984; Patrick, 2001; Wexler, 2007).

1.4.2 *Bacteroides fragilis* as a commensal organism

The normal microbiota of the GI tract are advantageous to the human host. The presence of *Bacteroides spp.* has been shown to be crucial for normal development and function of the host's gastrointestinal tract. As a commensal, *Bacteroides spp.* can benefit the host by providing nutrients, excluding invading pathogens and encouraging proper physical development of the gut and immune system.

Of all the areas colonised by bacteria in the human body, the gastrointestinal tract contains the largest and most diverse population of bacteria. The gut has the highest bacterial cell density of any ecosystem (500-1000 different species) with the majority of the bacteria in the gut being anaerobes (>99%), while the wet mass of human faeces is 35% bacteria, an indication of the high numbers present in the gut (van Houte and Gibbons, 1966; Moore and Holdeman, 1974; Xu & Gordon, 2003). Two divisions that dominate the gut are the *Bacteroidetes/Chlorobi* and the *Firmicutes*, with *Bacteroides* species accounting for between 26.5 to 30% of all organisms isolated from faecal samples (Moore and Holdeman, 1974; Salyers, 1984; Eckberg *et al*., 2005). The *Bacteroides* group was predominantly isolated from faecal samples of subjects with conventional western diets, with 10^{11} organisms per gram of dry weight faeces (Finegold, 1977). The large numbers of *Bacteroides* present may be due efficient competition for resources against other bacteria present in the normal microbiota. The sheer numbers of *Bacteroides* species can protect the host from pathogen colonisation; *B. fragilis* may out-compete other species for nutrients and wall attachment sites, while producing substances that are hazardous such as volatile fatty acids and causing the release of free bile acids.

As commensals *Bacteroides* species provide calories for the host; plant polysaccharides which are indigestible to the human host, such as hemicellulose, cellulose, pectin and some starches are degraded by the anaerobic bacteria of the gut. The microbiota process carbohydrates primarily into short chain fatty acids (SCFAs) that can be metabolised by the host; Namavar *et al*., (1989) found *B. fragilis* to associate with the mucosal lining – this proximity would ensure that the host received the SCFAs. Wostmann *et al.*, (1983) observed that germ free (GF) rats have to consume 30% more calories than conventionally raised rats (CONV-R) in order to maintain their body weight. CONV-R rats have 42% more body fat than GF rats even though the GF rats were consuming more calories, while GF rats that were colonised with faecal microbiota increased their body weight to that of CONV-R rats. When faecal samples from the CONV-R rats and the GF rats that were colonised with faecal microbiota were assayed for bacteria, they were found to contain similar profiles, with *Bacteroides* being one of the most well represented (Backhed *et al.*, 2004). *Bacteroides spp*. are involved in the breakdown of many plant polysaccharides that are undigested by the host: *B. thetaiotaomicron* has the largest number of putative glycosyl hydrolases of any bacterium, indicating that the *Bacteroides* may have a large role in supplying calories to the host (Xu & Gordon., 2003). Little is known about carbohydrate metabolism in *B. fragilis* but analysis of the *B. fragilis* genome suggests a large portion is devoted to carbohydrate catabolism and metabolism. Recently Wexler *et al*., (2002) identified a *B. fragilis* outer membrane protein (Omp121) with homology to *B. thetaiotaomicron* SusC. Xu & Gordon., (2003) noted that another *B. fragilis* Omp (Omp71) had 15 homologs in *B. thetaiotaomicron*, with 12 of those clustered around *B. thetaiotaomicron SusC*: they proposed that Omp121 and Omp71 could be members of carbohydrate utilisation pathway.

Not only do bacteria provide nutrients from the breakdown of indigestible polysaccharides, but they also provide immune stimulating molecules. *B. fragilis* processes dietary lignans such as Secoisolariciresinoldiglucoside (SDG) making them available for the host (Clavel *et al.*, 2006). SDG is an antioxidant and has been implicated in diabetes type 1 and type 2, and in the prevention of breast cancer by

inhibiting or slowing the growth of tumours, (Chen at al., 2003; Prasad, 2000; Prasad, 2001). The *Bacteroides* group, specifically *B. fragilis*, are important for the correct development and function of the immune system in the gastrointestinal tract. In rabbits that had a germ free appendix, the combination of *B. fragilis* and *B. subtilis* induced gut-associated lymphoid tissue (GALT) development and somatic diversification of the pre-immune system. *B. fragilis* was taken up by M cells but did not illicit an immune response, instead *B. fragilis* enhanced *B. subtilis* uptake, allowing *B. subtilis* to promote an immune response (Rhee *et al.,* 2004). *B. fragilis* polysaccharide A (PSA) also induces lymphoid organogenesis in previously germ free mice, and encourages T cell production. The presence of PSA regulates the ratio of immune effectors Th1/Th2 cell imbalances seen in germ free mice, which is essential for appropriate function of the immune response (Mazmanian *et al.,* 2005). Interestingly PSA is the only carbohydrate that is internalised by antigen presenting cells (APC) and displayed by the major histocompatability complex II (MHCII) to T cells (Cobb *et al.,* 2004). Germ free mice have incomplete vascularisation of the gut, but *B. thetaiotaomicron* can stimulate angiogenesis in these mice (Stappenbeck *et al.,* 2002). Germ free mice were inoculated with either *B. thetaiotaomicron*, *E. coli*, *B. infantis* or conventional mixed microbiota; two controls of germ free or conventionally raised mice were used. The *B. thetaiotaomicron* inoculated group had the largest increase in mRNAs studied; with colipase, angiogenin-3, and sprr2a (small proline rich protein-2) showing the greatest increases (Hooper *et al.,* 2001). Colipase stimulates the production of pancreatic triglyceride lipase and pancreatic lipase-related protein-2, both of which are involved in lipase metabolism, while angiogenin-3 contributes to angiogenesis and Sprr2a is involved in preserving the physical structure of the mucosal barrier. Recently it has been shown that *B. fragilis* acts as an anti-inflammatory against colitis caused by *H. hepaticus*, by suppressing pro-inflammatory interleukin-17 production from intestinal immune cells (Mazmanian *et al.,* 2008).

1.4.3 Pathogenicity – disease and medical relevance of *Bacteroides fragilis*

Though many benefits are conferred to the host by the presence of *B. fragilis*, it is also an opportunistic pathogen. *B. thetaiotaomicron* is the most prevalent *Bacteroides* species in the gut, but it is not the major clinical isolate; *B. thetaiotaomicron* accounts for 15-29% of the microbiota and causes 13-17% of *Bacteroides* infections, while *B. fragilis* makes up only 4-13% of the gastrointestinal microbiota but causes 63-81% of *Bacteroides* infections (Patrick, 2001). *B. fragilis* is the main obligate anaerobe clinical isolate found in abscess, soft tissue infections and bacteraemia but can also be isolated from intra-abdominal, vaginal, pilonidal, perianal and brain abscesses. *B. fragilis* can cause a number of infections including meningitis, septic arthritis, pelvic inflammatory disease, peritonitis and is the most frequent cause of anaerobic septicaemia (Patrick, 2001: Patrick & Duerden, 2006) .If anaerobic septicaemia is left untreated the mortality rate is about 60% (Nobles, 1973). This can be greatly reduced by application of appropriate antibiotics, lowering the mortality rate to 19% (Redondo *et al.*, 1995).

B. fragilis infection occurs when it gains access to normally uncolonised sites within the body, usually following ruptures of the intestine wall. Most initial stages of infections involving *B. fragilis* are polymicrobial; during the acute phase of infection (initial 20 hours), the facultative anaerobes (generally *E. coli*) are predominant while the aerotolerance of *B. fragilis* allows the organism to persist during this stage of infection (Sund *et al.*, 2008). The aerobic bacteria reduce the oxygen present in the environment, providing favourable conditions for *B. fragilis* to proliferate and become the dominant organism. As the infection progresses an abscess is formed; if the abscess metastasizes, the bacterial infection can transfer to other parts of the body via release of bacteria into the bloodstream, potentially resulting in bacteraemia. The pathogenicity of *B. fragilis* is due to several factors; penetration/destruction of host tissues, avoidance of the host immune system, aerotolerance, antibiotic resistance and the ability to exploit the host as a source of nutrients.

Avoidance of the host immune system and alteration of the host immune response allows *B. fragilis* to persist and grow during infection. The *B. fragilis* capsule is implicated in avoidance of the host immune system, with phase variation of surface polysaccharides observed within *B. fragilis* cultures. Three *B. fragilis* capsular types have been observed: the large capsule (LC), small capsule (SC) and an electron-dense-layer (EDL) forming a micro capsule (MC) (Patrick *et al*., 1986). There are many gene clusters believed to be involved in production of extracellular polysaccharides (10 annotated from NCTC9343 genome, PSA-J). The antigenic variation of eight extracellular polysaccharides from NCTC9343 (PSA-H) is produced by invertible promoters (switch on/off) that are upstream of seven of the polysaccharide biosynthesis gene clusters (Cerdeno-Tarraga *et al*., 2005). The expression of different PSs would allow avoidance of the host immune system and invading phage. The *B. fragilis* capsule has also been implicated in resistance to complement mediated killing, phagocytic uptake and killing by polymorphonuclear leukocytes (PMN) (Simon *et al*., 1982; Reid & Patrick 1984). Rotstein *et al*., (1989) observed that *B. fragilis* supernatant impaired PMN uptake of *B. fragilis* and also *E. coli*; the supernatant reduced the microcidal action of PMNs superoxide dismutase. Rotstein *et al*., (1989) proposed that SCFAs in the *B. fragilis* supernatant could be responsible as Eftimiadi *et al*., (1987) had observed that SCFAs decrease lysozyme release and efficacy of lysozyme action as well as inhibiting the polymorphonuclear leukocyte chemotactic response. The synergy of the infection is demonstrated by the presence of *B. fragilis* having a protective effect on *E. coli*; protecting other bacteria such as aerobic *E. coli* would allow *B. fragilis* to benefit through the aerobic bacteria's reduction of the environment. During infection an abscess is formed, this protects the host by isolating the invading pathogen but at the same time protects bacteria within the abscess from the host immune system. The human immune response to *B. fragilis* NCTC9343 is activated in response to LPS via Toll-like receptors TLR2 (Erridge *et al.*, 2004). *B. fragilis* PSA is one of the molecules thought to induce abscess formation by stimulating interleukin-2 production (IL-2) and ultimately the T cell proliferation inflammatory response (Stingele *et al.*, 2004).

Importantly *B. fragilis* is the only bacterium that can act as the sole infecting organism in inducing formation of abscesses (Onkerdonk *et al.,* 1977).

Adhesion to host cells is an important virulence factor that aids an invading pathogen to colonise a host. *B. fragilis* can bind several types of host cell. *B. fragilis* can adhere to epithelial cells and cause hemagglutination (Hofstad., 1984; Oyston and Handley., 1991). Gibson *et al*., (1998) observed that *B. fragilis* NCTC9343 adhered to murine mesothelial cells (MMCs), more than other *Bacteroides* species, such as *B. distasonis* and *B. thetaiotaomicron*. Furthermore purified *B. fragilis* capsular polysaccharide complex (CPC) consisting of both PSA and PSB, bound to MMCs and peritoneal macrophages as did PSA on its own. Two *B. fragilis* strains were shown to have lectins that attached to glucosamine and galatosamine (Rogemond and Guinet, 1986). Ninety to ninety five percent of *B. fragilis* strains isolated from abscesses had the highest levels of efficient binding to polymorphonuclear neutrophils, epithelial cells and intestinal cells (Guzman *et al*., 1997).

The destruction of host cell tissues is an important virulence factor as it provides nutrients for the infecting bacteria. To use the host as a source of nutrients *B. fragilis* produces a variety proteins; hyaluronidase and chondriotin sulfatases (attacks the host extracellular matrix by hydrolysing heparin and chondroitin sulphate), collagenase (breaks down collagen), fibrinolysin (breaks down fibrin found in the abscess), DNase and sialidase (degrades neuramic acid of epithelial cells and intestinal mucosa). Sialidase is also involved in adhesion and hemaglutination and interestingly is active in more than 90% of *B. fragilis* strains (Nakano *et al*., 2006). *B. fragilis* can use haemoglobin-haptoglobin complexes as a source of iron and the synergy of infection is demonstrated by *B. fragilis* using *E. coli* haemoglobin protease (Hbp) to provide a haeme source. Hbp scavenges haemoglobin and releases haeme, with *B. fragilis* showing enhanced growth in the presence of Hbp and haeme. Interestingly *E. coli* strains isolated from abscesses were all positive for Hbp (Otto *et al* 2002).

B. fragilis also causes diarrhoeal diseases with the first cases of Enterotoxigenic *B. fragilis* (EBTF) being described in 1984 (Myers *et al.*, 1984). It has been estimated that EBTF is present in 6.5% of the population, with sewage samples showing 9.3% of the *B. fragilis* strains isolated as EBTF positive (Sears *et al.*, 1995; Shoop *et al.*, 1990). These strains produce the enterotoxin fragilysin, a zinc metalloprotease that disrupts intercellular adhesion and increases permeability of the epithelium barrier by cleaving the intercellular tight junction proteins such as ZO-1 and E-cadherin (Moncrief *et al.*, 1995; Obiso *et al.*, 1997). As fragilysin has been found in intestinal strains, it has been proposed that fragilysin could contribute to the pathogenesis of *B. fragilis* in extra-intestinal infections by inducing haemorrhaging, or by causing the host immune system to destroy its own cells (Obiso *et al.*, 1997).

The emergence of *B. fragilis* as a diarrhoeal causing organism has added to the clinical importance of this species. Lassman *et al.*, (2007) reported that the Mayo clinic, during the period 1993-2004, had a 74% increase in anaerobic bacteraemia cases, with the *B. fragilis* group being the major group isolated. It should also be considered that reporting of *B. fragilis* infection may be underestimated due to the difficulty in isolating and culturing these bacteria. Therefore *B. fragilis* is an important pathogen that can cause disease and fatalities, all the more worrying are the high rates of resistance to antibiotics within the group (discussed in section below).

1.4.4 Historical view of antibiotic resistance in the *Bacteroides* group

Bacteroides fragilis has historically been shown to acquire antibiotic resistance. The first case of penicillin resistance was seen in the 1960s, this was followed by tetracycline resistance in the 1970s, with more than 60% of strains resistant to tetracycline (Sutter *et al.*, 1972). The first clinically significant clindamycin resistant *B. fragilis* strain was isolated in 1976 (Slaki *et al.*, 1976). The first metronidazole resistant *B. fragilis* strain was documented in 1978, which appeared after a course of long term treatment (Ingham *et al.*, 1978). At present, most *B. fragilis* are resistant to

β-lactams and aminoglycosides, and are frequently resistant to tetracycline, clindamycin and colistin. Because of this, the main drug that is effective in treatment of *B. fragilis* infections is metronidazole.

1.4.5 Recent cases of metronidazole resistance in the *Bacteroides* group

Indiscriminate use of the drug may have led to the recent rise in cases of metronidazole resistant strains. Since the introduction of metronidazole in the late 1960s, it has been the drug of choice for anaerobic infections, in particular for *B. fragilis* infections. Metronidazole is used to treat protozoan diseases and anaerobic bacterial infections. It is cheap, has good penetration and has been used widely since the 1970s. The worldwide rate of resistance to metronidazole is low at less than 0.5% (Falagas and Siakavellas, 2000), but there has been an increase in resistance to metronidazole in the *Bacteroides* group (Fang *et al.,* 2002). Five percent of Taiwanese *B. fragilis* strains assayed were metronidazole resistant with MICs greater than 32mg/l (Teng *et al.,* 2002). Mexico has also seen rates of metronidazole resistance reach above 5%, while the USA detected rates of 2% in 1989 (Leanos *et al.,* 1989 cited by Garcia-Rodriguez *et al.,* 1995; Wasilauskas., 1989).

The importance of resistance is highlighted by a case of an appendix wound infection, caused by *B. fragilis*, that did not respond to metronidazole treatment. The *B. fragilis* isolated from the patient was found to have MICs greater than 32mg/L (Rotimi *et al.,* 1999 cited by Brazier *et al.,* 1999). Another case of metronidazole resistant *B. fragilis* was reported where a patient having an elective laparotomy developed peritonitis and a small abscess causing a bowel obstruction. Two metronidazole resistant *B. fragilis* strains were isolated: one from blood cultures and the other from a pus sample, both with MICs of 8mg/L (Turner *et al.,* 1995). Recently an incidence of bacteraemia due to *B. fragilis* was reported in which the strain had simultaneous resistance to metronidazole, β-lactams, β-lactamase inhibitors, carbapenems, macrolides, and tetracyclines. The patient was prophylacticly treated with metronidazole and cefuroxime, prior and post surgery (endoscopic retrograde

cholangeopancreatography), but the patient developed a retroperitoneal infection. Several antibiotics were used to treat the patient, but after 22 days of infection the patient died (Whareham *et al.*, 2005). If Metronidazole resistance within the *Bacteroides* group continues to grow, any lower abdominal surgery would pose a major threat to the life of the patient.

1.4.6 Metronidazole mode of action

Metronidazole (2-methyl-5-nitroimidazol-1-ethanol) belongs to the 5-nitroimidazole group of antibiotics; they have a heterocyclic structure composed of an imidazole based ring with nitro group (NO_2) at position five on the ring. They are pro-drugs that are inactive until reduced. The reduction of the drug is almost exclusive to anaerobic cells, so only organisms employing anaerobic respiration are susceptible to metronidazole. It is believed that the nitro group of metronidazole is chemically reduced producing a cytotoxic nitro-radical anion intermediate. The main target of the metronidazole intermediate is DNA, causing ss nicks which ultimately become DSBs, though it might also inactivate enzymes and disrupt membranes.

Metronidazole is believed to enter the cell by passive diffusion. When metronidazole is applied to aerobic cultures, equilibrium is reached between extra and intra cellular pro-drug concentrations. Conversely in anaerobic cultures there is a continuous flow of the pro-drug into the cell; possibly due to a favourable diffusion gradient caused by reduction of the pro-drug (Ings *et al.*, 1973; Tally *et al.*, 1978). In metronidazole sensitive *B. fragilis* strains there was a reduction in concentration of metronidazole in the extracellular medium (Diniz *et al.*, 2000).

There is a correlation between anaerobic respiration and activation of metronidazole. Enzymes involved in redox reactions such as pyruvate/ketoredutaseses, hydrogenases, ferredoxin and flavodoxin, are believed to contribute to the reduction of metronidazole. There is a high association with the presence of pyruvate flavodoxin/ferridoxin oxoredutase (POR) and susceptibility to metronidazole. This was observed when comparing resistant and sensitive strains of *B. fragilis* (Britz and

Wilkinson, 1979). Ferrodoxin or flavodoxin that is reduced by POR can donate an electron to the NO_2 group of metronidazole which is required for bactericidal action (Castelli *et al.,* 1997). This implies that reduction of the pro-drug is essential to its activity. Metronidazole can also be reduced and activated in poorly oxygenated environments, typically seen in abscesses or necrotic centres of tumours.

The activated nitro radical anion intermediate can donate electrons to DNA, causing ss breaks resulting in DSBs and ultimately the death of the organism (Simmms-Cendan, 1996). The nitro-radicals act by non-specific association with DNA. In susceptible members of the *Bacteroides* group, application of metronidazole causes a high level of chromosomal degradation immediately after treatment (Diniz *et al.,* 2000). DNA synthesis in *B. fragilis* susceptible strains stops after metronidazole uptake, but RNA and protein synthesis continue for at least 1 hour, implying that DNA damage is the main action of metronidazole (Sigeti *et al.,* 1983 cited by Diniz *et al.,* 2000).

The amount of DNA damage is so large that the endogenous DNA repair systems cannot cope with the damage, thus leading to cell death. This is reflected by the viability of cultures after treatment with metronidazole, for example *Clostridium pasteurianium* cultures treated with metronidazole showed 0.1% viability compared to untreated. A similar reduction in viability was also measured in *Bacteroides* cultures (Church *et al.,* 1991; Diniz *et al.,* 2000). The importance of DNA damage is shown by the observation that recombination deficient mutants of *E. coli* are more sensitive to metronidazole compared to wild type strains (Jackson *et al.,* 1984) *B. thetaiotaomicron recA* mutants were also more sensitive to metronidazole than wild type (Cooper *et al*., 1997). Over expression of genes involved in repair can confer resistance; a mutant *recA* gene from an *H. pylori* metronidazole strain could confer increased resistance to metronidazole in *E. coli* (Chang *et al*., 1997). Casanueva *et al.,* (2008) showed that a *B. fragilis* transcriptional regulator (*reg*) could improve the survival of metronidazole sensitive *E. coli* DNA repair mutants; though the mechanism of protection is unknown it was proposed that it may regulate DNA repair processes.

1.5 Aims

The main aim of this project was to identify the genes involved in the pre-synaptic processing of DNA DSBs in *B. fragilis* and characterise the biochemical functions of the proteins.

- Identification of possible genes involved in *B. fragilis* pre-synapsis: confirm/discount putative genes role in DSB repair: introduction of putative *B. fragilis* presynapsis genes into *E. coli recBCD* null strain. Determination of rescue of *E. coli recBCD* null recombination deficient phenotype through expression of *B. fragilis* proteins (Sections 3.2 and 3.3)

- Characterisation of *B. fragilis* recombination deficient strain: Production of a *B. fragilis* recombination deficient strain, determine if mutants displayed a DSB repair deficient phenotype (sensitivity to DNA damaging agents such as UV and metronidazole) (Section 3.4).

- Examine role of other genes implicated in DSB repair: aimed to delete other genes, for example *recF* and *sbcCD*, and characterise mutants (Section 3.4).

- Assign possible function to subunits/domains of pre-synapsis complex via *in silico* analysis (Section 3.5).

- Purification of *B. fragilis* pre-synapsis proteins (Section 4.2).

- Determine the biochemical activity of *B. fragilis* pre-synapsis proteins: establish the preferred DNA substrate and cofactor requirements (Section 4.3).

Chapter 2: Materials and Methods

2.1 Growth Media

L-Broth: 10g/l Difco Bacto-Tryptone, 5g/l Difco Bacto-yeast extract, 5g/l NaCl, 1l dH_2O, pH adjusted to 7.2 with NaOH

L-agar: L-Broth containing 15g/l Oxiod agar

BHI-S: 37g/l Difco Bacto Brain Heart Infusion, 10ml/l 5% L-Cysteine, 10ml/l 10% NaHCO3, 10ml/l Haemin/Menadione, 1l dH_2O

BHI-S agar: 1.5% w/v Difco Bacto Agar 37g/l Difco Bacto Brain Heart Infusion 10ml/l 5% L-Cysteine, 10ml/l 10% NaHCO3, 10ml/l Haemin/Menadione, 1l dH_2O

Supplements

5% L-Cysteine: 5% w/v L-Cysteine, made up to volume with dH_20 and sterilized with 0.22 micron filter.

10% $NaHCO_3$: 10% w/v $NaHCO_3$, made up to volume with dH_20 and autoclaved

Haemin/Menadione: 50mg Haemin, 1.74g K_2HPO_4, 0.4g NaOH, 99ml dH_2O, autoclaved before addition of 1ml of 5mg/ml Menadione (made up to volume with alcohol).

Defined Minimal Medium

Prepared as described by Van Tassell and Wilkins 1978; solutions 1-6 were mixed in the following quantities: 340ml Solution 1, 500ml Solution 2, 100ml Solution 3, 40ml Solution 4, 10ml Solution 5, 10ml Solution 6. Bacteriological Agar was added to a final volume of 1.4%w/v. Plates were incubated under anaerobic conditions till reduced; indicated by colour change from pink to straw yellow.

Solution 1: 2g $(NH_4)_2SO_4$, 0.5g Sodium citrate, 5µl of 2.5mg/100ml Vitamin B12 stock, 7g KH_2PO_4, 8g K_2HPO_4, 10mg $MnCl_2.4H_2O$, 20mg $MgCl.6H_2O$, 340ml dH_2O, autoclaved

Solution 2: 0.3mg $FeCl_3.6H_2O$, 30mg $CaCl_2.2H_2O$, 1mg Resazurin, 500ml dH_2O, autoclaved

Solution 3: 10g Fucose, 100ml dH_2O, autoclaved

Solution 4: 10% $NaHCO_3$ (as previously described)

Solution 5: 5% L-Cysteine (as previously described)

Solution 6: Haemin/Menadione (as previously described)

Anaerobic Growth Conditions: Anaerobic cultures were grown in a Don Whitley Scientific (UK) MiniMacs anaerobic work station at 37°C with an anaerobic gas mix (10% Hydrogen 10% Carbon Dioxide and 80% Nitrogen).

Culture Growth: Monitored by measuring optical density at 600nm (OD_{600}) with a Hitachi U-2000 spectrophotometer (Hitachi Instruments, Japan).

2.2 Antibiotics and Reagents

Table 2.1 List of Antibiotics

Antibiotic	Stock solution	Working concentration	Preparation
Ampicillin	100mg ml^{-1}	100µg ml^{-1}	In sterile H_20
Chloramphenicol	34 or 20mg ml^{-1}	20µg ml^{-1}	In ethanol
Eryrthromycin	10mg ml^{-1}	10µg ml^{-1}	In ethanol
Kanamycin sulphate	50mg ml^{-1}	30µg ml^{-1}	In sterile H_20
Naladixic acid	50mg ml^{-1}	30µg ml^{-1}	In 50mM NaOH
Metronidazole	0.1 - 10mg ml^{-1}	0.05-5µg ml^{-1}	In sterile H_20
Spectinomycin dihydrochloride	50mg ml^{-1}	25µg ml^{-1}	In sterile H_20
Streptomycin sulphate	100mg ml^{-1}	200µg ml^{-1}	In sterile H_20
Tetracycline hydrochloride	10mg ml^{-1}	10 or 5µg ml^{-1}	In 50% ethanol
Trimethoprim	5mg ml^{-1}	125µg ml^{-1}	In methanol

Table 2.2 List of Reagents

Reagent	Stock solution concentration	Working solution concentration	Preparation
IPTG	2M	0.1 – 1mM	In sterile H_20
PMSF	250mM	1mM	In ethanol
Lysozyme	1mg ml^{-1}	100µg ml^{-1}	In sterile H_20
DTT	1M	As required	In 10mM sodium acetate pH5.2

2.3 Strains

Table 2.3 List of Strains

Strain	Genotype	Source
DH5α	F$^-$ Φ80*lac*ZΔM15Δ(*lac*ZYA-*arg*F)U169 *rec*A1 *end*A1 *hsd*R17(r_k^-, m_k^+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ$^-$.	Invitrogen
JM109	*end*A1, *rec*A1, *gyr*A96, *thi*, *hsd*R17 (r_k^-, m_k^+), *rel*A1, *sup*E44, Δ(*lac-pro*AB), [F´ *tra*D36, *pro*AB, *laq*I^{q}ZΔM15].	Stratagene
ER2566	F- λ- *fhuA2 [lon] ompT lacZ*::*T7 gene1 gal sulA11* Δ(*mcrC-mrr*)*114*::*IS10R*(*mcr-73*::*miniTn10–TetS*)2 *R*(*zgb-210*::*Tn10*)(*TetS*) *endA1* [*dcm*].	NEB
BL21(DE3)	F- *ompT hsd*S_β (r_β^-, m_β^-) *gal dcm* (DE3).	Novagen
BL21(DE3) pLysS	F- *omp*T, *hsd*S_B (r_β^-, m_β^-) *dcm gal* λ(DE3) pLysS, Cm^r.	Promega
BL21-CodonPlus(DE3)-RIL	F- *ompT hsdS*(r_β^-, m_β^-) *dcm*+ Tetr *gal* λ(DE3) *endA* Hte [*argU ileY leuW* Cam^r]	Stratagene
BL21-CodonPlus(DE3)-RP	F- *ompT hsdS*(r_β^-, m_β^-) *dcm*+ Tetr *gal* λ(DE3) *endA* Hte [*argU proL* Cam^r]	Stratagene
Rosetta-gami(DE3)pLysS	Δ(*ara–leu)7697* Δ*lacX74* Δ*phoA PvuII phoR araD139 ahpC galE galK rpsL*(DE3) F'[*lac+ lacI q pro*] *gor522*::Tn*10 trxB* pLysSRARE2 (CamR, KanR, StrR, TetR)4	Novagen
NM1035	NK402 *hsdR*(A957V) Δ*recD*	Garry Blakely
MG1655	F$^-$ LAM$^-$ *rph*-1. λ$^-$, *ilv*G, *rfb*-50	*E. coli* Genetic Resource Center
FP101	MG1655 Δ*recBCD*::Kan	This study
S17-1 (λpir)	*recA thi pro hsd* RM+ RP4::2-Tc::mu::Km Tn 7 λ pir	Simon et al 1983
NCTC9343	*B. fragilis* wild type strain – clinical isolate	Sheila Patrick
FP201	NCTC9343 Δ*addAB*::*ermF*	This Study
FP202	NCTC9343 Δ*addB*	This Study
FP203	NCTC9343 Δ*sbcCD*	This Study

2.4 Plasmids

Table 2.4 List of Plasmids

Plasmid	Description	Source
pUC19	Cloning vector, Amp^R	Yanish-Perron et al., 1985
pTYB2	Expression vector for fusion of protein to Chitin Binding Domain protein. Amp^R	NEB
pMAL-c2X	Expression vector for fusion of protein to Maltose Binding Protein. Amp^R	NEB
pTRC99a	Expression vector, Derivative of pKK233-2 Amp^R	Amersham (GE Healthcare, UK)
pLYL01	Expression vector, Amp^R Mob+ (Tc^R only in *Bacteroides*)	Salyers et al., 1995
pEP185.2	Suicide vector, Mob+ pir dependent ori RK6 Cm^R	Kinder et al 1993
pGB909	As pEP185.2 with I-SceI cleavage site	Zubin Thaker
pGB910	pGB909 containing *ermF*	Garry Blakely
pGB920	pLYL01 containing I-SceI under control of fucose promoter	Zubin Thaker
pFP101	As pTRC99a with *addAB* in SacI/XbaI	This Study
pFP102	As pTRC99a with *addA* in SacI/XbaI	This Study
pFP103	As pTRC99a with *addB* in SacI/XbaI	This Study
pFP104	As pTRC99a with *recD* in NcoI/BamHI	This Study
pFP105	As pFP104 with *addA* in BamHI/PstI	This Study
pFP106	As pLYL01 with *addAB* in SmaI	This Study
pFP107	As pMAL-c2X with *addA* in BamHI/PstI	This Study
pFP108	As pMAL-c2X with *addB* in BamHI/SalI	This Study
pFP109	As pMAL-c2X with *recD* in EcorI/BamHI	This Study
pFP110	As pTYB2 with *addA* in NheI/XhoI	This Study
pFP111	As pTYB2 with *recD* in NdeI/XhoI	This Study
pFP201	As pGB909 with *addAB* deletion cassette in EcoRV	This Study
pFP203	As pGB910 with *addA* deletion cassette in EcoRV	This Study
pFP204	As pGB910 with *addB* deletion cassette in EcoRV	This Study
pFP202	As pGB910 with *recF* deletion cassette in EcoRV	This Study
pFP205	As pGB910 with *sbcCD* deletion cassette in EcoRV	This Study

2.5 DNA Oligonucleotides

Table 2.5 List of DNA Oligonucleotides

Primer Title	Sequence
Bf rexA For	TTCTCGAGATGAGCGAACTCATTGTCTATAAAG
Bf rexA Rev	TTGGATCCGCGCGACCAATTGGAGAAATG
Bf recD for	TTCATATGATAAATAACTATTTAGAGCGGC
Bf recD Rev	TTTGGATCCTCACTCTATTTGCTCTTCCGG
pTrc99aRA	TTGAGCTCATGAGCGAACTCATTGTCTATAAAGC
pTrc99aRD	TTCCATGGTAAATAACTATTTAGAGCGG
pTrcDAAfor	TTT TTG GAT CCA TGA GCG AAC TCA TTG TCT ATA AAG C
pTrcDAArev	TTT TTT TCT GCA GAT TTT ACA GTA AAG GTC CTG GCG
RexBpTrcFor:	TTGAGCTCATGAAAACATTTCTCCAATTGG
RexBpMalFor:	TTGGATCCAAAACATTTCTCCAATTGGTCG
RxBtrcRev	TTTCTAGACCGCTGTTATATATGGATAGGAGG
RA	TTTCTAGAGCGCGACCAATTGGAGAAATG
RA I FOR	TTGCTAGCGAACTCATTGTCTATAAAG
RA I REV	TTCTCGAGTACTTTTTCGATTATTTCTTCTTC
RA P FOR	TTGGATCCATGAGCGAACTCATTGTCTATAAAGC
RD I FOR	TTCATATGATAAATAACTATTTAGAGCGGC
RD I REV	TTCTCGAGCTCTATTTGCTCTTCCGG
RD P FOR	TCTTCGAATTCATGATAAATAACTATTTAGAGCGGC
RDbluntHelicase3	GTTATGCGTTTGGCGGAGAGGGGCGCGCAACCGGCTAAGT
RDbluntHelicase5	ACTTAGCCGGTTGCGCGCCCCTCTCCGCCAAACGCATAAC
RD3overhangHelicase3 (Chapter 4 Section 3.7.1)	CTCGCGTCGCTCAGTCACTCGCTCCTTCGCCTTCTCGCGGGTTA TGCGTTTGGCGGAGAGGGGCGCGCAACCGGCTAAGT
RD3overhangHelicase5 (Chapter 4 Section 3.7.1)	ACTTAGCCGGTTGCGCGCCCCTCTCCGCCAAACGCATAAC
RD5overhangHelicase3 (Chapter 4 Section 3.7.1)	GTTATGCGTTTGGCGGAGAGGGGCGCGCAACCGGCTAAGT
RD5overhangHelicase5 (Chapter 4 Section 3.7.1)	ACTTAGCCGGTTGCGCGCCCCTCTCCGCCAAACGCATAACCCG CGAGAAGGCGAAGGAGCGAGTGACTGAGCGACGCGAG
recDNitail	CGATGCACGCGTAAAAGCTGTCTCCGAAAGAGATTTTACAGC
recDNo500	ATCGTACATGCTATGTCTTACG

rexANITAIL	GTGTCCAATAGACACCATACATCTGGAGGCTTTATAGACAATGAGTTC
rexBNiTAIL	GAGTGAATTCCTCTTTGTCATCGTCTATTTTACAGTAAAGGTCCTGCG
recDCi	ACAGCTTTTACGCGTGCATCG
recDCo500	ATACATGGAAGGATTGAAAGGC
recFNi	TAGAAAATCCGACTTACTACGATCGAGCAGATTGGTTTTTCC
recFNo500	TTTAGTACCGCTAAACTGATCG
recFCi	GAGCTATCAATGAAATGGAGG
recFCo500	TAGAAAATCCGACTTACTACG
sbcCDNitail	GCGATGGAGTTGCACTTGCACATCGGCTGTATGTAGTATACG
sbcCDNo500	GAAGACCACTAATGATACTCC
sbcCDCi	GTGCAAGTGCAACTCCATCGC
sbcCDC0500	GTAATCCCGATTCGGAAGCGG
RxBNi	CTATTTTACAGTAAAGGTCCTGCG
RxBNo500	ATCTGCTGATTTTATCAAGGGGCT
RxBNo1000	ACTGGAATTTCATACGGTACTCCT
RxBCi	ACGATGACAAAGAGGAATTCACTC
RxBCo500	CGGTAGTTAGTGTTAAAGATAGCG
RxBCo1000	AACTGATTATTCGGGTCAGTCATC
RxANi	GGAGGCTTTATAGACAATGAGTTC
RxANo500	ACGATTATCTGTTCCCTTACCCTA
RxANo1000	TTAGAAAGCATCCGAAACTTACCG
RxACi	CACAGGTTATCTGTGGTATGTAGA
RxACo500	GATATCGGCGATAAATATCACCCA
RxACo1000	ATCCCATGGTGATATTGACATTCC
RxANiermF	GCAATAGCGGAAGCTATCGGGGAGGCTTTATAGACAATGAGTTC
RxACiermF	TCAAGATGATTTTTTTGGTTTGGTCACAGGTTATCTGTGGTATGTAGA
RxBNiermF	GCAATAGCGGAAGCTATCGGCTATTTTACAGTAAAGGTCCTGCG
RxBCiermF	TCAAGATGATTTTTTTGGTTTGGTACGATGACAAAGAGGAATTCACTC
RecD-700-For:	TGCCTAAGATAAAAGTAAGCGG
RecD-700-Rev:	TCCATTCCGTCTTTTTCGTAGC
RexA-650-For:	CCGAATGTAATCCATAATTATCG

RexA-650-Rev:	TACATTTTCAAACTGTTGAGCG
RexA-1300-For:	TTGATTGTAGGTGATGTCAAGC
RexA-1300-Rev:	CTGTACGGATAGAAAAATATCC
RexA-1950-For:	TCGGAGTATGATCTCAATTCG
RexA-1950-Rev:	CTCTAACAGTTCATACAGAGG
RexA-2600-For:	CAGGTATATGAATTAGGTGACC
RexA-2600-Rev:	CATAGATTCCATATTGACCGG
Sq_rB_F500:	AATTGGGTGATATTTATCGCCG
Sq_rB_F1050:	ATATCACCATGGGATTTCCG
Sq_rB_F1519:	ACATTCCTTTTCATGGTGAACC
Sq_rB_F2003:	AAACACCTATCTTGATTGTCGG
Sq_rB_F2500:	ATAGCAAAGAAGGTACACTCC
Sq_rB_R500:	ATATCGGCGATAAATATCACCC
Sq_rB_R1000:	AAAACAGGGAGAAGTAAAGCC
Sq_rB_R1500:	AACGTTTCAGAGTTTCTATGCG
Sq_rB_R2000:	ATAGGTGTTTAATGCTGAAGGC
Sq_rB_R2500:	TATATTTGCCGGTACTTTGGG

2.6 Bacterial Methods

2.6.1 Preparation of Competent Cells – Rubidium chloride method

A colony of the desired strain was used to inoculate 5ml of LB with appropriate antibiotics. This was grown overnight with shaking at 37°C. A 1l flask containing 250ml of LB with 20mM $MgSO_4$ was pre-warmed before a 1/100 dilution of the overnight culture. This was incubated 37°C with 160 rpm shaking until an OD_{600} of 0.4 was reached. The culture was centrifuged at 4,500 x g in a Sorvall RC-5B centrifuge, with GSA rotor (Sorvall Ltd, UK) for 5 minutes at 4°C. The supernatant was discarded and the pellet was resuspended in 100ml ice-cold TFB1 followed by a 5 minute incubation on ice. The cells were pelleted by centrifugation as described above. The supernatant was discarded and the pellet was resuspended in 10ml ice-cold TFB2 followed by a 1 hour incubation on ice. Aliquots of the cells (200µl/tube) were snap-frozen in liquid nitrogen and stored at -80°C.

TFB1: 30mM CH_3COOK, 10mM $CaCl_2$, 50mM $MnCl_2$, 100mM RbCl, 15% v/v Glycerol, pH adjusted to 5.8 with 1M acetic acid then filter sterilised with a 0.22µm filter.

TFB2: 10mM MOPS, 75mM $CaCl_2$, 10mM RbCl, 15% v/v Glycerol, pH adjusted to 6.5 with 1M potassium hydroxide then filter sterilised with a 0.22µm filter.

2.6.2 Chemical Competent Cells – Calcium chloride method

A colony of the desired strain was used to inoculate 5ml of LB with appropriate antibiotics. This was grown overnight with shaking at 37°C. 10ml of LB in a 100ml flask was pre-warmed before a 1/100 dilution of the overnight culture. The culture was incubated at 37°C with 160rpm shaking until an OD_{600} of 0.4 was reached. One ml of the culture was centrifuged at 13,000 rpm in an Eppendorf 5415 D centrifuge (Eppendorf UK limited) for 2 minutes. The supernatant was discarded and the pelleted cells were resuspended in 500µl of ice cold 50mM $CaCl_2$. The resuspended cells were centrifuged at 13,000 rpm in an Eppendorf 5415 D centrifuge for 2 minutes. The supernatant was discarded and the pelleted cells were resuspended in 200µl of ice cold 50mM $CaCl_2$. The cells were incubated on ice for 1 – 24 hours before transformation.

2.6.3 Transformation – Heat shock method

Competent cells (e.g. *E. coli* DH5α) were defrosted on ice. 2-10µl of DNA were added to 50µl of competent cells and left to incubate on ice for 30 minutes. The cells were heat shocked at 37^{o}C for 30 seconds followed by incubation on ice for 5 minutes before addition of 500µl of pre-warmed LB. The cells were incubated at 37^{o}C for the appropriate time dependent on the antibiotic resistance. The culture was centrifuged in an Eppendorf 5415 D centrifuge at 13,000 rpm for 2 minutes. The supernatant was discarded and the pelleted cells were resuspended in 50µl of pre-warmed LB. The resuspended cells were spread on appropriate media with antibiotics and incubated overnight at 37^{o}C.

2.6.4 Transformation – Electroporation method for *B. fragilis*

5ml of BHI-S were inoculated with a colony of the desired *B. fragilis* strain and incubated anaerobically overnight at 37^{o}C. The overnight culture was diluted 1/40 by inoculating pre-warmed BHI-S and was grown at 37^{o}C anaerobically to an OD_{600} of 0.4. Ten millilitres of the culture were centrifuged using a Sigma ®3-16K, rotor 1180 (Sigma Centrifuges, Germany) at 26000g for 15minutes at 4^{o}C. The supernatant was discarded and the pelleted cells were washed in 10ml of dH_2O, which had been left overnight in an anaerobic environment. This was followed by centrifugation (Sigma ®3-16K, rotor 1180) at 26000g for 15minutes at 4^{o}C. The supernatant was discarded and the pelleted cells were resuspended in 10ml of anaerobically incubated dH_2O. These cells were centrifuged as previously described, the supernatant was discarded and the pelleted cells were washed in 5ml of dH_2O. The cells were centrifuged as previously described, the supernatant was discarded and the pelleted cells were washed with 100µl of dH_2O and left on ice till required. 40µl of the cells were placed in a pre-chilled electroporation cuvette (BIO_RAD Gene Pulser® (Bio-Rad Laboratories Ltd, UK) Cuvette 0.1cm gap electrode), to which 0.75-2µl of DNA were added. The cells were electroporated in a BIO_RAD MicroPulser™ with a voltage of 2kV. One ml of pre-warmed BHI-S was added to the cells followed by anaerobic incubation at 37^{o}C. The cultures were centrifuged in an Eppendorf 5415 D centrifuge at 13,000rpm for 2 minutes then resuspended in 50µl of pre-warmed BHI-S. The cells were spread on appropriate media/antibiotic plates and incubated for 2 days in an anaerobic atmosphere at 37^{o}C.

2.6.5 Conjugation – *E. coli* to *B. fragilis*

A colony of the desired donor strain (based on *E. coli* S17-1 λPir) was inoculated into 5ml of LB with appropriate antibiotics. The culture was grown aerobically overnight with 160rpm shaking at 37°C. A colony of the appropriate recipient *B. fragilis* strain was inoculated into 5ml of BHI-S containing appropriate antibiotics; cultures were incubated overnight anaerobically at 37°C. The overnight cultures were

used to inoculate appropriate fresh medium. The *E. coli* strain was diluted 1/100 into 10ml LB and incubated with 160rpm shaking at 37°C till an OD_{600} of 0.2. The *B. fragilis* strain was inoculated 1/40 into 20ml of BHI-S and incubated anaerobically at 37°C until an OD_{600} of 0.2 was reached. The cultures were mixed in a donor/recipient ratio of 10:1. 250µl of the mixture were placed on sterile 0.45µm filters on BHI-S agar plates. The plates were incubated aerobically at 37°C for 2 hours before anaerobic incubation at 37°C overnight. The discs were removed from the plates and placed in tubes containing 5ml of BHI-S and briefly vortexed. The discs were removed and the cells were pelleted by centrifugation (Sigma ®3-16K, rotor 1180) at 2600g for 10 minutes. The supernatant was discarded and the pellet was resuspended in 50µl of fresh BHI-S and spread onto BHI-S plates containing appropriate antibiotics. The plates were incubated anaerobically at 37°C for 2-4 days.

2.6.6 P1 Transduction

A recipient *E. coli* strain was grown overnight with 160rpm shaking at 37°C. The overnight cultures were diluted 1/100 into 10ml LB and incubated with 160rpm shaking at 37°C till an OD_{600} of 0.4 was reached. One ml aliquots were removed and centrifuged (Eppendorf 5415 D centrifuge) at 13,000 rpm for 2 minutes. The supernatant was removed and the cell pellet was resuspended in 1ml of LB. One hundred microlitres of cells were aliquoted into tubes containing 100µls of 50mM $CaCl_2$ and 100µls of 100mM MgSO4. Either 10µls, 1µl or 0.1µls of P1 lysate were added to the resuspended culture. These were incubated at 37°C for 20 minutes, after which 0.5ml of pre-warmed LB and 200µls of 1mM Na citrate were added, followed by a further incubation at 37°C for 30-60 minutes. Each culture was centrifuged as previously described. The pellet was resuspended in 0.1ml and spread on plates with appropriate selection then incubated overnight at 37°C.

2.6.7 Growth Curve

E. coli: 5ml of LB with appropriate antibiotics were inoculated with a colony of the desired strain and grown overnight with shaking (160rpm) at 37°C. A 1/100 dilution

was made of the overnight culture into 20ml of LB in a 250ml flask. This was incubated at 37°C with shaking (160rpm).

B. fragilis: 5ml of BHI-S with antibiotics if appropriate was inoculated with a colony and incubated overnight at 37°C anaerobically. This was diluted 1/100 into 20ml of BHI-S and incubated anaerobically at 37°C.

The OD at 600nm was measured for a culture at appropriate time intervals. When a spectrophotometer reading was above an OD_{600} of 1, the sample was diluted in appropriate medium so that a reading below 1 could be attained. The measurements were then multiplied by the dilution factor to give the optical density. Measurements were taken until stationary phase was reached. To measure doubling time, the exponential phase was extended to ensure all cells were growing exponentially by diluting cultures 1/10 into fresh medium when an OD_{600} of 0.2 was reached. All growth cultures were performed in triplicate.

2.6.8 UV Sensitivity Test – Plate method

Overnight cultures of strains were serially diluted in bacterial buffer (0.03%w/v KH_2PO_4, 0.07%w/v $Na_2H\ PO_4$, 0.04%w/v NaCl and 0.002%w/v $MgSO_4.7H_2O$). Ten micro litres of each dilution were spotted onto BHI-S plates. The plates were irradiated with various doses of UV (0-100 KJ/m^2); the plates were incubated anaerobically at 37°C for 24-48hrs. All experiments were performed in triplicate.

2.6.9 UV Sensitivity Test – Culture method

Spot Test: A colony from the desired *B. fragilis* strain was inoculated into 5ml of BHI-S and incubated anaerobically at 37°C until stationary phase. The culture was diluted 1/100 into 20ml of fresh medium and incubated at 37°C until an OD_{600} of 0.5 was reached. The culture was pelleted by centrifugation (Sigma ®3-16K, rotor 1180) at 26000g for 15minutes. The pellet was resuspended in 40ml of bacterial buffer. 5ml of the suspension were aliquoted into an open Petri dish and irradiated with various

doses of UV (0-25 KJ/$m^{2)}$. The suspension was then serially diluted and 10µl of each dilution was spotted on BHI-S plates. The plates were protected from light and incubated anaerobically at 37°C for 24-48hrs. This allowed appropriate parameters to be selected for further testing. All experiments were performed in triplicate.

Full Plate: A colony from the desired *B. fragilis* strain was inoculated into 5ml of BHI-S and incubated anaerobically at 37°C until stationary phase. The culture was diluted 1/100 into 20ml of fresh medium and incubated at 37°C until an OD_{600} of 0.5 was reached. The cells were pelleted by centrifugation (Sigma ®3-16K, rotor 1180) at 26000g for 15minutes. The pellet was resuspended in 40ml of bacterial buffer. 5ml of the suspension were aliquoted into an open Petri dish and irradiated with various doses of UV (0-25 KJ/$m^{2)}$ then serially diluted. One hundred µls of each dilution was plated on BHI-S plates. The plates were protected from light and incubated anaerobically at 37°C for 24-48hrs. All experiments were performed in triplicate.

2.6.10 *B. fragilis* SOS Induction Test

A colony from the desired *B. fragilis* strain was inoculated into 5ml of BHI-S and incubated anaerobically at 37°C till stationary phase was reached. The culture was diluted 1/100 into 20ml of fresh BHI-S and incubated anaerobically at 37°C until OD_{600} of 0.2. To determine cell morphology prior to antibiotic addition a sample was removed for microscopy. Cultures either had no addition of antibiotic or metronidazole was added to a culture at either 5 or 0.5µg/ml; followed by incubation anaerobically at 37°C for 2 hours before a second sample was removed for microscopy. All experiments were performed in triplicate. The samples were then viewed using a Leitz Wetzlar Metalux II microscope (Leitz Wetzlar now Leica-microsystems, Germany); images were acquired with a Hamamatsu digital c4742-95 camera (Hamamatsu Photonics K.K., Japan) using Improvision Openlab software (PerkinElmer Inc., USA).

2.7 DNA Methods

2.7.1 Plasmid DNA

Plasmid DNA was purified using the Qiagen QIAprep Spin Mini Prep kit or the Qiagen QIAfilter Plasmid Midi Kit (Qiagen, Germany). If using the Mini Prep kit, 1-25ml of overnight culture were used as per manufacturer's instructions, and eluted using 30-50µls of elution buffer. If the Midi Kit was used, 100ml of overnight culture were used as per manufacturer's instructions; DNA was eluted using 500µls of 1xTE buffer. DNA was stored at -20°C.

1x TE buffer: 10mM Tris-HCl, 1mM EDTA pH8

2.7.2 Chromosomal DNA

To purify chromosomal DNA, Promega Wizard SV Genomic DNA Purification System (Promega, USA) or the BIO-RAD AquaPure Genomic DNA isolation kit was used. Kits were used as per the manufacturer's instructions.

2.7.3 DNA Quantification

The concentration of DNA in a sample was assayed using a UV spectrophotometer at 260nm. An UV_{260}nm reading of 1.0 is equivalent to 50µg/ml of double stranded DNA. To determine the purity of the sample, the concentration of DNA to protein ratio, UV_{260}/UV_{280} was used: a ratio of 1.8-2 suggested that DNA was of a reasonable level of purity.

2.7.4 Synthetic DNA Oligonucleotides

All oligonucleotides were ordered from MWG (Eurofins MWG Operon, Germany). The dessicated oligonucleotides were resuspended in 1xTE to give a stock solution

of concentration 100pM. Stock solutions were diluted with dH_2O to give working solutions of 10pM. All solutions were stored at -20°C.

2.7.5 Polymerase Chain Reactions

PCR reactions were performed in a Techne Progene Thermal Cycler (Bibby Scientific Limited, UK) or an Eppendorf Mastercycler gradient cycler. A standard PCR contained 1x polymerase buffer, 0.2mM dNTPs (Promega), 0.2pM forward primer, 0.2pM of reverse primer, 100µl of DNA and 1.5 to 3 units of Taq polymerase (NEB – New England Biolabs Inc., USA), dH_2O was added to give a final volume of 50µl.

Initial denaturation was performed at 94°C for 3 minutes. This was followed by 30 cycles of denaturation at 94°C for 1 minute, annealing at primer specific temperatures for 1 minute and extension at 72°C for the time appropriate for the template size and for the enzyme used: in the case of Taq polymerase an extension time of 1 minute per kb of DNA was used.

Depending on the complexity of the template or the further use of the product other polymerases were used: Phusion (Finzymes – Thermo Fisher Scientific, USA) and Pfu (Promega). For these reactions standard conditions were adjusted as per the manufacturer's instructions. Primers, enzymes and buffers were removed using Qiagen PCR purification kit as per the manufacturer's instructions.

2.7.6 Agarose Gel Electrophoresis

DNA was separated by agarose gel electrophoresis. Gels were made to 0.8-1.5% w/v Agarose (Invitrogen – Invitrogen Corp., USA) in 1xTAE with microwave heating until dissolved. Gels were cast in appropriate sized trays. When the gels were set they were placed in Bio-Rad Electrophoresis tanks (mini, midi or subcell) GT trays containing 1x TAE buffer. DNA samples were mixed with 6 x DNA loading buffer II (Maniatis et al., 1982) to a final concentration of 1x. Appropriate ladders (NEB)

were loaded. The current was supplied by a BIO-RAD Power-Pak 300. Standard gels were run at 80 volts for 60 minutes. The gel was stained in ethidium bromide (1µg/ml) for 20 minutes, then destained in dH_2O for 20 minutes. The DNA was visualised using a UVP CCD camera connected to a UVP VisiDoc-It Imaging system (UVP LLC, USA) with a Sony UP-895MD Video Graphic Printer (Sony, Japan).

1x TAE buffer: 40mM Tris, 1mM EDTA pH8, 19mM Glacial acetic acid

6x DNA loading buffer II: 0.25% w/v Bromophenol blue, 0.25% w/v Xylene cyanol, 15% w/v Ficoll

2.7.7 Gel Purification of DNA

To remove unwanted PCR products from a PCR sample, agarose gel DNA purification was used. The DNA was electrophoresed through 1% low melting point agarose. The DNA was viewed using a UVP long-wave UV transilluminator and the desired band was excised from the gel then purified with the Qiagen QIAquick gel extraction kit as per the manufacturer's instructions.

2.7.8 Ligation of DNA

Standard ligations were carried out in 20µl volumes. Ten ng of vector DNA were present with appropriate molar ratios of insert DNA ranging from 1:1 to 1:12 dependant on type of ligation. NEB or Promega T4 DNA ligase was used with 1-4 Weiss units per reaction dependant on the type of ligation (blunt cloning required higher concentrations of ligase). T4 DNA ligase buffer was used each reaction at 1x concentration. Generally the reactions were incubated overnight at 16°C. The ligation was used immediately or frozen till required.

2.7.9 DNA Digestion by Restriction Endonucleases

A standard digest was performed in 20µl reaction volume. One µl of NEB or Promega restriction enzymes were used with 2µl of 10x reaction buffer, specific to the enzymes, in each reaction; BSA was dependant on enzyme present and was used at a 1/100 dilution. Reaction temperature was generally 37°C though some enzymes required 25°C. Incubations were for 1 hour. Enzymes and buffers were removed using Qiagen PCR purification kit as per the manufacturer's instructions.

2.7.10 Dephosphorylation of DNA

To prevent digested DNA religating it was necessary to dephosphorylate the DNA. Immediately after the digestion reaction, 1µl of NEB Antarctic Phosphatase and Antarctic Phosphatase buffer to 1x concentration was added to the reaction and incubated at 37°C for 30 minutes. Enzymes and buffers were removed using Qiagen PCR purification kit as per the manufacturer's instructions.

2.7.11 Phosphorylation of DNA

Phosphorylation was required for some DNA products prior to ligation. A standard reaction contained 1µl of NEB T4 polynucleotide Kinase, 1x T4 polynucleotide Kinase buffer, 1x T4 DNA ligase buffer, and 26µl of DNA. The reaction was incubated at 37°C for 30 minutes. Enzymes and buffers were removed using Qiagen PCR purification kit as per the manufacturer's instructions.

2.7.12 Single Colony Gels

Single colonies from transformation plates were patched onto fresh plates with appropriate selection and incubated at 37°C overnight. The patch of cells was removed with a sterile toothpick and placed in a 0.5ml tube containing 130µls of SCFS buffer. The cells were left at room temperature for 30 minutes. The toothpick was removed and the suspension was centrifuged for 30 minutes in an Eppendorf

5415 D centrifuge at 13,000 rpm. 50µls of the supernatant were electrophoresed through a 0.8% gel at 60 volts for 2 hours. Supercoiled plasmid DNA was used as a control. DNA was visualised as previously described (2.7.6).

SCFS Buffer: 2.5% v/v Ficoll, 1.25% w/v SDS, 0.015% w/v Bromophenol Blue, 10µg/ml RNase A, made to volume in 1xTAE buffer

2.7.13 DNA Sequencing

Sequencing reactions were performed in a Techne Progene Thermal Cycler or an Eppendorf Mastercycler gradient cycler. The ABI PRISM Big DyeTM Ready Reaction Cycle Sequence Kit (Applied Biosystem, UK) was used. Reactions contained 20-200ng DNA, 2µls BIG DYE, 3.2µls of primer (at 1pmol/ml) and final volume made up to 20µls by addition of dH_2O. A standard sequencing program consisted of an initial denaturation of 96°C for 1 minute followed by 30 cycles of denaturation at 96°C for 30 seconds, annealing at 50°C for 30 seconds and extension at 60°C for 4 minutes. Reactions were processed by the University of Edinburgh SBS sequencing service on an ABI Prism 3100 DNA Sequencer.

2.7.14 Radiolabelling of DNA

5′ end-labelling of DNA with ^{32}P: To 5′ end-label DNA, ^{32}P γATP (Amersham Redivue 0.37MBq/µl) was used. The labelling reactions were performed in a Techne Progene Thermal Cycler. A standard reaction contained 5µl ^{32}p γATP, 1µl Promega Polynucleotide Kinase, 2µls of 10x Polynucleotide kinase buffer, 5µls DNA and 7µls dH_2O. The reaction was incubated at 37°C for 30 minutes. 25mM EDTA pH8 was added before a further incubation at 72°C for 10 minutes. To remove unincorporated ^{32}P, the sample was placed in a Microspin G25 column (Amersham) and centrifuged at 3,000rpm for 2 minutes in an Eppendorf 5415 D centrifuge. The activity of the labelled DNA was measured: the counts per second (cps) for 1µl of the sample should be between 500-1000cps. This was used straight away or stored at -20 °C.

2.7.15 Chromosomal Gene Deletion

Marked gene deletion method: DNA sequences consisting of 500bp upstream and downstream of the gene to be deleted were amplified by PCR. The primers contained tail regions that were homologous to either end of the erythromycin gene (*ermF*). Crossover PCR allowed the upstream and downstream sequences to be joined to either side of the erythromycin resistance gene: producing a deletion cassette. The deletion cassette was cloned into a suicide vector (pGB909) which contained an I-SceI cleavage site. The suicide vector was conjugated into *B. fragilis*; colonies that were erythromycin resistant were selected as these colonies were single crossovers with the chromosome. A plasmid (pGB920) that expressed I-SceI meganuclease (inducible by fucose) was introduced into the *B. fragilis* single crossover strains, with selection for resistance to tetracycline. The selected colonies were plated on defined minimal media containing tetracycline and fucose. Fucose allowed induction of the I-SceI Meganuclease, thus producing a DNA double strand break on the chromosome. Repair of the break would involve homologous pairing of DNA regions (either the upstream or downstream region) effectively removing the gene completely. Colonies were screened for erythromycin resistance, chromosomal DNA was extracted and screened by PCR for the presence of the gene or deletion cassette.

Unmarked gene deletion method: The production of unmarked deletions was as described above. The deletion cassette did not contain *ermF* but consisted only of the upstream and downstream region of the desired gene. The deletion cassette was cloned into a suicide vector (pGB110) which contained an I-SceI cleavage site and erythromycin resistance. The protocol was followed as above until selection for double crossovers: colonies were screened for erythromycin sensitivity, chromosomal DNA was extracted and screened by PCR for the presence of the gene or deletion cassette.

2.8 Protein Methods

2.8.1 SDS-Polyacrylamide Gel Electrophoresis

Protein samples were resolved using non-denaturing polyacrylamide gels. Gels were cast in glass plates for use in Bio- Rad Mini Protean 3 electrophoresis tanks. A standard gel was prepared contained 12% acrylamide, 1x resolving buffer and 0.1% SDS made up to volume with dH_2O. Immediately before application to the plates 37.5µls of 10% APS and 3µls of TEMED were applied and mixed. Approximately 3.3ml of running gel were poured into the plate and layered with 250µls of Isopropanol. Once the gel was set the Isopropanol was removed. 25µls of APS and 2µls of TEMED were added to the stacking gel before addition to the plates. Before the staking gel polymerised, combs were placed in the plates to provide wells for the protein samples. Once the stacking gels were polymerised the combs were removed and the gels were placed in Bio-Rad Mini Protean 3 electrophoresis tanks containing 1x Running buffer.

Protein samples were prepared by incubation at 100°C for 10 minutes with NEB 1x SDS sample buffer. Cell samples were prepared by centrifugation of 0.5ml – 1ml culture for 2minutes at 13,000rpm. The supernatant was removed and the pellet was resuspended in an appropriate amount of NEB 1x SDS sample buffer (for a 0.5ml culture at OD_{600} of 0.4, pellets were resuspended in 30-40µls of loading buffer) then incubated at 100°C for 10 minutes. 20µls of the sample were loaded into the wells along side NEB protein ladder. The gel was electrophoresed at 200V until the dye reached the bottom of the well (about 45 minutes). The gels were washed in dH_2O to remove SDS, then stained for 30 minutes with Coomassie stain, followed by destaining for 30 minutes with Coomassie destain and left overnight in dH_2O to rehydrate. Gels were dried onto blotting paper with a Bio-Rad gel dryer at 80°C for 25 minutes.

4x Resolving Buffer: 1.5M Tris pH 8.8

4x Stacking Buffer: 0.5M Tris pH 6.8

10 x Running Buffer: 0.25M Tris base, 1.9M Glycine

1x Running Buffer: 1/10 dilution of 10 x Running buffer with SDS added to final concentration of 0.1% w/v

Coomassie Stain: 45%v/v dH_2O, 45%v/v Methanol, 10%v/v Glacial acetic acid, 0.125%w/v Coomassie brilliant blue

Coomassie Destain: 45%v/v dH_2O, 45%v/v Methanol, 10%v/v Glacial acetic acid

2.8.2 Protein Purification – Preparation of Crude Cell Lysates

The appropriate strain was incubated at 37°C with shaking in 500-2000ml of LB until an OD_{600} of 0.4 was reached. Protein expression was induced by addition of 1mM IPTG with incubation at 18°C overnight with shaking. The cells were pelleted by centrifugation at 6000rpm at 4°C in a Sorvall RC-5B centrifuge with a GSA rotor. The supernatant was discarded and the pellet was resuspended in 5-20ml of lysis buffer containing lysozyme. The cells were incubated on ice for 30 minutes before sonication at 8 microns with 10 x 10 seconds bursts interrupted by 20 second intervals to stop overheating. The cells were then centrifuged at 1300rpm for 20 minutes at 4°C in a Sorvall RC-5B centrifuge with a S34 rotor. The supernatant was removed and kept on ice or stored in 50 % glycerol at -20°C until required.

An adaptation of this method was used to improve lysis and solubility of proteins. The above protocol was followed till the sonication step. At this point a Brij lysis method was followed. Brij 58 (Sigma-Aldrich Co Ltd, UK) was applied to the cells slowly over 30 minutes with stirring at 4°C to a final concentration of 0.3% was reached. The cells were centrifuged and the supernatant was removed. Ammonium

sulphate was slowly added to the supernatant over 2 hours with gentle stirring at 4°C to a final concentration of 50%. The precipitated proteins were recovered by centrifugation.

2.8.3 Protein Purification – FPLC

A Pharmacia FPLC system consisting of a pump (p5000), gradient programmer (GP250) and valve (v7) (Pharmacia Ltd., UK) was used for ion exchange chromatography to purify proteins. Mono-S, Mono-Q, Q-sepharose, Tricorn Superose 12 10/300 GL and Hydroxyappatite columns (GE Healthcare, UK) were used. In each case the columns were equilibrated with the appropriate buffer before use. Flow rates used were dependant on column type. Elutes were monitored for protein concentration by UV_{280}nm absorbance.

2.8.4 Hydroxylappatite Column Packing

All hydroxyappatite columns were prepared in-house. DNA grade Bio-Gel® HTP from Bio-Rad was used. Before packing the column the Bio-Gel must be rehydrated. One gram of Bio-Gel hydrated to a volume of 2-3ml. The bio-gel was mixed with degassed buffer in a ratio of 1:6 HTP to buffer w/v, and allowed to settle and the fines decanted. The gel was then packed into a 2ml column.

2.8.5 Purification of Chitin Binding Domain (CBD) Tagged Protein

The Impact system from NEB was used. The *recD* and *addA* genes were ligated into pTYB2 to produce pFP111 and pFP110, respectively, then introduced into host strain. A 1/100 dilution of an overnight of each strain was used to inoculate 1l of fresh LB in a 5l flask. The culture was grown at 20°C with shaking to an OD600 of 0.4 before addition IPTG to a final concentration of 1mM. The culture was centrifuged at 5000g for 10 minutes at 4°C in a Sorvall RC-5B centrifuge with a GSA rotor. The supernatant was discarded; the pellet was removed and stored at -20°C. The pellet was resuspended in 50ml of cell lysis buffer then sonicated with amplitude of 8

microns with 10 rounds of 10 seconds of sonication followed by 20 seconds of rest. The lysate was centrifuged at 20,000g for 30 minutes in a Sorvall RC-5B centrifuge with a S34 rotor. The soluble lysate was removed and loaded onto an equilibrated chitin column. All column steps took place at 4°C. Unbound proteins were removed by washing the column with 10 bed volumes of column buffer. Cleavage buffer was loaded onto the column and the flow of the column was stopped for 16-40 hours, the presence of thiols allowed intein self-cleavage to occur. Flow was started again and proteins were eluted in the cleavage buffer.

Cell Lysis Buffer: 20mM Tris-HCl pH 8.0, 500mM NaCl, 1mM EDTA, 20µM PMSF

Column Buffer: 20mM Tris-HCl pH 8.0, 500mM NaCl, 1mM EDTA,

Cleavage Buffer: 20mM Tris-HCl pH 8.0, 500mM NaCl, 1mM EDTA, 50 mM DTT

2.8.6 Purification of Maltose Binding Protein (MBP) Tagged Proteins

To produce MBP fusion proteins the gene of interest was ligated into pMAL-c2x (NEB). The proteins were over-expressed as previously described (2.8.2). The cells were pelleted by centrifugation at 6000rpm at 4°C in a Sorvall RC-5B centrifuge with a GSA rotor. The supernatant was discarded and the pellet was resuspended in 5-20ml of MBP lysis buffer. The cells were incubated on ice for 30 minutes before sonication at 8 microns with 10 x 10 seconds bursts interrupted by 20 second intervals to stop overheating. The cells were then centrifuged at 1300rpm for 20 minutes at 4°C in a Sorvall RC-5B centrifuge with a S34 rotor. The supernatant was removed and kept on ice. 5ml of NEB Amylose Resin were loaded into a Qiagen 5ml polypropylene column. The column was equilibrated with 8 bed volumes of MBP Column Buffer. The lysate was added to the column and allowed to flow through by gravity. The column was washed with 10 volumes of MBP Column Buffer to remove unbound proteins. 5ml of MBP elution buffer were used to elute the fusion protein: 0.5ml fractions were collected.

MBP Lysis Buffer: 100mM NaCl, 20mM Tris-HCl pH7, 0.1mM EDTA

MBP Column Buffer: 420mM Tris-HCl pH7, 200mM NaCl, 1mM EDTA, 1mM Sodium azide, 1mM DTT

MBP Elution Buffer : 420mM Tris-HCl pH7, 200mM NaCl, 1mM EDTA, 1mM Sodium azide, 1mM DTT, 10mM Maltose

2.8.8 Protein Quantification (Bradford Assay)

To quantify proteins in a sample, the Bio-Rad Protein Assay Kit (Bradford Reagent) was used as per the manufacturer's protocol. A series of protein standards (1.25-25µg/ml) were prepared from diluted BSA (10mg/ml) in 800µl of dH_2O. To each standard 200µls of protein reagent were added, mixed and left to incubate at room temperature for 5 minutes. The OD_{595} was measured for each standard. A curve of standard protein concentration vs. OD_{595} was made. The protein sample was prepared in the same way as the protein standards and the OD_{595} was measured. The standard curve was used to estimate the concentration of protein in the sample.

2.8.9 Dialysis

Proteins were dialysed to remove salts or glycerol. Thermo Scientific Snakeskin® pleated dialysis tubing MWCO 3.5K from Pierce was used. Samples were placed in the tubing and both ends were sealed. The tubing was placed in a volume of buffer 100 times that of the protein sample. This was left at 4°C with gentle stirring for 30 minutes. The buffer was changed and incubated at 4°C with gentle stirring for 30 minutes a further 3 times before the sample was removed.

2.8.10 Concentrating Protein Samples

Proteins were concentrated using centrifugal filter columns. Two columns were used Vivaspin 6, 5000 MWCA by Vivascience Sartorious Group (VWR, UK) and Centricon YM-30, 30000 MWCO by Millipore (Millipore Ltd, UK). Both columns were used as per the manufacturer's instructions.

2.8.11 Mass Spectrophotometry

Protein samples were run on a 12% SDS PAGE gel and stained with Coomassie stain. After destaining, the required bands were excised from and incubated in 300µls of 200mM NH4HCO3 (ABC) with 50% acetonitrile at 30°C for 1hour. The gel slice was then transferred to 300µls of 20mM DTT, 200mM ABC and 50% acetonitrile for 1 hour at 30°C followed by three washes in 300µls of 200mM ABC and 50% acetonitrile. The band was then removed and placed in 100µls of 50mM idoacetamide, 200mM ABC and 50% acetonitrile and left in the dark for 20 minutes. The gel slice was washed 3 times with 500µls of 20mM ABC, 50% acetonitrile. The gel slice was then cut into 2mm by 1mm pieces and placed in a 0.5ml tube to be centrifuged at 13,00rpm for 2 minutes. The bands were immersed in acetonitrile until the gel pieces turned white. The acetonitrile was removed and the gel pieces were allowed to dry. Twenty nine µl of 50mM ABC and 1µl of trypsin were added to the dried gel pieces and incubated at 4°C until the gel pieces swelled. The tubes were sealed and incubated at 32°C for 16-24 hours. The samples were placed in a sonication water bath for 5 minutes before briefly centrifuging to remove debris. A gold MALDI sample plate was prepared by cleaning with methanol. Two and a half µl of matrix were spotted on to the plate sample grid followed by addition of 2.5µl of sample. The plates were left to air dry. The samples were analysed using the Applied Biosystems Voyager MALDI-TOF mass spectrophotometer and associated software. Spectra were analysed using Mascot software (Matrix Science Ltd., UK).

2.8.12 Exonuclease Assay

A master mix was prepared containing all reactants except AddAB. The reaction mix was aliquoted into 0.5ml tubes and incubated for 2 minutes at 37°C before addition of AddAB. The tubes were incubated for a further 10 minutes before the reaction was stopped by the addition of stop buffer. The DNA was loaded onto a 1% agarose gel and electrophoresed for 90 minutes at 60 volts before visualising with ethidium bromide. If 32p labelled DNA was used as the substrate, the gel was dried onto nitrocellulose and exposed to film overnight at -70^{o}C with a Dupont Cronex Lightning-Plus intensifying screen (Sigma-Aldrich Co Ltd, UK). The film (Amersham Biosciences Hyperfilm-HP) was developed in a Konica SRX-101A developer (Konika Minolta, Japan).

Standard Reaction: 25mM Tris-acetate pH7.5, 1mM DTT, 1mM ATP, 0.5mM or 2mM $Mg(Oac)_{2,}$ 0.0025pMol-0.5nMol AddAB, 10nMol DNA, made up to 20µls volume with dH_2O

Stop Buffer: 100mM EDTA, 2.5% SDS w/v, made up to volume with 6x DNA loading Buffer

2.8.13 *In Silico* Analysis

In silico analysis was used assign function to subunits/domains of AddAB. Various programs were used for sequence alignment, best match and domain assignation: COBALT (constraint based alignment tool) provided by NCBI (http://www.ncbi.nlm.nih.gov/tools/cobalt/), Interpro scan provided by EMBL-EBI (http://www.ebi.ac.uk/Tools/pfa/iprscan/), CDART (Conserved Domain Architecture Retrieval Tool) multi-domain search scan provided by NCBI (http://www.ncbi.nlm.nih.gov/Structure/lexington/lexington.cgi) and Pfam (protein families) provided by Sanger Institute (http://pfam.sanger.ac.uk/), PROSITE scan provided by the Swiss Institute of Bioinformatics (http://expasy.org/prosite/). Finally sequence alignments/similarities/domains were deduced by eye.

Chapter 3: Identification of Genes required for DSB repair in *Bacteroides fragilis*

3.1 Introduction

DNA repair is an essential process for all organisms since unrepaired DSB will be fatal. DSBs occur frequently in bacteria due to exogenous sources and endogenous processes such as replication; it is estimated that one DSB will occur per generation in *E. coli* (Cox 2002). It is essential that DSBs are repaired quickly and efficiently to prevent mutations and chromosome degradation. In bacteria the repair steps are divided into pre-synapsis, synapsis and post-synapsis. The pre-synaptic stage processes the DSB into a 3′ single-strand overhang, the substrate required for strand invasion in the synapsis stage and the eventual repair of the DSB.

There are three identified pre-synapsis systems involved in recombination in bacteria; represented by the AdnAB, AddAB and the RecBCD protein complexes. Each system functions in a similar manner but differ in the physical composition of the machinery. At the time of starting this project, the RecBCD type recombination machinery was believed to be restricted to Gram negative bacteria while the AddAB/RexAB system was restricted to Gram positive bacteria.

This chapter investigates the genes encoding proteins involved in pre-synaptic DSB repair in *B. fragilis* NCTC9343. It is important to note that initial work was undertaken in *E. coli* as *B. fragilis* NCTC9343 was considered genetically intractable at the start of the project. Homologues were initially identified through analysis of the *B. fragilis* NCTC9343 genome, and their role in pre-synapsis was then either confirmed or refuted via in vivo *E. coli* work. Once the *B. fragilis* recombination genes had been identified, a new method for genetic manipulation of *B. fragilis* NCTC9343 (produced by other work in the lab) was employed to produce deletion mutants. In this chapter the phenotype of a *B. fragilis* strain in which the genes encoding AddAB homologues had been deleted was observed. The strain showed a deficiency in recombination and sensitivity to DNA damaging agents.

3.2 Identification of putative NCTC9343 pre-synapsis genes

The *E. coli recBCD* system is considered the paradigm for DSB repair in Gram negative bacteria. The initial BLAST search to find *B. fragilis* genes that encoded proteins with homology to the *E. coli* RecBCD failed to identify any open reading frames (ORFs) with significant homology to either RecB or RecC from *E. coli*. One ORF was identified that encoded a protein with 37% identity to *E. coli* RecD (BF0679, figure 3.2.1). The *B. fragilis* gene is 1416 bp potentially producing a protein of 472 amino acids, predicted to be an ATP-dependant exodeoxyribonuclease.

The original annotation of the *B. fragilis* genome identified a gene BF2192 (figure 3.2.1), the predicted product of which shared 25% identity with RexA protein from *Lactococcus lactis*. The BF2192 gene is 3171bp encoding a protein of 1057 amino acids, predicted to be a putative helicase. Cerdeno-Tarraga *et al*., (2005) proposed that *B. fragilis* might utilise a novel system that included both a RecD and an AddA homologue. The first aim was to test this hypothesis.

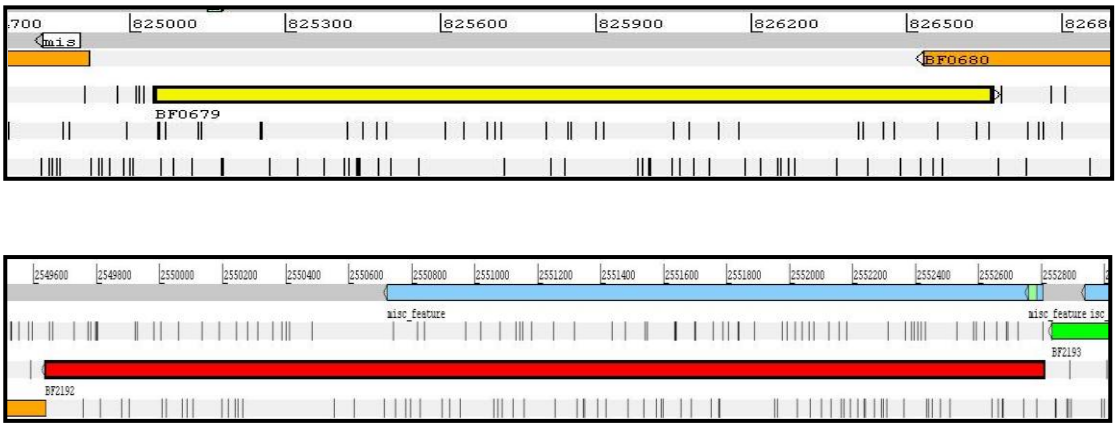


Figure 3.2.1.: Artemis *B. fragilis* NCTC9343 BF0679 and BF2192

Representation of possible recombination genes BF0679 and BF2192, taken from *B. fragilis* NCTC 9343 Artemis annotation (Cerdeno-Tarraga *et al*., 2005)

3.3 In vivo characterisation of putative *B. fragilis* genes in a defined genetic system

Initial work to identify the *B. fragilis* recombination genes was undertaken using *E. coli* as it has been well documented that recombination deficient *E. coli* display a distinct phenotype – such as a slower growth rate and sensitivity to DSB causing agents, including UV irradiation and metronidazole exposure. Analysis of the annotated genome highlighted candidate *B. fragilis* recombination genes BF0679 and BF9192; these genes were introduced into an *E. coli* recombination deficient strain and determined if they could rescue the UV sensitive phenotype.

3.3.1 Phenotypic Characterisation of an *E. coli* Δ *recBCD* strain

Our initial strategy was to generate an appropriate *E. coli* Δ *recBCD* strain and confirm its phenotype. The *E. coli* MG1655 genetic background was chosen as it was derived from a wild type K-12 strain. Compared to many other laboratory strains (such as those used for cloning) it has minimal genetic modification of its genome and its phenotype is close to that of the original wild type strain (Blattner *et al*., 1997).

P1 transduction was used to introduce Δ *recBCD::Km* into MG1655 to produce strain FP101. The *E. coli* MG1655 strain was transduced with a P1 lysate (JJC1086 - kindly provided by D. Leach) so that the *recBCD* genes were replaced with the kanamycin gene. Because the FP101 strain was to be used for identifying the *B. fragilis* recombination genes the phenotype of the strain was confirmed by testing for UV sensitivity and growth defects.

3.3.1.1 Phenotypic characterisation of an *E. coli* Δ *recBCD* strain

Recombination deficient *E. coli* strains are sensitive to UV radiation (Chaudry & Smith 1984). The sensitivity of the *recBCD* null FP101 strain to UV radiation was confirmed. Serial dilutions of stationary phase cultures in bacterial buffer were

performed; 10µl of each dilution was spotted on LB plates. This was replicated so there was one control set of plates without UV exposure, and one set of plates that were exposed to 80J/m^2 UV radiation. All plates were kept in the dark to prevent photo-reactivation and incubated overnight at 37°C. Experiments were performed in triplicate, and a typical example is shown (figure 3.3.1.1.1)

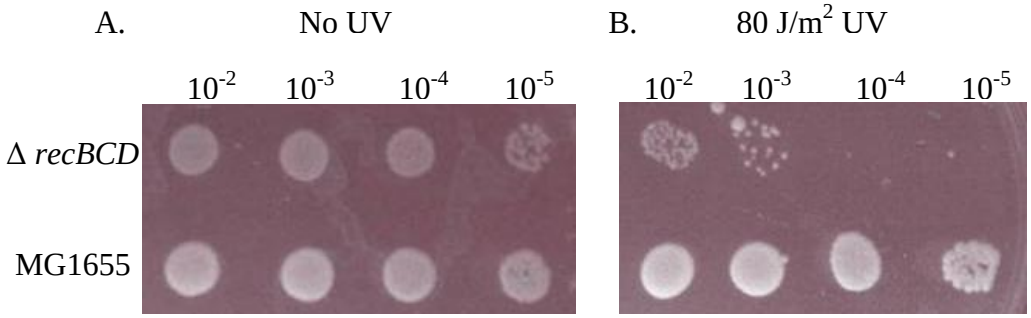


Figure 3.3.1.1.1: plate assay for UV sensitivity of MG1655 compared to Δ *recBCD* strain

Stationary phase cultures of Δ *recBCD* (FP101) and MG1655 were serially diluted and 10µl of each dilution were spotted on LB plates. One set of plates had no UV exposure while the replica set was irradiated at 80 J/m^2 UV. Plates were then incubated at 37°C for 24 hours. **A**. The control plate shows that Δ *recBCD* strain grows slightly less well than MG1655. **B**. The UV irradiated plates showed that Δ *recBCD* strain is sensitive to UV exposure compared to the wild type, MG1655.

In the absence of UV the Δ *recBCD* strain grew less well than the parental strain MG1655 (figure 3.3.1.1.1). Exposure to UV did not alter the viability of the MG1655 strain, whereas the FP101 strain that lacks the *recBCD* genes was significantly affected by the UV exposure (3 $\log_{10}$ reduction in colony forming ability). A strain lacking its recombination system would be unable to repair the excess DSBs caused by UV and would therefore die after exposure to UV.

In the UV test it was noted that on the control plates (no UV exposure – figure 3.3.1.1.1) the FP101 strain had a lower viability than the wild type MG1655. Bacteria that lack their pre-synapsis proteins are deficient for recombination. Such strains have a slower growth rate than wild type strains because of the inability to repair endogenous DSBs that occur during normal cellular processes.

Wild type MG1655 strains grown aerobically in LB at 37°C have doubling times of around 23 minutes; a recombination deficient strain (Δ*recBCD*) should have a longer doubling time than a recombination proficient strain. As the pTrc99a plasmid was used for expression of the *B. fragilis* NCTC9343 genes, the pTrc99a plasmid was introduced into wild type cells and into the deletion strain to test if its presence had any effect on growth. Overnight cultures of MG1655 and FP101 were used to inoculate fresh LB. The cultures were incubated at 37°C with shaking. Optical density 600 nm (OD_{600}) was measured at regular interval times until cultures reached stationary phase. All experiments were performed in triplicate.

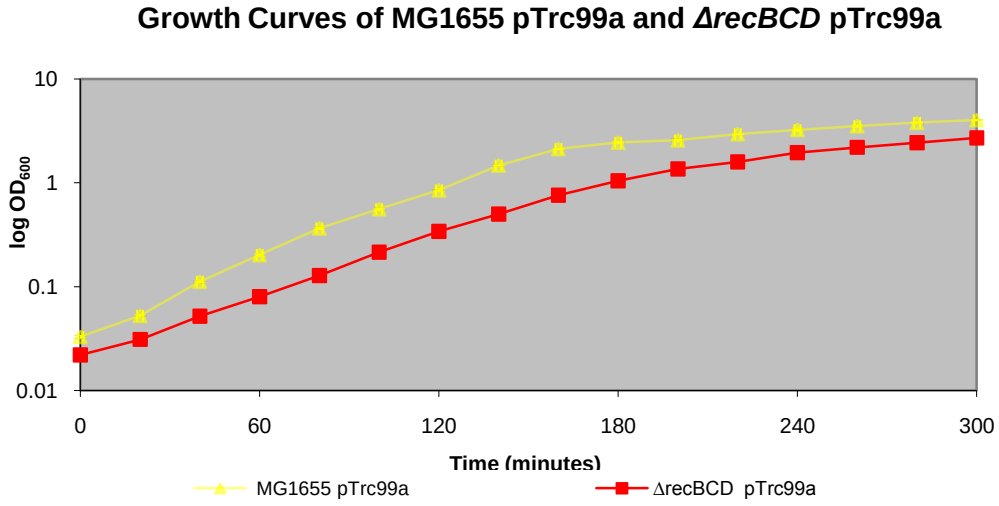


Figure 3.3.1.1.2: Growth Curves of MG1655pTrc99a and Δ*recBCD* pTrc99a

Graph showing the growth of MG1655pTrc99a ▲ and Δ*recBCD* pTrc99a ■ - standard deviation for the triplicate cultures are represented by error bars. The doubling time of MG1655 pTrc99a was 24 minutes whereas doubling time of the Δ*recBCD* pTrc99 strain was 31 minutes.

Figure 3.3.1.1.2 shows the growth curves of the wild type MG1655 strain and the Δ*recBCD* FP101 strain. The doubling time of MG1655 pTrc99a was 24 minutes. The *recBCD* null strain (FP101) had a slower growth rate than the wild type strain with a doubling time of 31 minutes. Under normal environmental conditions DSBs occur due to replication; these would be repaired by the RecBCD protein complex. In a strain that lacks *recBCD*, the repair of DSBs would not be as efficient: other proteins may substitute for RecBCD but cannot act as efficiently. Cells will become

overloaded with chromosomal breaks and will be unable to replicate and will ultimately die; this is reflected in the slower doubling time seen in FP101 (Δ*recBCD*) strain.

3.3.2.1 Identification of *B. fragilis* pre-synapsis proteins - Do BF2192 and BF0679 encode analogues of *recBCD*?

Recombination deficient (*recBC* null) *E. coli* strains are hypersensitive to UV radiation (Chaudry & Smith 1984), therefore our hypothesis was that the introduction and expression of *B. fragilis* pre-synapsis genes in an Δ *recBCD E. coli* strain would rescue the growth defect and UV sensitive phenotype.

The *B. fragilis* NCTC9343 BF0679 and BF2192 genes were both cloned into the *E. coli* expression vector pTrc99a giving plasmid pFP105 (figure 3.3.2.1.1). The FP101 (Δ*recBCD*) strain was transformed with the pFP105 plasmid and was tested for UV sensitivity. Stationary phase cultures were serially diluted in bacterial buffer, then 10µl of each dilution were spotted on replica LB plates, one set of plates was exposed to 80Jm^2 UV while the control set was not, both sets of plates were kept in the dark and incubated overnight at 37°C. All experiments were performed in triplicate and a representative example is shown (figure 3.3.2.1.2).

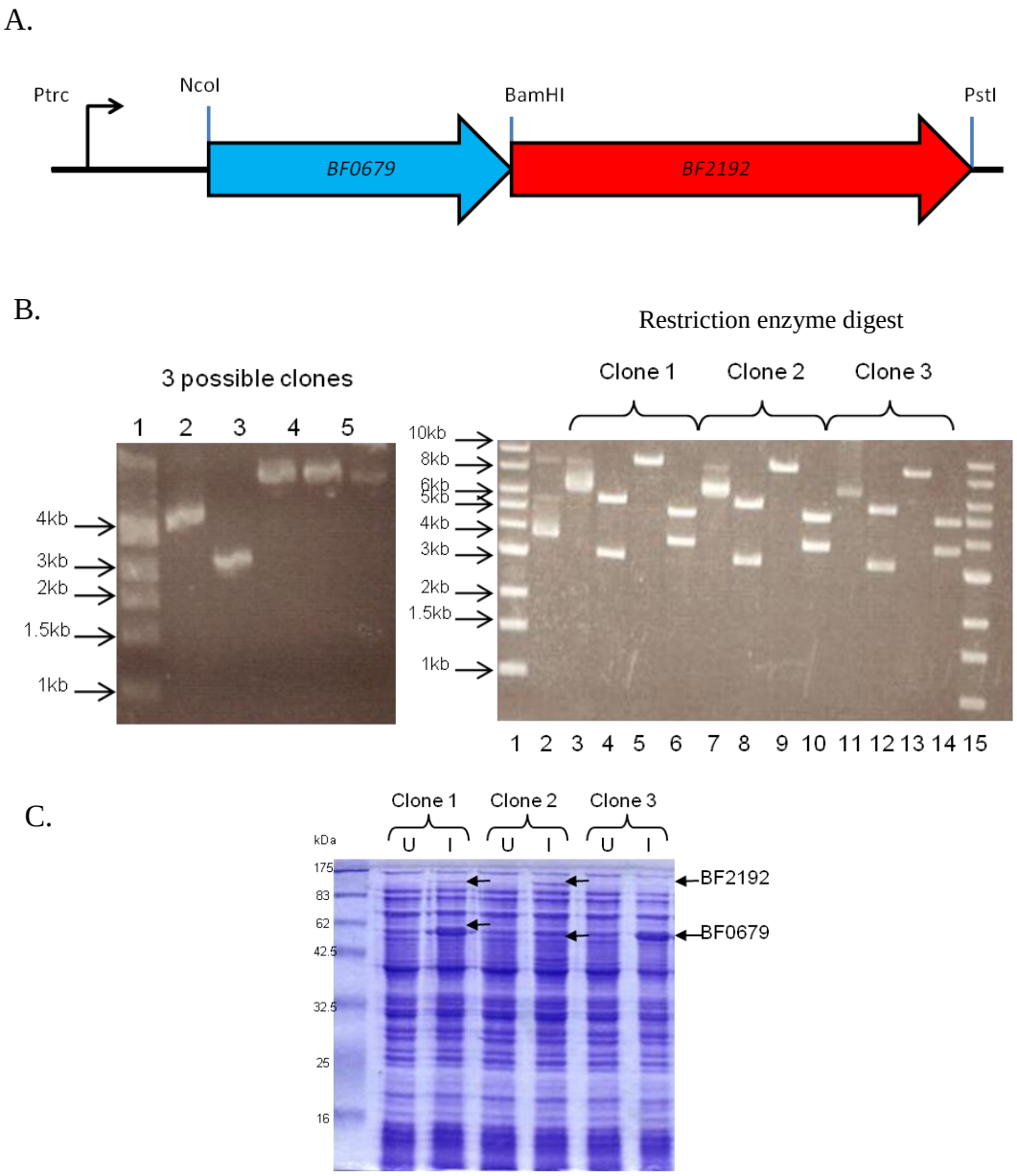


Figure 3.3.2.1.1: Production of pFP105 – BF2192 and BF0679 expression vector

A. The BF0679 gene was ligated into NcoI and BamHI of pTrc99a (giving plasmid pFP104) the BF2192 gene was then ligated into BamHI/PstI giving plasmid pFP105. B. The left panel shows 3 plasmids (lanes 4, 5 and 6) of larger size than the pTrc99a and pFP105 control (lane 2 and 3 respectively) suggesting the presence of a possible insert. The right hand side panel shows the restriction enzyme digests confirmed the presence of insert in the plasmid. Lane 2 – pFP104, lanes 3, 7 and 11 – clones 1, 2 and 3 digested with BamHI/PstI, releasing the 3.2kb insert and the 5.6kb plasmid, lanes 4, 8 and 12 - clones 1, 2 and 3 digested with BamHI, linearising the 8.8kb plasmid, lanes 5, 9 and 13 - clones 1, 2 and 3 digested with NcoI giving two fragments, lanes 6, 10 and 14 - clones 1, 2 and 3. C. SDS-PAGE gel showing the bands of the expected size for the expression products of both BF0679 and BF2192 from all three clones.

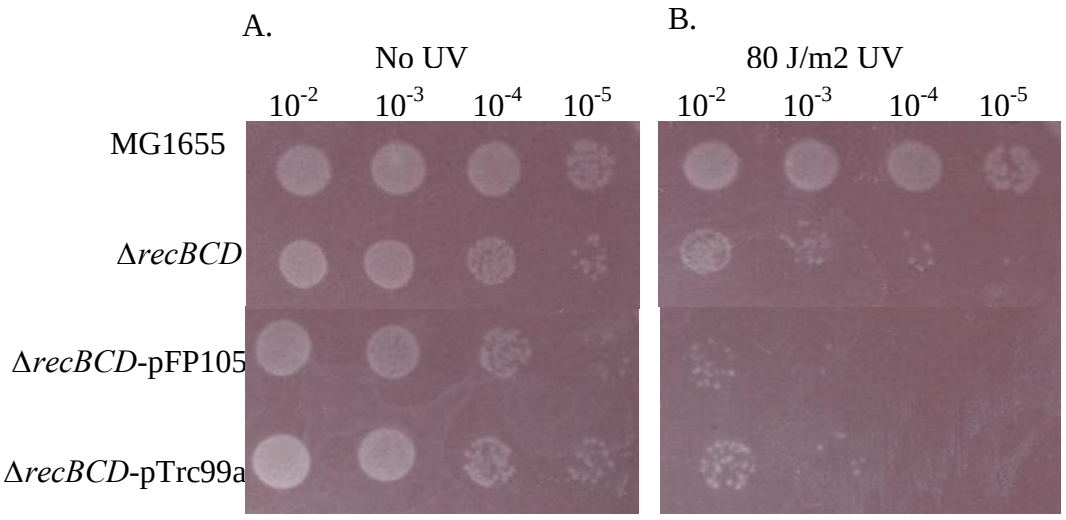


Figure 3.3.2.1.2: Plate assay for UV sensitivity of Δ*recBCD* strain expressing *B. fragilis* BF2192 and BF0679 proteins.

Stationary phase cultures of MG1655, Δ*recBCD*, Δ*recBCD* pTrc99a and Δ*recBCD* pFP105 (plasmid expressing *B. fragilis* BF2192 and BF0679 proteins) were serially diluted and 10µl of each dilution were spotted on LB plates. One set of plates was irradiated with 80 J/m^2 UV (Panel A.) while another received no UV exposure (Panel B.). All plates were incubated aerobically at 37°C for 24 hours. The Δ*recBCD*, Δ*recBCD* pTrc99a and Δ*recBCD* pFP105 strains were severely affected by UV exposure compared to the recombination proficient MG1655 strain (Panel B.). The Δ*recBCD* expressing *B. fragilis* BF2192 and BF0679 proteins (Δ*recBCD* pFP105) strain had no alleviation of UV sensitivity.

The strains were examined for differences in growth. The UV irradiated plate showed the viability of the wild type was relatively unaffected whereas the FP101 (*recBCD* null) strains showed a dramatic reduction, up to 3 $\log_{10}$ difference in number of colonies (figure 3.3.2.1.2). The Δ*recBCD* pFP105 strain expressing *B. fragilis* BF0679 and BF2192 did not show any alleviation in the UV sensitivity, and displayed a similar phenotype to the Δ*recBCD* strain (3 $\log_{10}$ reduction in viability). These results indicate that BF2192 and BF0679 are not sufficient to initiate repair of DSBs in the absence of *recBCD* in *E. coli*.

These results could also suggest that the genes were not expressed or the proteins were degraded; this was tested by analysis of the *E. coli* lysates using SDS PAGE (figure 3.3.2.1.1). Since both proteins were present, perhaps one or both of the

polypeptides were not functional in *E. coli*. Other component/components may be required for this system to function, for example they may require chaperone proteins to facilitate correct folding or directing to the site of DNA damage. The other interpretation of these results is that either one or both of the proteins are not involved in DSB repair. It is also possible that the *B. fragilis* proteins may not be capable of repairing the high level of DSBs caused by UV damage. There could be several factors affecting their efficiency, the proteins may not be expressed to produce functional proteins at an appropriate concentration or *E. coli* DNA may not contain enough *B. fragilis* specific appropriate chi sites for full functionality of DSB repair.

The UV test showed that the presence of *B. fragilis* BF2192 and BF0679 genes did not confer UV resistance to the recombination deficient Δ*recBCD* strain. To determine if these proteins could repair a lower level of damage associated with endogenous processes, the growth rates of the Δ*recBCD* pTrc99a and the Δ*recBCD* pFP105 strains were measured (figure 3.3.2.1.3 and figure 3.3.2.1.4). 5ml of LB were inoculated with a colony of the desired strain and grown overnight with shaking (160rpm) at 37°C. A 1/100 dilution was made of the overnight culture into 20ml of fresh LB, incubated at 37°C with shaking (160rpm). The OD at 600nm was measured at 20 minute time intervals until stationary phase was reached. All cultures were grown in triplicate.

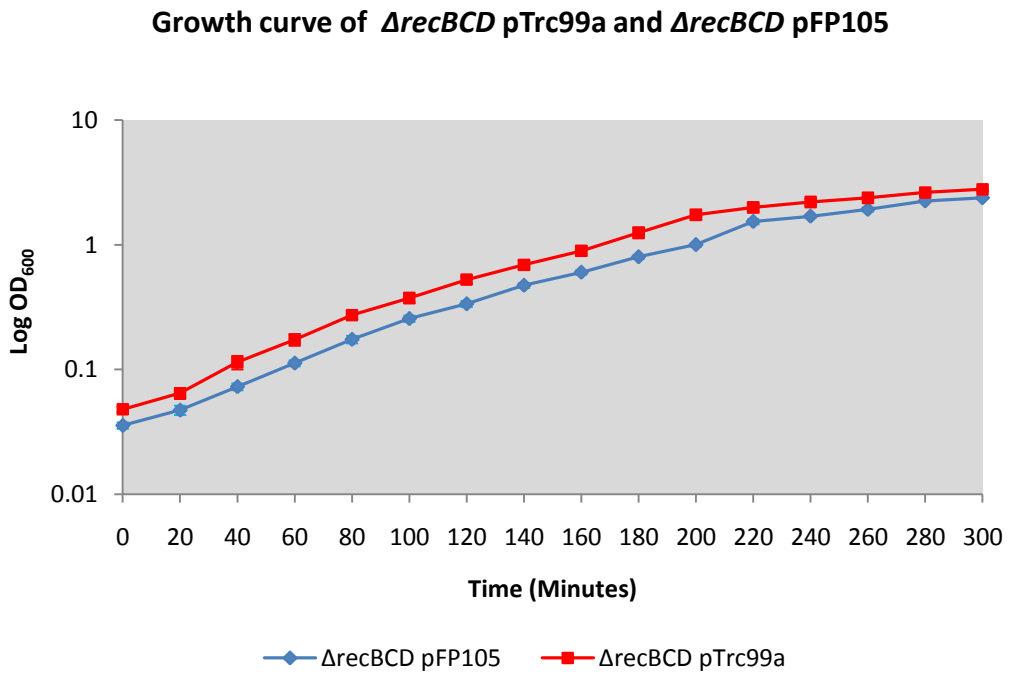


Figure 3.3.2.1.3: Growth Curves of *ΔrecBCD* pFP105 compared to *ΔrecBCD* pTrc99a

Growth curves of *ΔrecBCD* pTrc99a ■ and *ΔrecBCD* pFP105 ◆ - data points are means of triplicate cultures, standard deviation for triplicate cultures are shown by error bars. Comparison of the Δ*recBCD* strain (FP101) to the strain expressing the *B. fragilis* genes (FP101 pFP105) shows that there was no improvement of the growth defect.

The Δ*recBCD* strain expressing BF0679 and BF2192 (strain FP101 pFP105) did not grow any faster than the recombination null strain (figure 3.3.2.1.3). To determine the doubling times of the cultures, exponential phase was extended (figure 3.3.2.1.4). Cultures were prepared as previously described but when the cultures reached OD_{600nm} 0.2, a 1/10 dilution was performed and growth measurements were continued to be taken until OD_{600nm} 0.2 was reached again; as cultures were kept in exponential phase this allowed an accurate measurement of the doubling times.

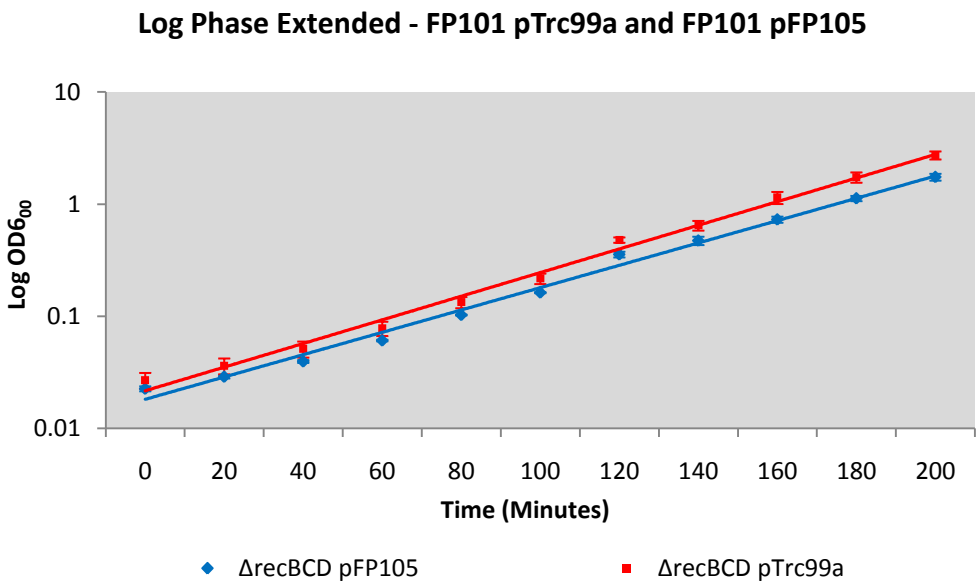


Figure 3.3.2.1.4: Extended exponential phase of Δ*recBCD* pFP105 and Δ*recBCD* pTrc99a

The doubling time of Δ*recBCD* pTrc99a ■ was 30 minutes whereas doubling time of Δ*recBCD* pFP105 ◆ strain was 32 minutes (standard deviation for triplicate cultures represented by error bars).

The expression of the *B. fragilis* genes in the Δ*recBCD* strain did not lead to an improvement in the growth defect. There is a slight difference in growth rates with the deletion strain that expresses the *fragilis* proteins having the lower growth rate. The doubling time for Δ*recBCD* pTrc99a is 30 minutes while the doubling time for Δ*recBCD* pFP105 (expressing BF0679 and BF2192) is 32 minutes, which is 6% longer. This difference may be due to expression of the proteins having a deleterious effect on other process in the cell. This result in combination with the UV irradiation tests suggests that BF0679 and BF2192 are not sufficient to initiate pre-synaptic processing of DSBs. This led to two possible interpretations either another protein was required for the *B. fragilis* pre-synaptic system or that one or both of the genes were not involved in pre-synaptic processing.

3.3.2.2 Does the presence of BF0679 rescue the phenotype of NM1035 (Δ*recD*)

To confirm if BF0679 was involved in DSB repair, BF0670 was introduced into an *E. coli* Δ *recD* strain. The NM1035 strain (Δ*recD rnhA*) was used as it has an observable phenotype of filamentous cells. BF0679 was ligated into pTrc99a giving plasmid pFP104 (figure 3.3.2.2.1). The RecD expressing plasmid was introduced into the Δ *recD* strain (NM1035). Cultures were prepared as previously described. The cultures were grown to mid-exponential phase (~OD_{600nm} 0.4) and the cell morphology was observed (figure 3.3.2.2.2).

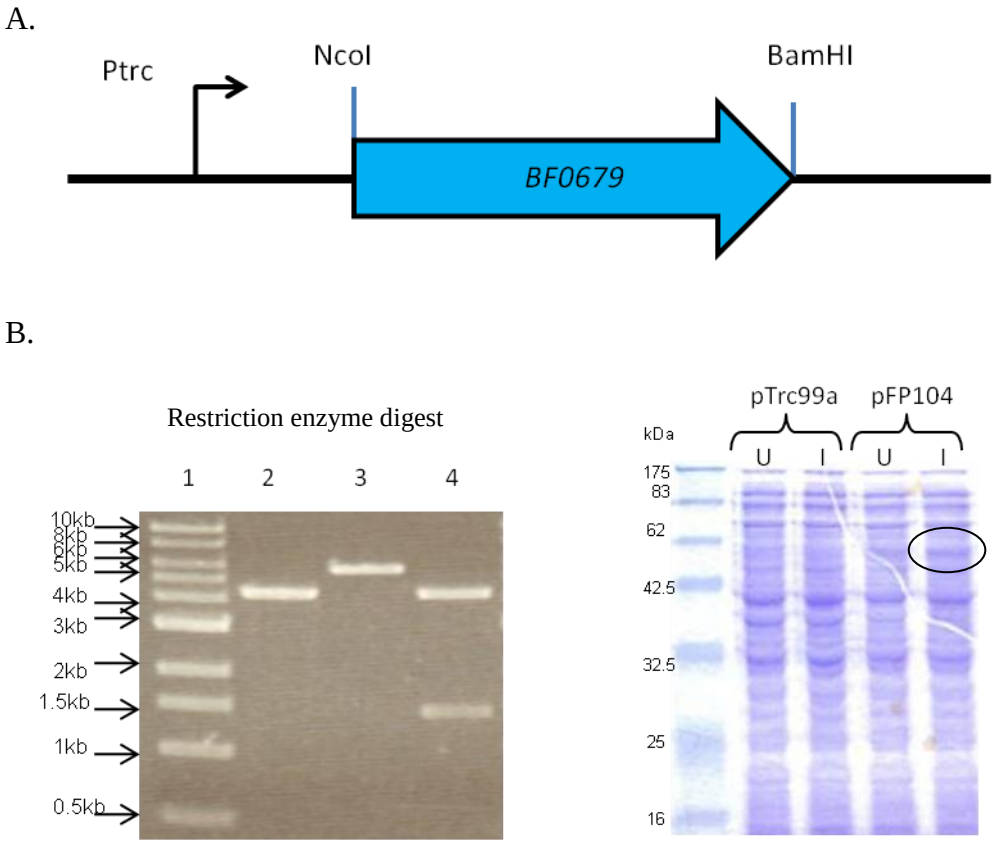


Figure 3.3.2.2.1: Production of pFP104 – BF0679 expression vector

A. The *BF0679* gene was ligated into NcoI and BamHI of pTrc99a giving plasmid pFP104. B. The restriction enzyme digests confirmed the presence of insert in the plasmid. Lane 2 – pTrc99a digested by BamHI, linearising the 4.2 kb plasmid, lane 3 – pFP104 digested with BamHI giving linearised 5.6 kb plasmid, lane 4 – pFP104 digested with BamHI/NcoI, releasing the 1.4kb insert and the 4.2kb plasmid. The right hand panel SDS PAGE gel shows a band (encircled) of the expected size for the expression product of BF0679.

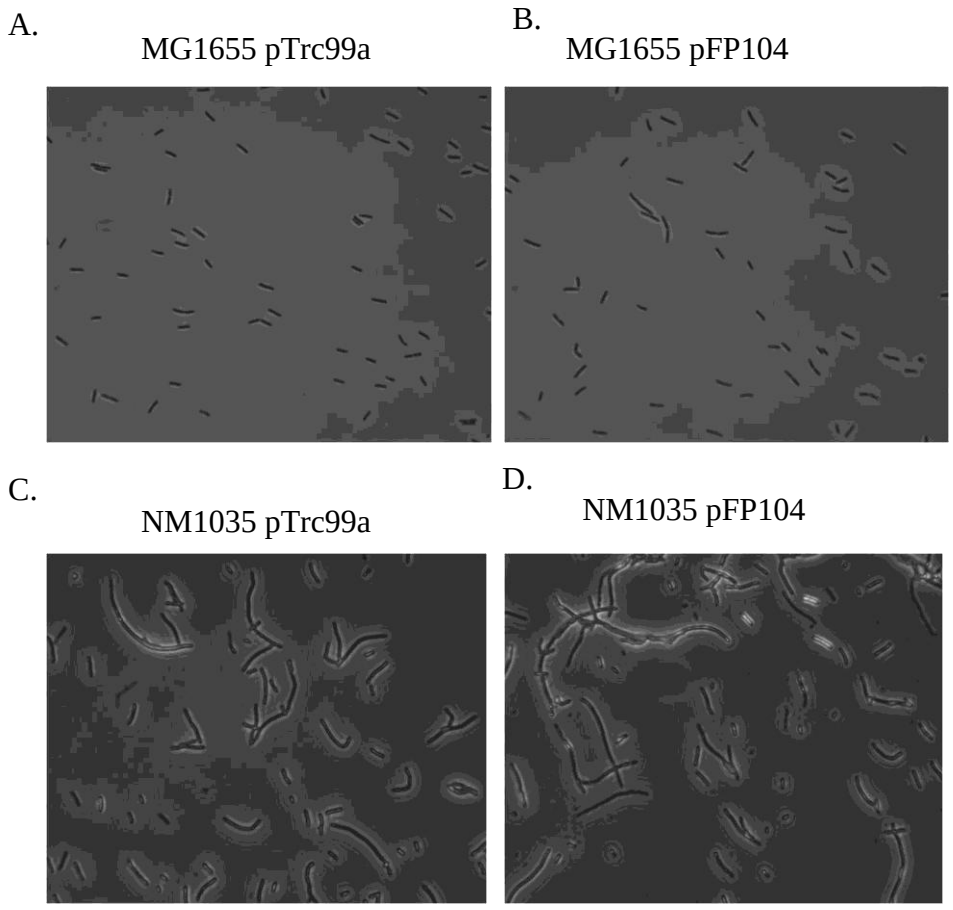


Figure 3.3.2.2.2: MG1655 and NM1035 expressing *B. fragilis* BF0679

Photomicrographs of cultures of MG1655, MG1655 pTrc99a, NM1035 pTrc99a and NM1035 pFP104 were grown to an OD of OD_{600nm} 0.4. The presence of *B. fragilis* BF0679 does not affect the morphology of MG16655 or NM1035. A and B. Wild type and wild type containing the BF0679 expression plasmid, display normal morphology. C The NM1035 strain displays a filamentous morphology. D. The presence of the BF0679 expression plasmid does not alter the filamentous morphology of the NM1035 strain.

The MG1655 strain containing *B. fragilis* BF0679 (expressed from the pFP104 plasmid) showed no difference in cell morphology compared to the wild type. The NM1035 and NM1035 harbouring the pFP104 plasmid both displayed filamentous cells, showing that the *B. fragilis* BF0679 protein product did not rescue the filamentous phenotype of the *E. coli* Δ *recD rnhA* strain (NM1035).

3.3.3.1 Identification of *B. fragilis* pre-synapsis proteins - Do BF2192 and BF2191 encode the *B. fragilis* RecBCD analogues?

Further analysis of the *B. fragilis* genome highlighted an overlapping gene downstream from the AddA encoding gene, BF2192. The BF2191 gene is 2871bp potentially producing a protein of 957 amino acids (figure 3.3.3.1.1). The gene product shared 17% identity with RexB from *L. lactis*. Many of the two component AddAB type DSB repair complexes genes are co-transcribed. Even though AddAB systems were believed to be restricted to Gram positive bacteria and the homology was low, this suggested that BF2191 and BF2192 encoded the pre-synaptic processing proteins in *B. fragilis*. To determine if BF2191 and BF2192 together had a role in DSB the genes were introduced into a Δ *recBCD E. coli* background and the growth rate and UV sensitivity measured.

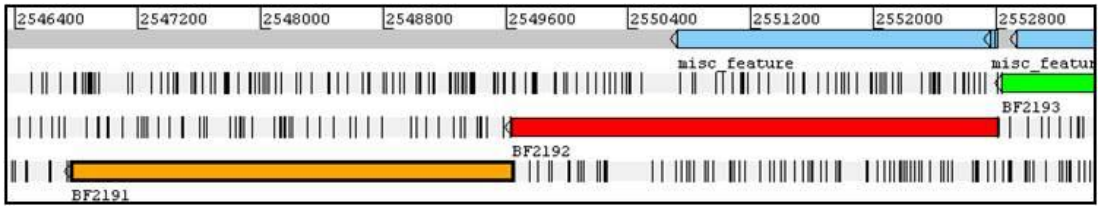


Figure 3.3.3.1.1: Artemis *B. fragilis* genes BF2191 and BF2192

Representation of possible recombination genes, BF2191 and BF2192 (taken from *B. fragilis* NCTC 9343 Artemis annotation, Cerdeno-Tarraga *et al*., 2005). The highlighted gene BF2191 overlaps the BF2192 gene.

3.3.3.2 Does the presence of BF2192 and BF2191 rescue the Δ*recBCD* UV phenotype?

The *B. fragilis* NCTC9343 BF2191 and BF2192 genes were cloned into the *E. coli* expression vector pTrc99a giving plasmid pFP101 (figure 3.3.3.2.1). The Δ *recBCD E. coli* strain was transformed with pFP101 and tested for UV sensitivity/resistance (figure 3.3.3.2.2). Overnight stationary phase cultures were serially diluted in bacterial buffer, then 10µl of each dilution were spotted on replica LB plates, one set of plates was exposed to 80J/m^2 UV while the control set was not. All plates were

kept in the dark and aerobically incubated overnight at 37°C. All experiments were performed in triplicate and a representative example is shown (Figure 3.3.2.2).

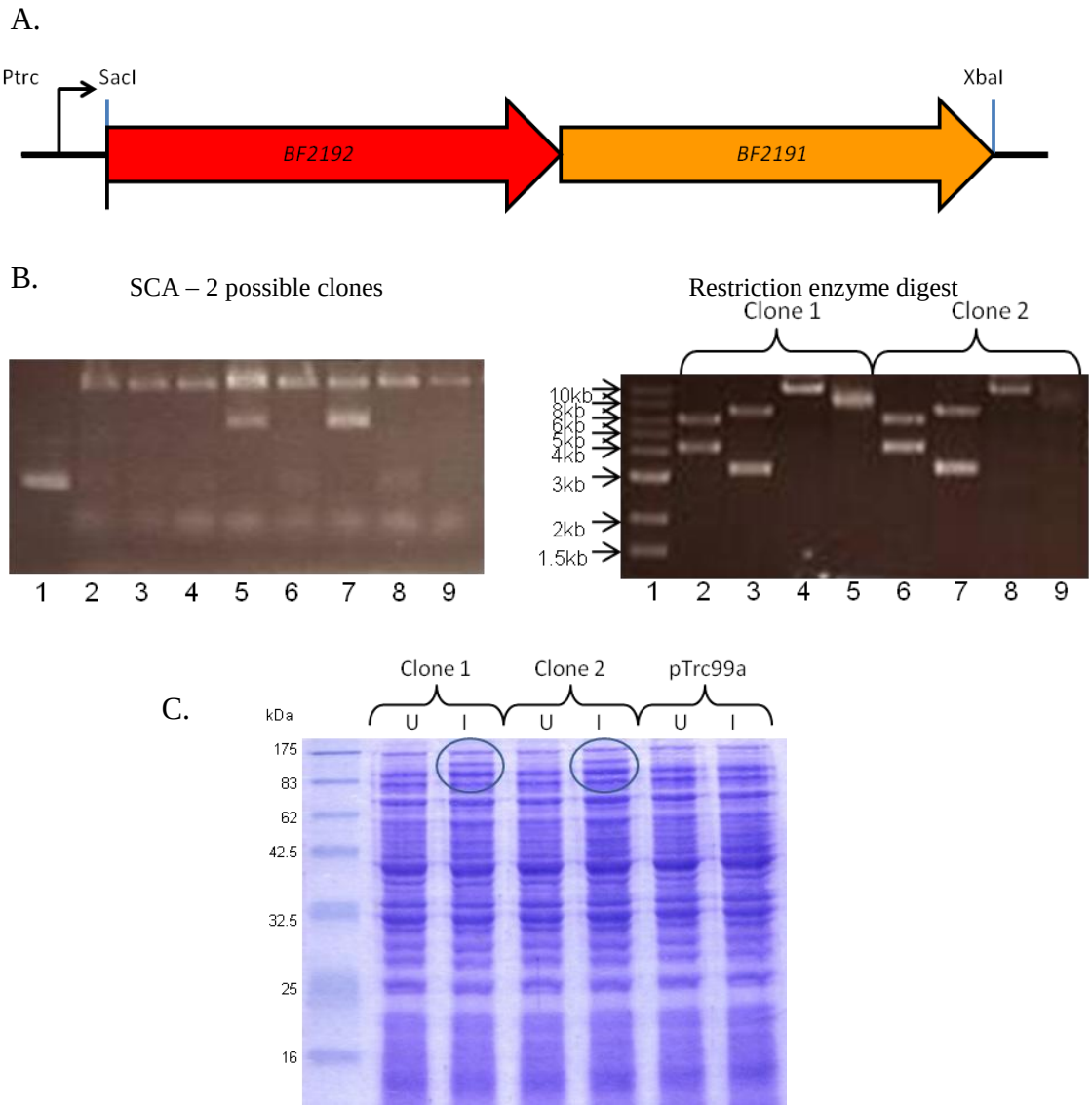


Figure 3.3.3.2.1: Production of pFP101 – BF2192 and BF2191 expression vector

A. BF2191 and BF2192 were amplified together by PCR, the genes were then ligated into SacI and XbaI sites of pTrc99a giving plasmid pFP104. B. The single colony screen shows two plasmids running higher (lanes 5 and 7) than the control plasmid (lane1), suggesting possible inserts. The restriction enzyme digests confirmed the presence of insert in the plasmids. Lanes 2 and 6 – plasmids digested by XbaI/SacI giving two fragments of 4.2 and 6kb, Lanes 3 and 7 – plasmids digested by NdeI giving two fragments of 3 and 7kb. Lanes 4 and 8 – plasmids digested by SacII linearising the plasmid by cutting the *addAB* insert, lanes 5 and 9 uncut plasmid. C. SDS PAGE gel showing protein banding from two clones. Highlighted by circles are band of appropriate size for expression products.

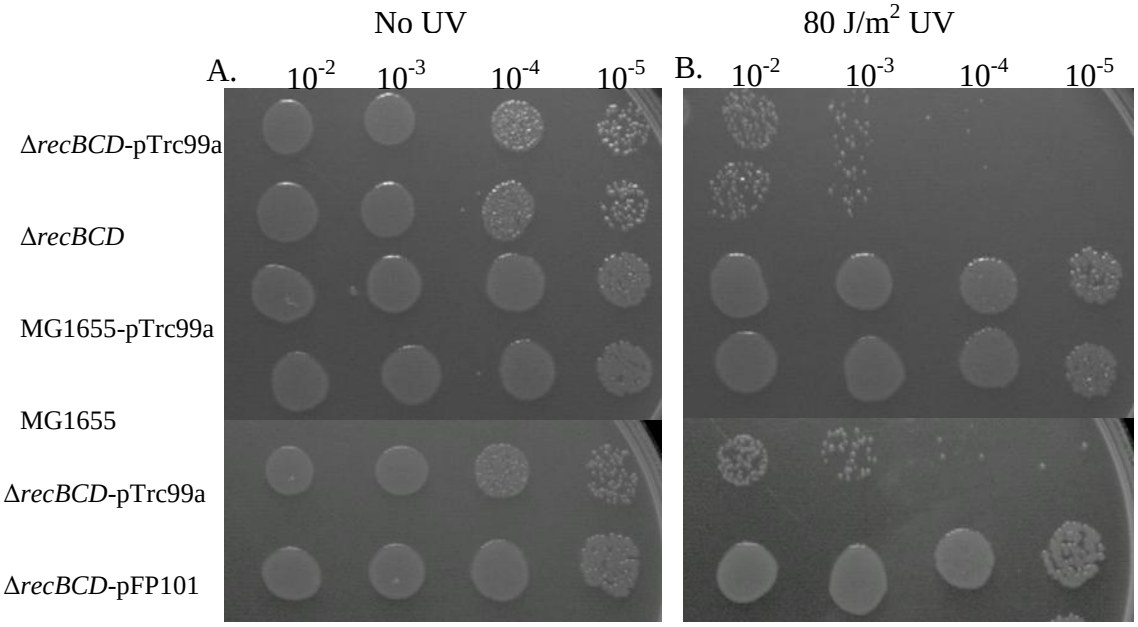


Figure 3.3.3.2.2: plate assay for UV sensitivity of Δ *recBCD* strain expressing *B. fragilis* BF2191 and 2192

A. Stationary phase cultures of MG1655, MG1655 pTrc99a, Δ*recBCD*, Δ*recBCD* pTrc99a and Δ*recBCD* pFP101 were serially diluted; 10µl of each dilution were spotted on LB plates and grown aerobically at 37°C for 24 hours. B. A replica plate was produced; this was irradiated with 80 J/m^2 UV then grown aerobically at 37°C for 24 hours. Comparison of the UV exposed plates (panel B) to the non UV exposed plates (panel A) shows that UV irradiation had a significant effect on viability of the Δ *recBCD* strain, while the MG1655 wild type was unaffected. The Δ *recBCD* strain expressing the *B. fragilis* genes (Δ*recBCD*-pFP101) displayed resistance to UV irradiation similar to wild type.

In keeping with the earlier experiment in section 3.3.2.1, under the control conditions (no UV exposure) the Δ *recBCD* and Δ *recBCD* pTrc99a strains had a lower viability than the wild type MG1655 (about 1 $\log_{10}$ difference). The plate exposed to UV radiation showed that the Δ *recBCD* pFP101 and Δ *recBCD* pTrc99a strains had a 3 $\log_{10}$ reduction in viability compared to the wild type MG1655, which was unaffected. The presence of the pTrc99a plasmid has no effect on the viability of the cells in either the MG1655 or the Δ *recBCD* strains under both conditions. Figure 3.3.3.2.2 shows that the Δ *recBCD* pFP101 strain expressing *B. fragilis* BF2191 and BF2192 was unaffected by the UV exposure and viability is on a par with wild type MG1655. These results showed that the *recBCD* null strain containing a plasmid expressing both of the *B. fragilis* genes rescued the UV sensitivity. This

demonstrated one or both of these genes had a role in DSB repair and were functional in *E. coli*.

3.3.3.3 Does the presence of BF2192 and BF2191 rescue the Δ *recBCD* growth defect?

In keeping with previous methodology, the effect of the *B. fragilis* BF2191 and BF2192 genes on growth (figure 3.3.3.3.1) was measured. 5ml of fresh LB were inoculated with a colony of the desired strain and grown overnight with shaking (160rpm) at 37°C. A 1/100 dilution was made of the overnight culture into 20ml of fresh LB in a 250ml flask and incubated at 37°C with shaking (160rpm). The OD at 600nm was measured for cultures at 20 minute intervals until stationary phase was reached. To measure doubling time, exponential phase was extended to ensure all cells were growing exponentially by diluting cultures 1/10 into fresh medium when an OD_{600} of 0.2 was reached. All experiments were performed in triplicate.

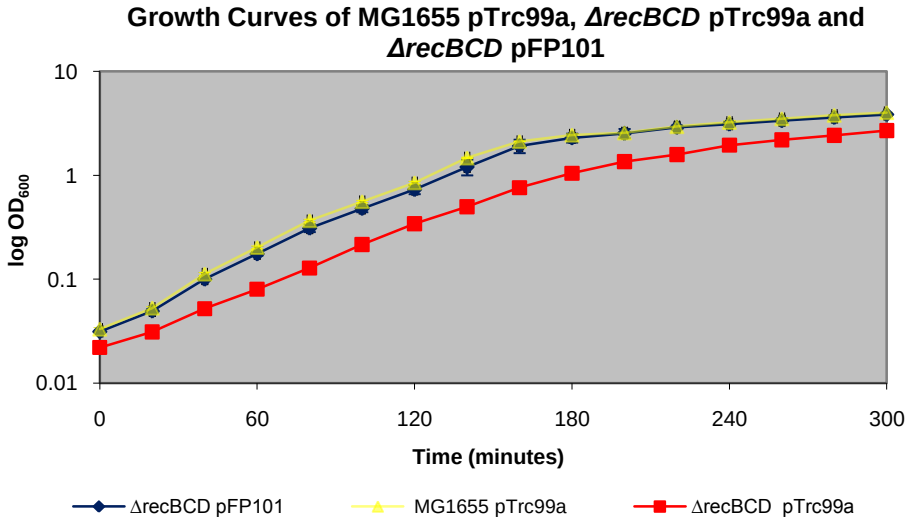


Figure 3.3.3.3.1: Growth of MG1655 pTrc99a, Δ*recBCD* pTrc99a, Δ*recBCD* pFP101.

Comparison of growth rates between MG1655 pTrc99a ▲, Δ*recBCD* pTrc99a ■ and Δ*recBCD* pFP101 ◆ - grown aerobically in LB at 37°C. The standard deviation for triplicate cultures is represented by error bars. Samples were taken every 20 minutes and absorbance measured at OD 600nm. MG1655 pTrc99a and Δ*recBCD* pFP101 cultures had comparable growth rates whereas the Δ*recBCD* pTrc99a strain had a slower growth rate than the other strains. MG1655 pTrc99a had a

doubling time of 23 minutes. Δ*recBCD* pTrc99a had a doubling time of 31 minutes. Δ*recBCD* pFP101 had a doubling time of 24 minutes.

There was a significant difference in growth rate of the Δ*recBCD* strain when the *B. fragilis* genes were expressed. MG1655 pTrc99a and Δ*recBCD* pFP101 cultures had comparable growth rates, showing that presence of the *B. fragilis* genes rescued the growth defect of the Δ*recBCD* strain. Doubling time for Δ*recBCD* pFP101 was 24 minutes which was similar to that of the wild type MG1655 at 23 minutes. In comparison, the Δ*recBCD* pTrc99a strain had a doubling time of 31 minutes, significantly slower than the other two strains. Overall the growth of the Δ*recBCD* strain was 33% slower than MG1655, and is similar to results from section 3.3.3. These results show that the presence of BF2192 and BF2191 rescued the growth defect of the Δ *recBCD E. coli* strain FP101. As with the previous experiment (section 3.3.3.2) this shows that the *B. fragilis* machinery can function in *E. coli* and implies a role in DSB repair for the proteins encoded by BF2191 and BF2192.

3.3.3.4 Identification of *B. fragilis* pre-synapsis proteins - Are both BF2192 and BF2191 products required for DSB repair?

To determine if both proteins were required, or whether either could act on its own for DSB repair, each protein was expressed separately in recombination deficient *E. coli* strains. Two plasmids were produced; BF2191 was cloned into pTrc99a giving pFP103, and BF2192 was cloned into pTrc99a giving pFP102. FP101 strain (Δ *recBCD*) was transformed with plasmids expressing either BF2192 (pFP102) or BF2191 (pFP103) (figure 3.3.3.4.1 and 3.3.3.4.2). Both strains were tested for rescue of the UV sensitivity and growth defects (figure 3.3.3.4.1.1 and figure 3.3.3.4.2.1).

A.

B.

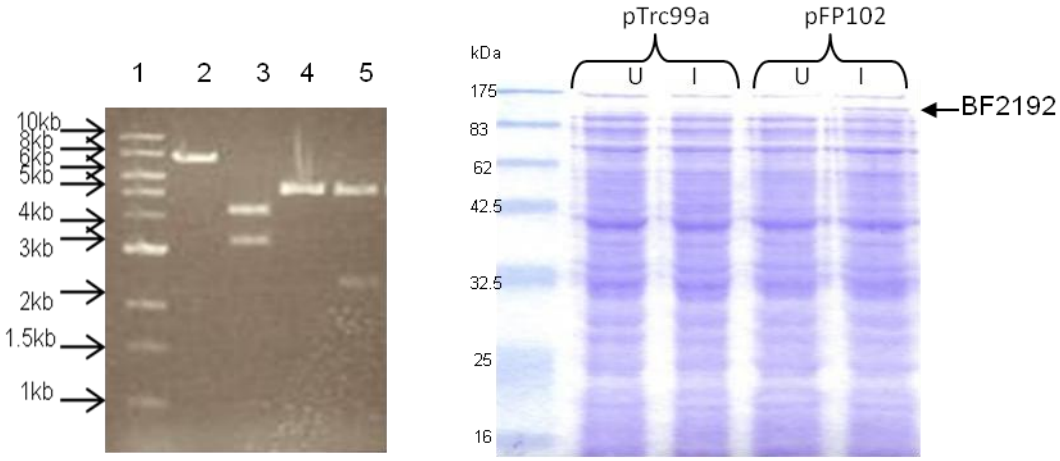
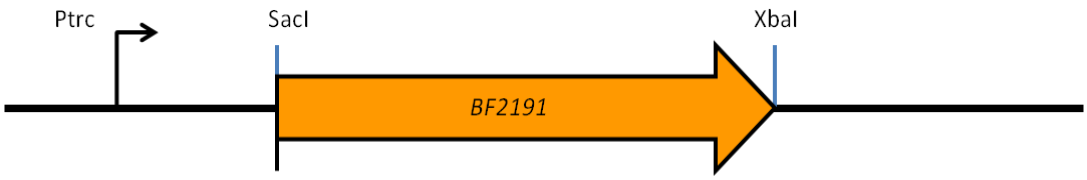


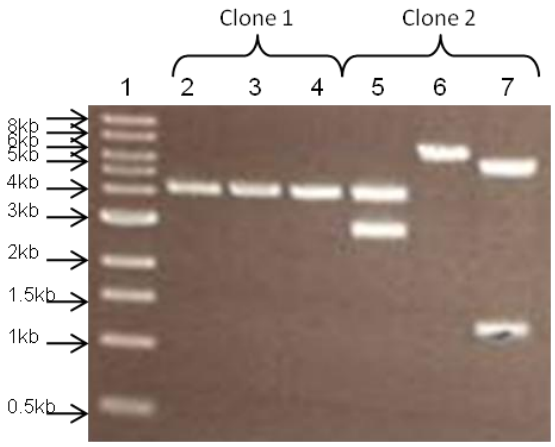
Figure 3.3.3.4.1: Production of pFP102 – BF2192 expression vector

A. The *BF2192* gene was amplified by PCR and ligated into SacI and XbaI sites of pTrc99a giving plasmid pFP102. B. Restriction enzyme digests confirmed the presence of insert in the plasmid; Lane 2 – pFP102 digested by XbaI, linearising the 7.4 kb plasmid, lane 3 – pFP102 digested by XbaI /SacI, releasing the 3.2kb insert and the 4.2kb plasmid, lane 4 – pFP102, lane 5 – pFP102 digested by NcoI, giving two fragments of 5.1 and 2.3 kb. Expression of BF2192 was confirmed by SDS PAGE gel.

A.



B.



C.

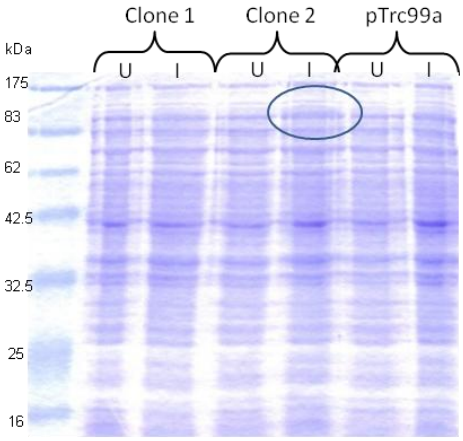


Figure 3.3.3.4.2: Production of pFP103 – BF2191 expression vector

A. The *BF2191* gene was amplified by PCR and ligated into SacI and XbaI sites of pTrc99a giving plasmid pFP103. B. Two possible clones were tested for insert by restriction enzyme digests; only clone 2 contained the insert. Lane 5 – pFP105 digested with SacI/XbaI, releasing the 2.8kb insert and the 4.2kb plasmid, lane 6 – pFP104 digested by SacI, linearising the 7 kb plasmid, lane 7 – pFP103 digested with HindIII giving two fragments of 1.4 and 5.6kb C. Expression of BF2191 was confirmed by SDS PAGE gel (band of appropriate size highlighted by circle).

3.3.3.4.1 Does the presence of BF2192 or BF2191 rescue the Δ *recBCD* UV phenotype?

Stationary phase cultures were serially diluted and 10µl of each dilution were spotted on LB plates. One set of plates was exposed to 80J/m^2 UV while the control set was not, followed by an overnight incubation at 37°C. All experiments were performed in triplicate and a representative example is shown (figure 3.3.3.4.1.1)

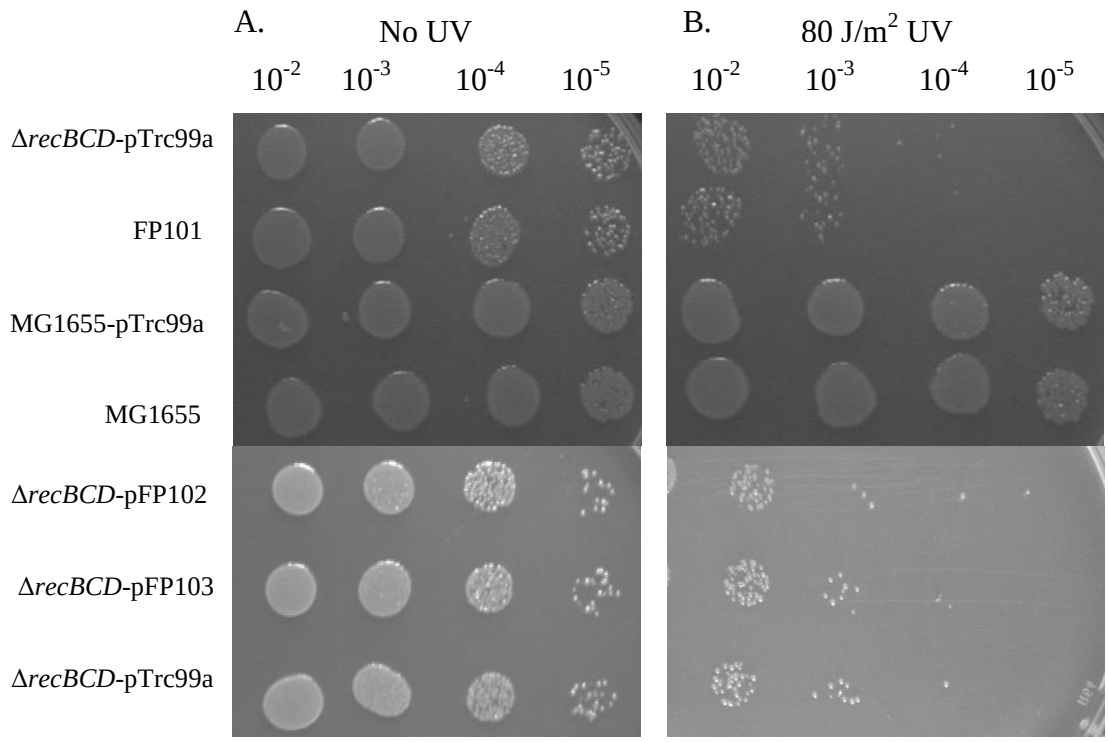


Figure 3.3.3.4.1.1: plate assay for UV sensitivity of Δ*recBCD* pFP103 and Δ*recBCD* pFP102

A. Stationary phase cultures of MG1655, MG1655-pTrc99a, Δ*recBCD*, Δ*recBCD*-pTrc99a, Δ*recBCD*-pFP102 and Δ*recBCD*-pFP103 were serially diluted; 10µl of each dilution were spotted on LB plates and grown aerobically at 37°C for 24 hours. B. A replica plate was produced; this was irradiated with 80 J/m^2 UV then grown aerobically at 37°C for 24 hours. Panel B shows that UV irradiation had a significant effect on viability of the Δ*recBCD* strains, while the MG1655 wild type was unaffected. Strains expressing the *B. fragilis* proteins separately did not show any alleviation of UV sensitivity.

In figure 3.3.5.5.1a the *recBCD* null strain was sensitive to UV radiation, with a 3 $\log_{10}$ reduction in viability. Both of the *recBCD* null strains that contained plasmids

expressing either BF2192 or BF2191 showed no rescue of the UV sensitivity. This demonstrated that the protein products of either of the genes on their own could not rescue UV sensitivity.

3.3.3.4.2 Does the presence of BF2192 or BF2191 rescue the Δ*recBCD* growth defect?

To determine if the individual proteins could repair endogenous DNA damage, such as stalled replication forks, their effect on growth rate was measured (figure 3.3.3.4.1.2). The experiment was performed as previously described.

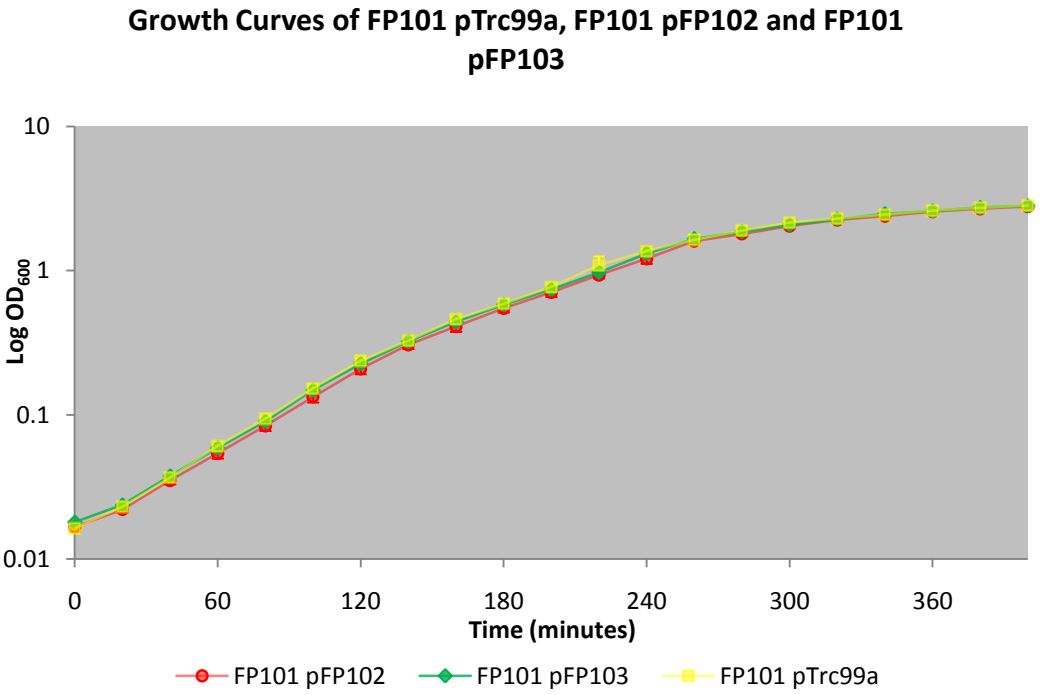


Figure 3.3.3.4.1.2: Growth of Δ*recBCD* pTrc99a, Δ*recBCD* pFP102 and Δ*recBCD* pFP103.

Comparison of growth rates for Δ*recBCD* pTrc99a ■, Δ*recBCD* pFP102 ●and Δ*recBCD* pFP103◆ - standard deviation for triplicate cultures represented by error bars. The cultures were grown aerobically in LB at 37°C and samples were taken every 20 minutes and absorbance measured at OD 600nm. All cultures had comparable growth rates, showing that presence of either BF2191 or BF2192 did not rescue the growth defect of the Δ*recBCD* strain. Δ*recBCD* pFP102 had a doubling time of 31 minutes, Δ*recBCD* pFP103 had a doubling time of 31 minutes and Δ*recBCD* pTrc99a had a doubling time of 30 minutes.

All strains had comparable growth rates with similar doubling times of around 30 minutes. The presence of BF2192 gene in the recombination deficient Δ *recBCD* strain did not improve the growth defect. Equally the introduction of BF2191 did not improve the growth defect of the strain. From experiments in section 3.3.5.3 it was shown that the presence of both genes alleviates the Δ*recBCD* growth defect, therefore it was concluded that both proteins were required for DSB repair. Collating the results of all the experiments in section 3.3, shows that both *B. fragilis* genes are required for DSB repair. From this point forward BF2191 will be referred to as *addB* and BF2192 will be referred to as *addA*.

3.4 Generating precise deletions in *B. fragilis* NCTC9343

Previously all experimental procedures had been carried out in *E. coli*, as at that time in the project *B. fragilis* NCTC9343 was genetically intractable. Other work within the laboratory had developed a procedure that allowed genetic manipulation of *B. fragilis* NCTC9343 (Patrick *et al*., 2009). The *E. coli* complementation experiments had identified the *B. fragilis* pre-synapsis genes; the next approach was to produce a *B. fragilis addAB* deletion using the newly devised deletion method and determine if it was deficient in recombination.

3.4.1 Deletion method for *B. fragilis addAB*

A schematic for generating a marked deletion in *B. fragilis* NCTC9343 is shown in figure 3.4.1.1. DNA sequences consisting of 500bp upstream and downstream of *B. fragilis addAB* genes were amplified by PCR using primers that contained tail regions that were homologous to either end of an erythromycin gene (*ermF*). Crossover PCR allowed the upstream and downstream *addAB* sequences to be joined to either side of the *ermF* gene: producing a deletion cassette. The deletion cassette was cloned into a suicide vector (pGB909) that contained an I-SceI recognition site. The suicide vector was conjugated from *E. coli* into *B. fragilis*; colonies that were erythromycin resistant were selected as these cells contained the plasmid integrated

into the chromosome. A plasmid (pGB920) that expressed the I-SceI meganuclease (inducible by fucose) was introduced into the *B. fragilis* single-crossover strains by electroporation, with selection for resistance to tetracycline. The selected colonies were streaked on defined medium containing tetracycline and fucose to allow induction of the I-SceI Meganuclease, thus producing a DNA double-strand break on the chromosome. Repair of the break would involve homologous pairing of DNA regions (either the upstream or downstream region) effectively removing the gene completely. Colonies were screened for erythromycin resistance; chromosomal DNA was extracted from erythromycin resistant colonies and screened by PCR for the presence of the gene or deletion cassette.

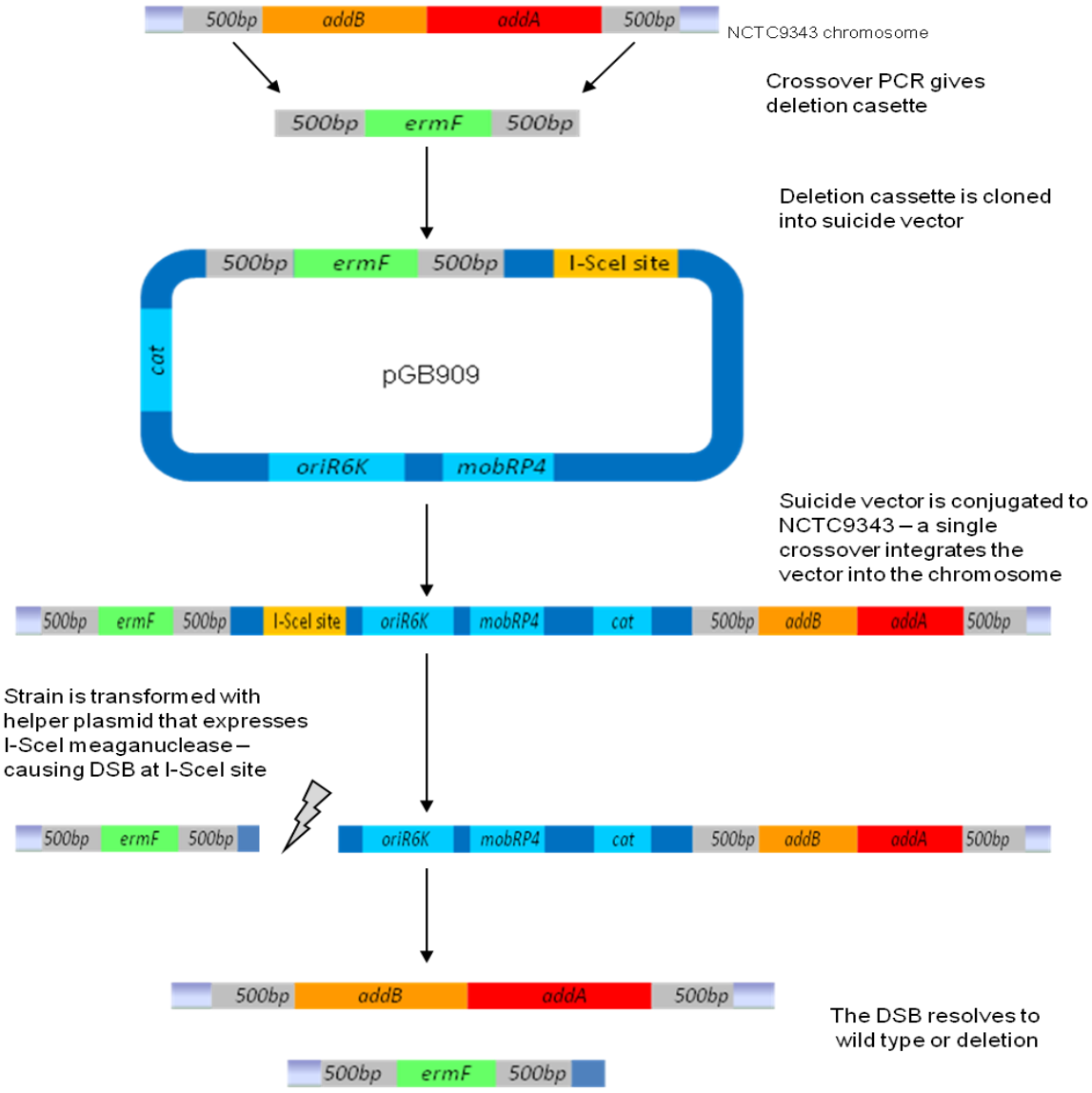


Figure 3.4.1.1: Representation of NCTC9343 marked deletion method

3.4.2 Identification of recombination deficient *B. fragilis*

Producing the *addAB* deletion strain proved problematic. After inducing the double-strand break erythromycin resistance colonies were selected; colony PCRs or chromosomal PCRs were then undertaken to detect the deletion cassette. A large number of colonies screened by PCR showed the deletion cassette and wild type *addAB* genes or only the wild type. This may have been due to several reasons; 1. Deletion of the recombination system in cells may be detrimental or produce non-viable cells; 2. There may have been recombination with the endogenous plasmid or at other points on the chromosome allowing the erythromycin gene to be integrated without removal of the target genes. One *addAB* deletion mutant was obtained. The *addAB* null strain was named FP201. The following sections characterize the mutant in relation to DSB repair and DNA damaging agents.

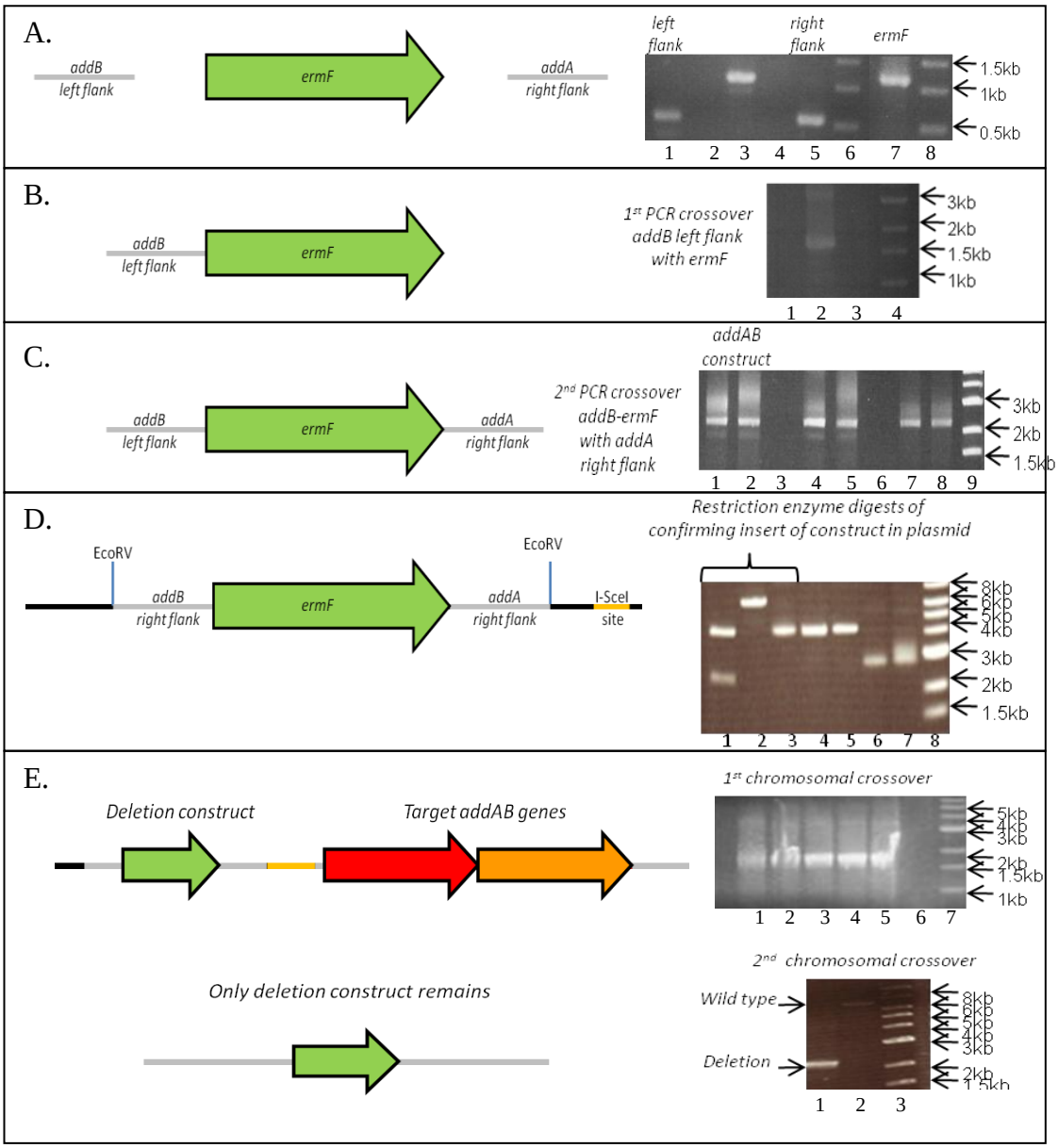


Figure 3.4.2.1: Production of deletion plasmid (pFP201) and *ΔaddAB* strain

A. The regions immediately to the left and right of the *addAB* operon and the *ermF* gene (agarose gel lanes 1, 5 and 7 respectively) were amplified using PCR. B. The left flank of *addB* was fused to the *ermF* gene by crossover PCR (agarose gel lane 2). 3 A second crossover PCR fused the *addA* right flank to the left flank/*ermF* fusion producing the deletion cassette (agarose gel lane 1 and 2). D. The cassette was ligated into the EcoRV site of pGB909. Presence of the insert was confirmed by restriction enzyme digest; lane 1- releasing fragments, lane 2- pFP201, lane3 linearised pFP201, lane 7- pGB909. E. The 1st chromosomal crossover integrates pFP201 into the chromosome (1st chromosomal crossover agarose gel – lanes 1, 2, 3, 4 and 5). The 2nd chromosomal crossover removes the *addAB* genes leaving only the deletion cassette (2nd chromosomal crossover agarose gel lane 1 shows the deletion, lane 2 is the wild type control).

3.4.2.1 UV Sensitivity of Δ *addAB B. fragilis*

To determine the phenotype of the *B. fragilis addAB* null strain, it was compared to NCTC9343 for viability after UV irradiation. Stationary phase cultures were serially diluted, then 10µl were spotted on a series of BHI-S plates. One set of plates were exposed to UV at varying intensities, while the other control set of plates were not. All plates were kept in the dark and incubated anaerobically for 1-2 days. All experiments were performed in triplicate and a representative example is shown (figure 3.4.2.1.1)

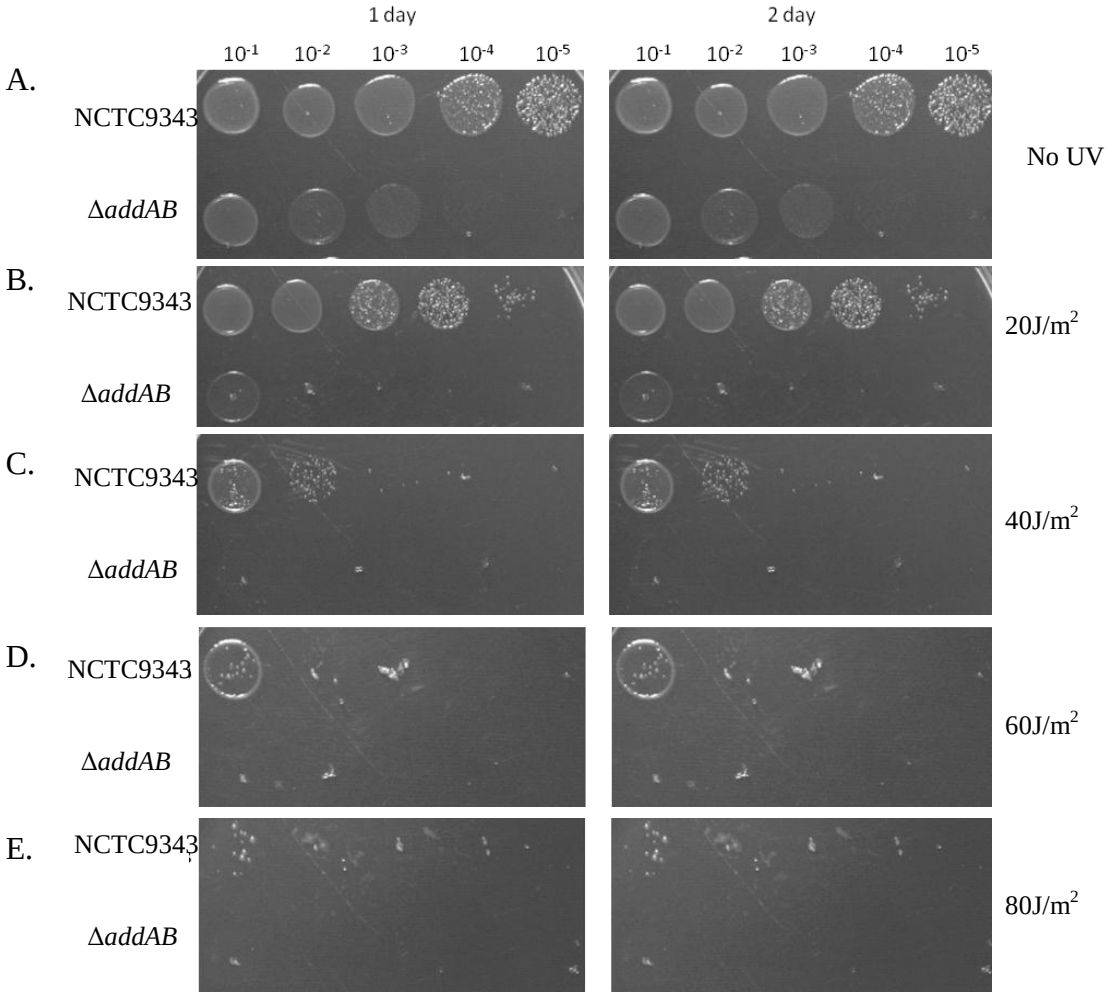


Figure 3.4.2.1.1: Assay for UV sensitivity of NCTC9343 and Δ*addAB* strain

Stationary phase cultures of NCTC9343and Δ*addAB* (strain FP201) were serially diluted and 10µl of each dilution were spotted on BHI-S plates and grown anaerobically at 37°C for 24 - 48 hours. Replica plates were made and exposed to various levels of UV; 20, 40**,** 60 and 80 J/m^2 UV irradiation. These plates were kept in the dark and grown anaerobically at 37°C for 24 - 48 hours. The results show FP201 (Δ *addAB*) is more sensitive to UV than wild type. When there is no UV exposure (A.) the Δ*addAB* has an inherent growth defect compared to wild type. At 20J/m^2 (B.) the Δ*addAB* is severely affected compared to wild type. At 40J/m^2 (C.) and above (D. E.) the recombination deficient strain shows no growth at all levels of dilution. The wild type showed better resistance to UV exposure than the recombination deficient strain.

In the absence of UV irradiation; strain FP201 (Δ*addAB*) had an inherent growth defect as its colony forming ability was 2 $\log_{10}$ lower than the NCTC9343 strain. Comparing this result to section 3.3.2 where the *E. coli* Δ *recBCD* strain (FP101) had only a small reduction in viability; the effect of deleting *addAB* in *B. fragilis*

appeared to be more severe than deleting *recBCD* in *E. coli*. Figure 3.4.4.1b shows the effect of 20J/m^2 UV irradiation on both strains; at this level the wild type had a 1 log_{10} drop in viability compared to the no UV control whereas the Δ *addAB* strain had a 2 log_{10} reduction in viability compared to the no UV control. At 40J/m^2 NCTC9343 there was a 3 log_{10} reduction compared to the no UV control and the Δ *addAB* (FP201) strain had a 4 log_{10} or more reduction compared to the no UV control. UV doses above 40J/m^2 are only informative for the wild type strain as the Δ *addAB* strain was completely killed at all dilutions.

It is interesting to note that NCTC9343 appeared to be more sensitive to UV than *E. coli* MG1655. A UV dose of 80J/m2 had hardly any affect on the MG1655 strain but *B. fragilis* wild type NCTC9343 had a significantly reduced viability (4-5 log_{10}). The results indicated that *B. fragilis* NCTC9343 was more sensitive to UV damage than *E. coli* MG1655 and that deletion of the *B. fragilis* recombination genes had a more severe effect than the deletion of the *E. coli* recombination genes. This confirms a role for the *B. fragilis addAB* genes in DSB repair in response to UV damage.

3.4.2.2 Growth of NCTC9343 and FP201 (Δ *addAB*)

It is expected that a recombination deficient strain will have a slower growth rate due to cell death following the inability to repair DSBs. The growth of the wild type strain NCTC9343 was compared to the *addAB* deletion strain (figure 3.4.2.2.1 and 3.4.2.2.2). A colony of the appropriate strain was used to inoculate 5ml of BHI-S; this was incubated overnight at 37°C anaerobically. The stationary phase cultures were then diluted 1/100 into 20ml of fresh pre-reduced BHI-S and incubated anaerobically at 37°C. The OD at 600nm was measured for cultures at 45 minute intervals until stationary phase was reached. All cultures were grown in triplicate.

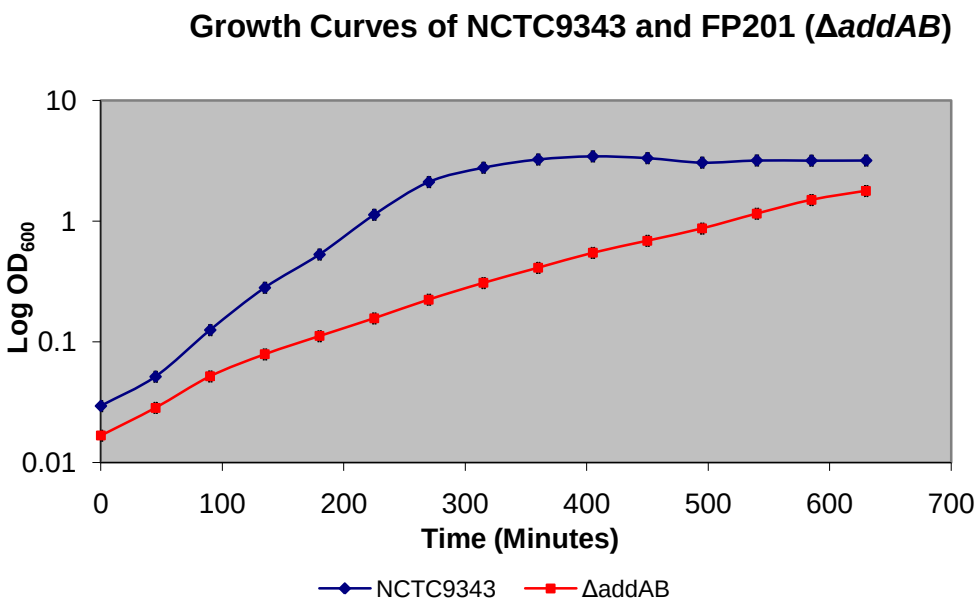


Figure 3.4.2.2.1: Growth curves of *B. fragilis* strains

Comparison of growth rates between NCTC9343 ◆and FP201 (Δ*addAB*) ■, grown anaerobically in BHI-S at 37°C. Samples were taken every 45 minutes and absorbance measured at OD 600nm. The Δ *addAB* strain had a slower growth rate than the NCTC9343 wild type. The standard deviation of triplicate cultures is represented by error bars.

There was a marked difference in the growth rates of wild type NCTC9343 and the *addAB* deletion strain (FP201). The FP201 Δ*addAB* strain had a slower growth rate than NCTC9343. NCTC9343 reached stationary phase considerably earlier than the FP201 *addAB* null strain, by 200 minutes. The curves at exponential phase appear to be biphasic; NCTC9343 showed a decrease at 45 minutes while the FP201 growth slowed at 90 minutes, and may suggest depletion of a limiting substrate. Due to the biphasic nature of the growth curves, doubling times were determined during extended log growth.

To measure the doubling times at low cell densities, the cultures were kept below OD_{600} of 0.2. The experiment was performed as previously described in this section but with two alterations; time intervals were 30 minutes and cultures were diluted 1/10 into fresh medium when an OD_{600} of 0.2 was reached, to ensure all cells were growing exponentially. All cultures were grown in triplicate.

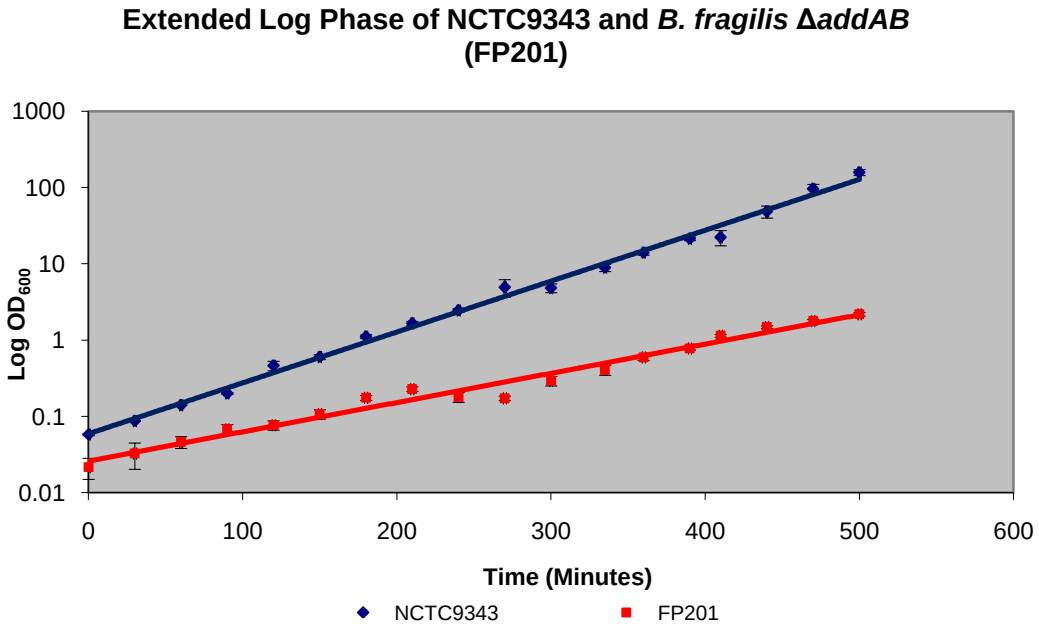


Figure 3.4.2.2.2: Extended exponential phase of *B. fragilis* NCTC9343 and Δ*addAB* strain (FP201)

Graph showing the extended exponential phase of NCTC9343 ◆ and the Δ*addAB* strain (FP201) ■ - standard deviation for triplicate cultures represented by error bars. Cultures were grown anaerobically in BHI-S at 37°C and samples were taken every 30 minutes and absorbance measured at OD 600nm. ODs were multiplied by the dilution factor. Doubling time of NCTC9343 was 44 minutes whereas the doubling time of *addAB* null strain was 75 minutes.

The extended exponential phase allowed the doubling times of each strain to be measured at lower cell densities. NCTC9343 had a doubling time of 44 minutes, in contrast to a doubling time of 75 minutes for the Δ *addAB* strain. It is interesting to compare this to the *E. coli recBCD* deletion strain where growth was 30% slower than the MG1655 wild type. This result supports the role of the *addAB* genes in normal growth of cells and shows how important the repair of DSBs is to replication of the chromosome.

3.4.2.3 Is the Δ *addAB* strain sensitive to metronidazole?

At present, the drug most effective for treating *B. fragilis* infections is metronidazole. The drug is believed to cause DNA DSBs – too many DSBs will kill even recombination proficient strains. This hypothesis was supported by the demonstration that repair deficient *E. coli* strains showed increased sensitivity to metronidazole when grown anaerobically compared to the wild type parent strain (Jackson *et al.*, 1984). Since the *B. fragilis* Δ *addAB* strain was sensitive to UV one would propose that a *B. fragilis* recombination deficient strain would be significantly affected by metronidazole compared to a recombination proficient strain.

To test the sensitivity of the Δ *addAB* strain to DNA damaging agents, cells were grown in the presence of metronidazole. Stationary phase cultures were serially diluted and 10µl of each dilution were spotted on a series of BHI-S plates that contained different concentrations of metronidazole. The plates were incubated anaerobically at 37°C for 48 hours. All experiments were performed in triplicate and a typical example is shown (figure 3.4.2.3.1).

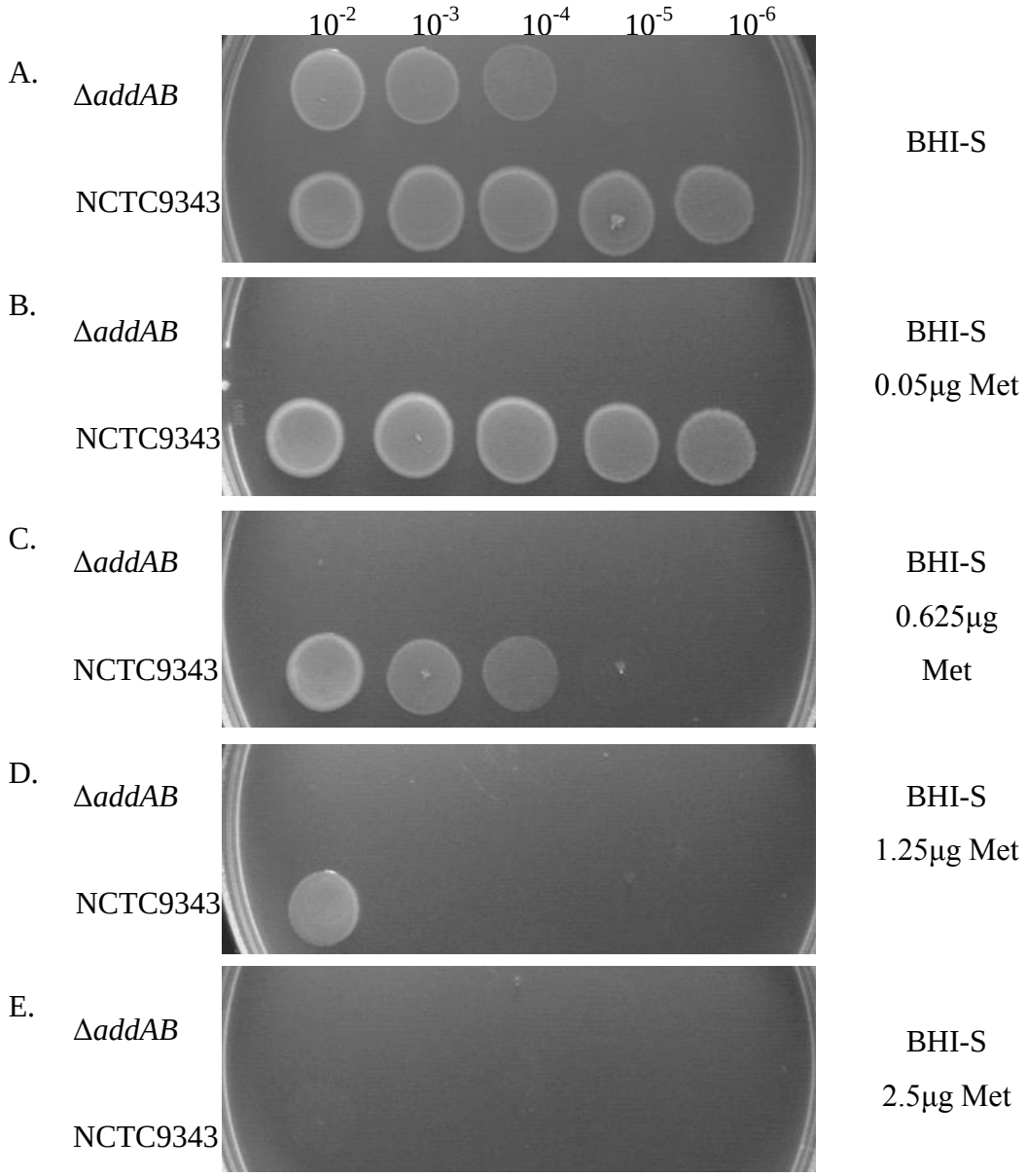


Figure 3.4.2.3.1: Assay for metronidazole sensitivity of NCTC9343 and Δ*addAB*

Stationary phase cultures of NCTC9343 and FP201 were serially diluted and 10µl of each dilution were spotted on BHI-S and grown at 37°C for 48 hours. Replica plates were made that contained different concentrations of metronidazole (0 (A.), 0.05(B.), 0.625(C.), 1.25(D.) and 2.5µg/ml (E.)). The Δ *addAB* was more sensitive to metronidazole than the wild type at all levels of metronidazole. A. With only BHIS the Δ *addAB* displays an inherent growth defect compared to wild type. B. In the presence of 0.05µg/ml of metronidazole the Δ *addAB* strain is severely affected – with no growth observed at all dilutions, whereas the wild type appears unaffected. C., D., and E. From 0.625 µg/ml of metronidazole we see inhibition of growth of the wild type. At 2.5 µg/ml there is no growth at any dilutions of wild type.

The control plate with no metronidazole showed a 1-2 $\log_{10}$ difference in growth of FP201 compared to NCTC9343. The noticeable difference in growth between the two cultures is due to FP201 lacking a recombination system, the inability to repair DSBs via homologous recombination means that a proportion of bacteria die due to the presence of DSBs. At the lowest concentration of metronidazole (0.05µg/ml) the wild type were unaffected, even though metronidazole was present and would be causing DSBs. The HR repair system is able to cope with this level of damage so viability was unaffected. The deletion strain (FP201) showed at least a 4 $\log_{10}$ reduction in viability compared to the wild type. Without the ability to repair DSBs, the presence of breaks caused by metronidazole had a significant effect on the survival of the *addAB* null strain. The higher levels were only informative for the wild type strain. At 0.6µg/ml metronidazole the effect of the DNA damage accumulating in the NCTC9343 strain caused a 2 $\log_{10}$ reduction in viability. At the level of metronidazole appropriate for sensitivity of wild type strains without resistance (2.5µg/ml) there was a 4 $\log_{10}$ or more reduction in the NCTC9343 strain compared to the no metronidazole control. These results indicate metronidazole may directly or indirectly cause DSBs and further supports the role of the *B. fragilis* AddAB proteins in DSB repair.

3.4.2.4 Are the *B. fragilis addAB* genes required for survival following exposure to oxygen?

B. fragilis is an anaerobic organism, but it able to survive short periods of oxygen exposure (Rolfe *et al*., 1977). To test if the *addAB* genes were required for survival the deletion strain was tested for oxygen tolerance. Stationary phase cultures were serially diluted and 10µl of each dilution were spotted on BHI-S plates - one set of plates were only grown anaerobically while another set had a 3 hour aerobic exposure before returning to grow anaerobically. All experiments were performed in triplicate and a representative example is shown (figure 3.4.2.4.1).

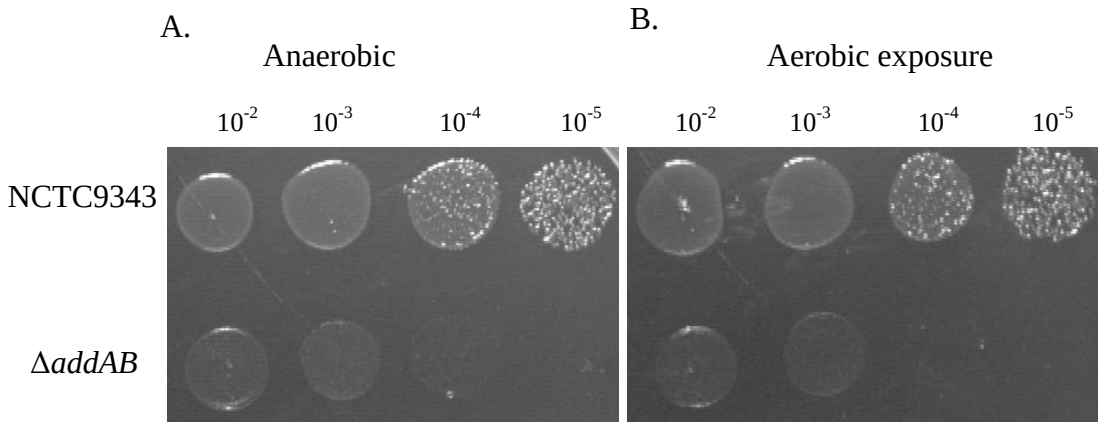


Figure 3.4.2.4.1: plate assay for oxygen sensitivity of NCTC9343 and Δ*addAB* (FP201)

A. Stationary phase cultures of NCTC9343and FP201 were serially diluted and 10µl of each dilution were spotted on BHI-S plates and grown anaerobically at 37°C for 24 hours. B. A replica plate was produced that was exposed to oxygen for 3 hours prior to anaerobic incubation. Comparison of the cultures exposed to oxygen (panel B) with the cultures that received no oxygen exposure (panel A) shows that there is no difference in growth due to exposure.

Both NCTC9343 and FP201 whether grown anaerobically or with an oxygen exposure showed no difference in viability (figure 3.4.2.4.1). This implies that the *addAB* genes are not involved in survival of oxygen exposure. Other genes such as superoxide dismutase or catalase, that are involved in the removal of harmful products due to oxygen exposure, may prevent DNA damage, this will be discussed more fully in chapter 5.

3.4.2.5 Does *B. fragilis* have an SOS system?

In *E. coli*, when there is a significant amount of DNA damage, a set of genes known as the SOS regulon are activated. The SOS response up-regulates genes involved in the repair of DNA damage but at the same time inhibits cell division allowing the cell time to repair the damage so that replication can continue. In *E. coli* as DNA damage increases, RecBCD processes DSBs into 3′ ss-DNA coated with RecA. When RecA is assembled on ss-DNA it becomes a co-protease (RecA*) and facilitates the self cleavage of LexA (the repressor of the SOS regulon) allowing the

SOS regulon to be activated; SfiA is induced which inhibits polymerisation of Ftz, leading to inhibition of septation and filament formation (reviewed by Janion 2008; Dajkovic and Lutkenhaus 2006; Michel 2005).

E. coli recBC strains when exposed to the DNA damaging agent naladixic acid (acting on DNA gyrase and thus causing breaks) display a normal morphology but also high levels of death due to the inability to repair the DNA damage (McPartland *et al*., 1980). If there is no pre-synaptic processing by RecBCD there will be no ss-DNA coated with RecA produced, therefore one would conclude that in the *recBC* null *E. coli* there is no SOS induction and no inhibition of cell division; therefore if *B. fragilis* has an SOS response then the NCTC9343 wild type strain would display a filamentous morphology when metronidazole was present in a culture. Conversely, if the AddAB proteins were involved in the SOS response then the *B. fragilis addAB* null strain would not display an SOS phenotype (filamentous cells) when exposed to metronidazole.

3.4.2.6 Effect of metronidazole exposure on morphology of *B. fragilis* strains

To determine if *B. fragilis* has a SOS system, the morphology of cells of wild type NCTC9343 and Δ *addAB* strains were observed following DNA damage (figure 3.4.2.6.1). High or low levels of metronidazole were added to cultures during exponential growth followed by anaerobic incubation at 37°C for 2.5 hours.

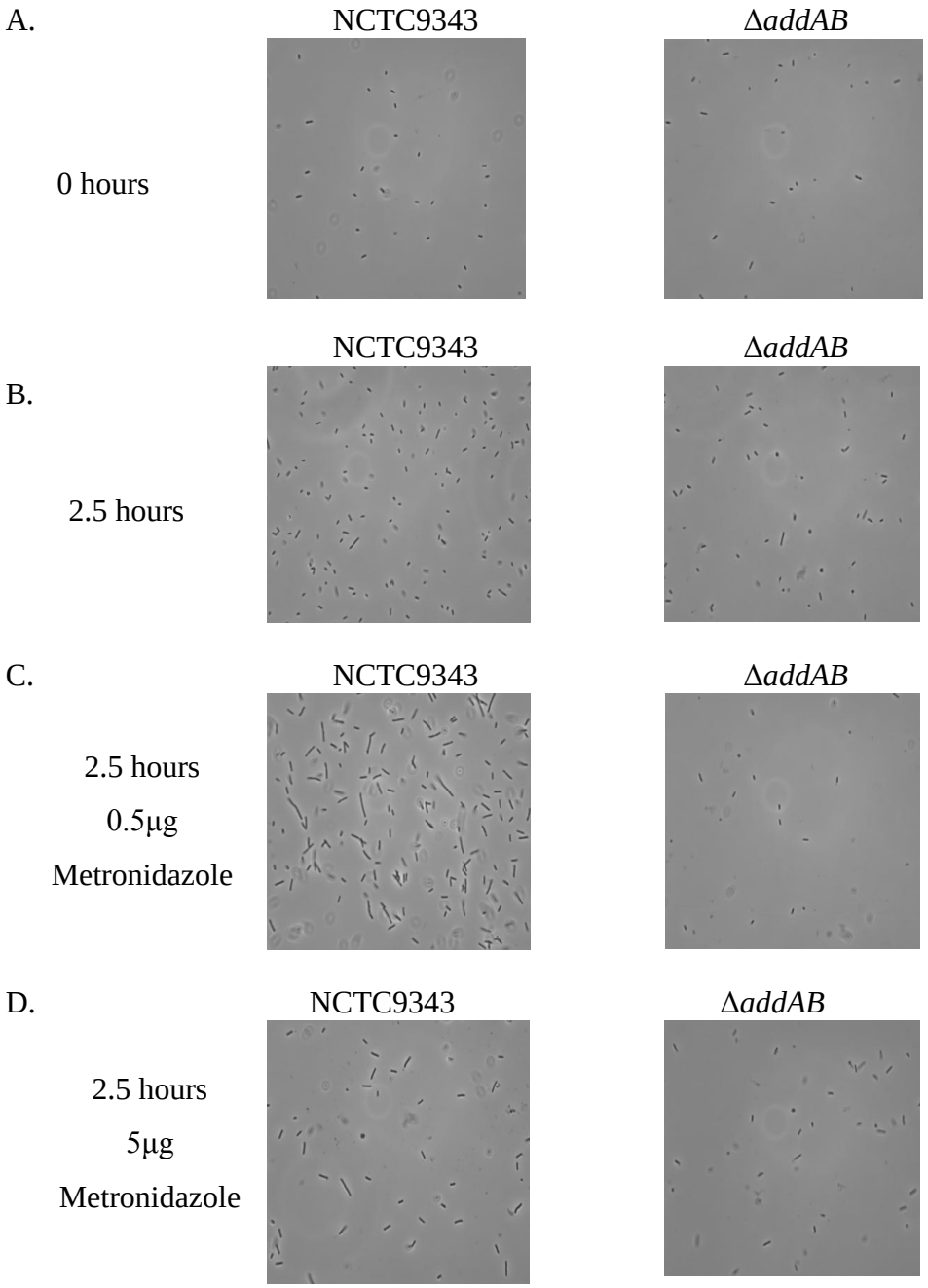


Figure 3.4.2.6.1: Morphology of NCTC9343 and the Δ*addAB* strain in the presence of Metronidazole

Photomicrograph of cells from wild type or Δ *addAB* strains following metronidazole treatment. A. wild type and Δ *addAB* strains 0 hours. B. At 2.5 hours wild type and Δ *addAB* that had not been subjected to metronidazole show normal morphology. C. At 2.5 hours after treatment with 0.5 μg/ml of metronidazole the Δ *addAB* strain displays normal morphology while the wildtype shows filamentous morphology. D. At 2.5 hours after treatment with 0.5 μg/ml of metronidazole the Δ *addAB* strain displays normal morphology while the wildtype shows filamentous morphology.

	0 hours	2.5hours	2.5hours 0.5Met	2.5hours 5Met
NCTC9343	**0.268**	**1.965**	**1.78**	**0.750**
Δ *addAB*	**0.260**	**0.814**	**0.370**	**0.337**

Table 3.4.2.6.1: OD_{600nm} of NCTC9343 and Δ*addAB* strain (FP201) during metronidazole exposure

Table shows OD600nm of cultures before and after addition of metronidazole.

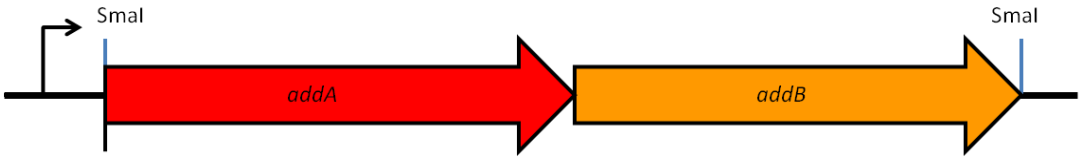
Growth of the Δ *addAB* strain at both concentrations of metronidazole was drastically affected, with significant cell death shown by the low final ODs (table 3.4.2.6.1) and reduced numbers of cells (figure 3.4.2.6.1). When the wild type was grown under sub-inhibitory conditions (0.5μg/ml Metronidazole) for 2.5 hours there is only a small difference in final OD compared to the no metronidazole control.

In contrast with the sub-inhibitory conditions, the inhibitory level of metronidazole (5μg/ml) caused cell death in the wild type culture and some filamentous cells were observed. There was significant chaining and filamentous cells present in the wild type cultures with sub-inhibitory levels of metronidaozle; this morphology indicates inhibition of FtsZ polymerisation, which is a characteristic of an SOS response. Conversely in the cultures containing the Δ *addAB* strain at sub-inhibitory levels of metronidazole there were no filamentous cells. These results show that *B. fragilis* NCTC9343 has a SOS-like response to DNA damage, and suggests that *B. fragilis* AddAB is required for SOS induction in response to DNA damage caused by metronidazole.

3.4.2.7 Attempts to complement the *addAB* mutation

The *B. fragilis addAB* genes, together with their promoter, were amplified by PCR and ligated into pLYL01 giving pFP106 (figure 3.4.2.7.1). Numerous attempts were made to introduce the *addAB* genes into the Δ *addAB* strain by electroporation, but all attempts failed. The wild type strain, however, could be successfully transformed with *addAB* plasmid. The failure to transform the Δ *addAB* strain may be due to the AddAB proteins having a role in transformation, as some papers have shown that in *E. coli* the *recBCD* genes are required for efficient transformation (Hoekstra et al 1980; Qishi and Colsloy cited by Kowalczyowski 1994). Another possibility may be due to the phenotype of the strain – a strain where a large majority of the cells die will have low transformation efficiency. Also if the transformation process itself is damaging to cells it will only increase the stress on the already fragile cells. As levels of cell death in the *B. fragilis* Δ *addAB* strain are higher than in the *E. coli* Δ *recBCD* strain this might explain why the *E. coli* Δ *recBCD* strain could be complemented but not the *B. fragilis* Δ *addAB strain* and why the wild type *B. fragilis* strain could be successfully transformed.

A.



B.

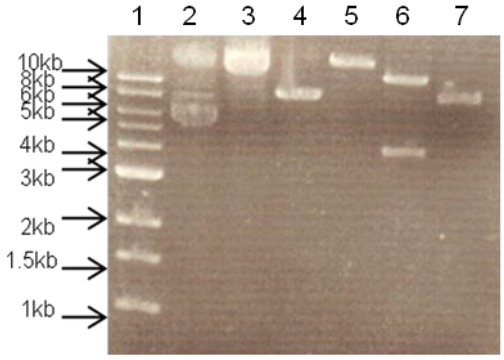


Figure 3.4.2.7.1: Production of pFP106 – AddAB expression vector

A. The *BF2191* and *BF2192* genes were amplified by PCR and ligated into SmaI of pLYL01 giving plasmid pFP103. B. Restriction enzyme digests confirmed the presence of insert in the plasmid. Lane 1-1kb ladder, lane 2 – pLYL01, lane 3- pFP106, lane 4 – pLYL01 digested with SmaI linearising the 8kb plasmid, lane 5 – pFP106 digested with BamHI, linearising the plasmid, lane 6 – pFP106 digested with by PstI, giving two fragments of 3.5 and 11.5kb, lane 7 – pLYL01 digested with PstI linearising the plasmid

3.4.2.8 Markerless Deletions of other putative genes involved in recombination - *addB recF* and *sbcCD*

To determine the phenotypes of strains containing mutations in either *addA* or *addB* the next step was to produce a markerless deletion of *addA* and *addB*. The annotation of the NCTC9343 genome sequence had also highlighted other genes that were possibly involved in *B. fragilis* recombination, *recF* and *sbcCD*. To determine their roles in recombination these would be deleted and their phenotypes characterised.

The protocol for producing these mutants was slightly different from the protocol used for the *addAB* mutant in section 3.4.1; these mutants would be markerless so instead of a cassette containing the erythromycin gene, the deletion plasmid (pGB910) contained *ermF*, therefore the deletion cassette only consisted of regions flanking the gene to be deleted. At the time of ending the project two *addB* null strains and four *sbcCD* null strains had been produced (figure 3.4.2.8.1 and 3.4.2.8.2). The *addA* and *recF* deletions were attempted numerous times but failed to produce the strains. Unfortunately due to time constraints only preliminary experiments were performed using these deletion mutants; the Δ *addB* strain displayed a similar sensitivity as the Δ *addAB* strain to UV and metronidazole exposure, while the Δ *sbcCD* strain showed no sensitivity to UV and metronidazole (data not shown).

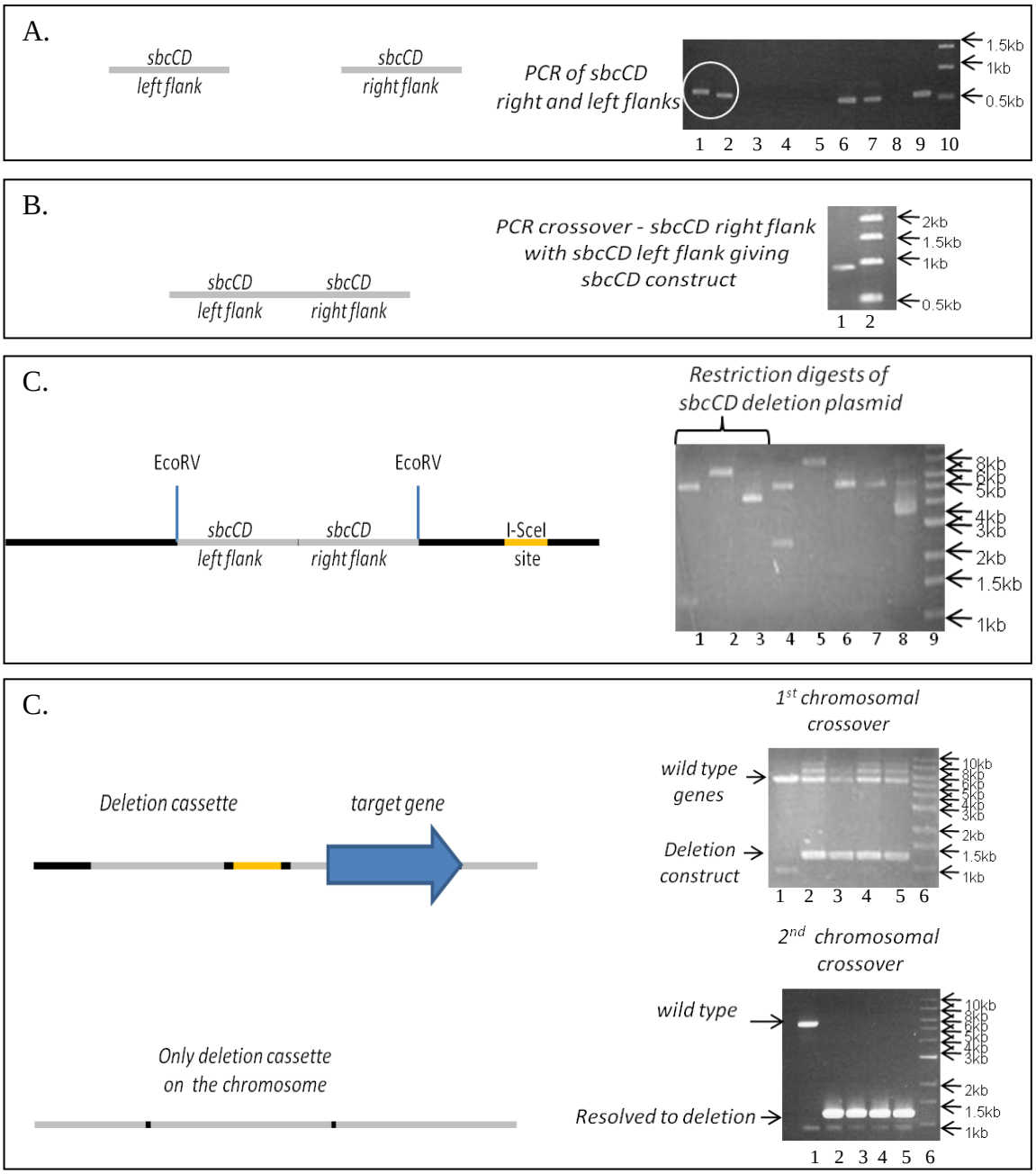


Figure 3.4.2.8.1: Production of deletion plasmid (pFP203) and *Δ sbcCD* strain

A. The regions immediately to the left and right of the *sbcCD* gene were amplified using PCR. Lanes 1 and 2 of the agarose gel show the left and right PCR products respectively. B. The 2 PCR products were fused together by crossover PCR, giving the deletion cassette (lane 1 of the agarose gel). C. The cassette was ligated into the EcoRV site of pGB910. Presence of the insert was confirmed by restriction enzyme digest; lane 1- releasing fragments, lane 2- pFP203, lane3 linearised pFP203, lane 7- pGB910. D. After integration of the plasmid into the chromosome (lanes 2, 3, 4 and 5 of the first chromosomal crossover event agarose gel) a 2nd induced chromosmal crossover event removes the *sbcCD* gene leaving only the deletion cassette (lanes 2, 3, 4, and 5 of the 2nd chromosomal crossover gel).

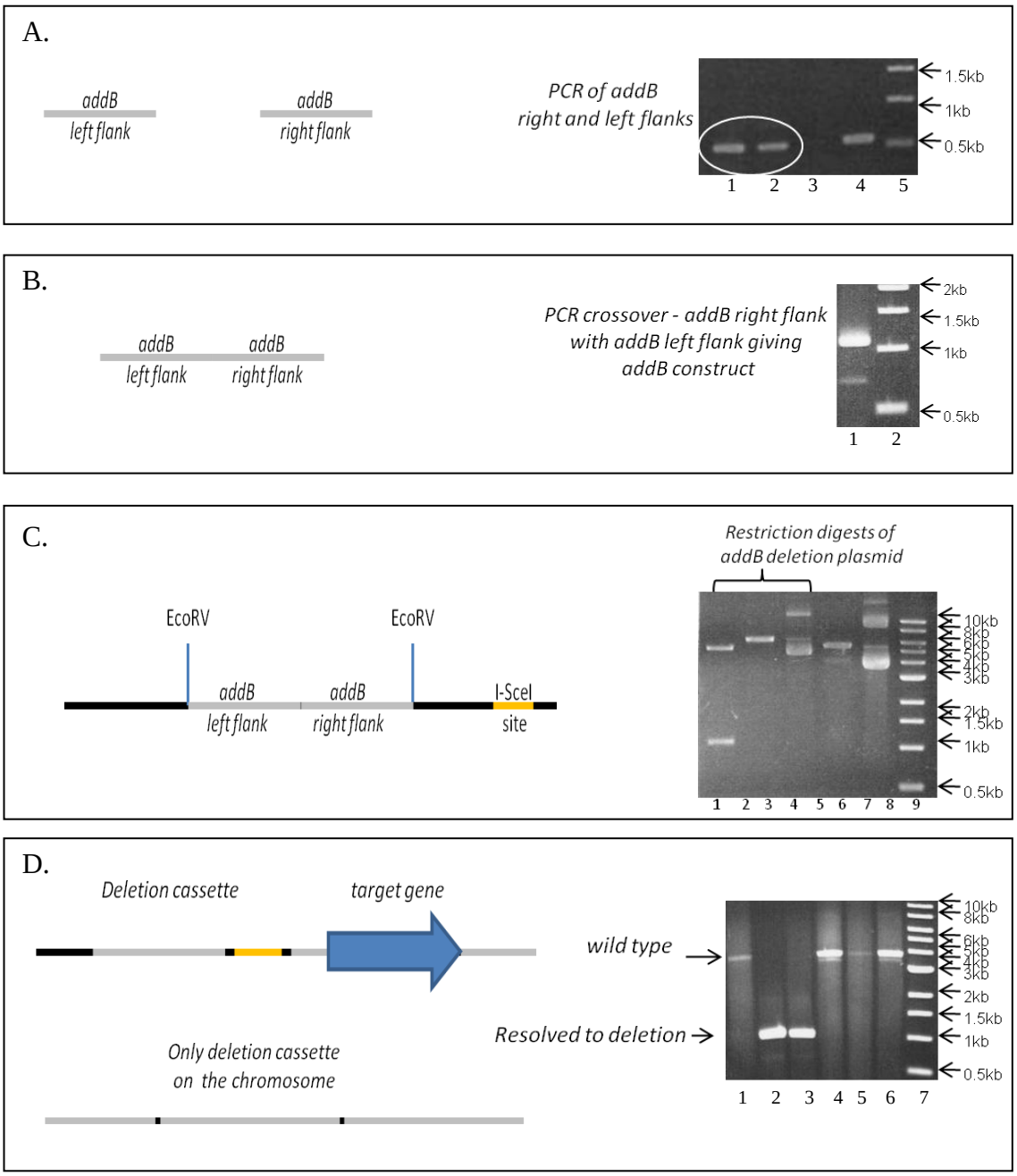


Figure 3.4.2.8.2: Production of deletion plasmid (pFP202) and *Δ addB* strain

A. The regions immediately to the left and right of the *addB* gene were amplified using PCR. The agarose gel lanes, 1 and 2 show the left and right PCR products respectively. B. The 2 PCR products were fused together by crossover PCR, giving the deletion cassette. Lane 1 of the agarose gel shows the fused PCR product. C. The cassette was ligated into the EcoRV site of pGB910. Presence of the insert was confirmed by restriction enzyme digest; lane 1- releasing 2 fragments, lane 2- pFP202, lane3- linearised pFP201, lane 7- pGB910. D. After integration of the plasmid into the chromosome a 2nd induced chromosomal crossover event removes the *addB* gene leaving only the deletion cassette. The agarose gel shows wild type (lanes 4, 5 and 6) and deletion crossover events (lanes 2 and 3).

3.5 *In Silico* Analysis

3.5.1 *In Silico* Analysis - BF0679

The *in vivo* work in *E. coli* provided evidence that *B. fragilis* putative *recD* was not required as part of *B. fragilis* pre-synaptic complex for repair of DSBs. It was also demonstrated that the *B. fragilis* putative *recD* could not rescue the phenotype of an *E. coli recD* null strain (NM1035). If the *B. fragilis* putative RecD is not a functional analogue of *E. coli* RecD, what is its function?

Rocha *et al*., 2005 proposed there are two types of RecD; divided into RecD1 and RecD2. They proposed that the RecD1 family of proteins is found in genomes containing RecBC and is part of the RecBCD complex involved in DSB repair, while the RecD2 family is present in genomes lacking RecBC. Using InterPro scans they characterised the RecD2 type as containing a RuvA-like domain within an N-terminus extension that is not present in RecD1. Montague *et al*., (2009) tried to determine the phylogenetic distribution of both RecD types. For their study they assigned RecD as either 1 or 2 based on size; RecD1 was assigned as short (450-655 amino acids) while the RecD2 was longer (more than 710 amino acids). The phylogenetic analysis led Montague *et al*., (2009) to propose that the RecD1 family are associated with RecBC, while RecD2 was found mainly with AddAB containing bacteria; therefore the association is phylogenetically distinct. They found no instances of RecD1 with AddAB containing bacteria.

To determine if *B. fragilis* RecD is type 1 or 2, the criteria in the above studies were used. *B. fragilis* RecD is short (472 amino acids in length) so this would assign it to the RecD1 family. An InterPro scan did not find any RuvA-like domains; the lack of this domain would also assign *B. fragilis* RecD as RecD1. A CDART multidomain scan of *B. fragilis* RecD gave the top two hits as TIGR01448 (E value = 4.88e-24) and RecD (E value = 1.24e-20). The TIGR01448 is a family described as RecD like proteins that are as 200aa longer than ReD1 and are not found associated with RecBC. Therefore *B. fragilis* RecD could be either D1 or D2. *B. fragilis* RecD

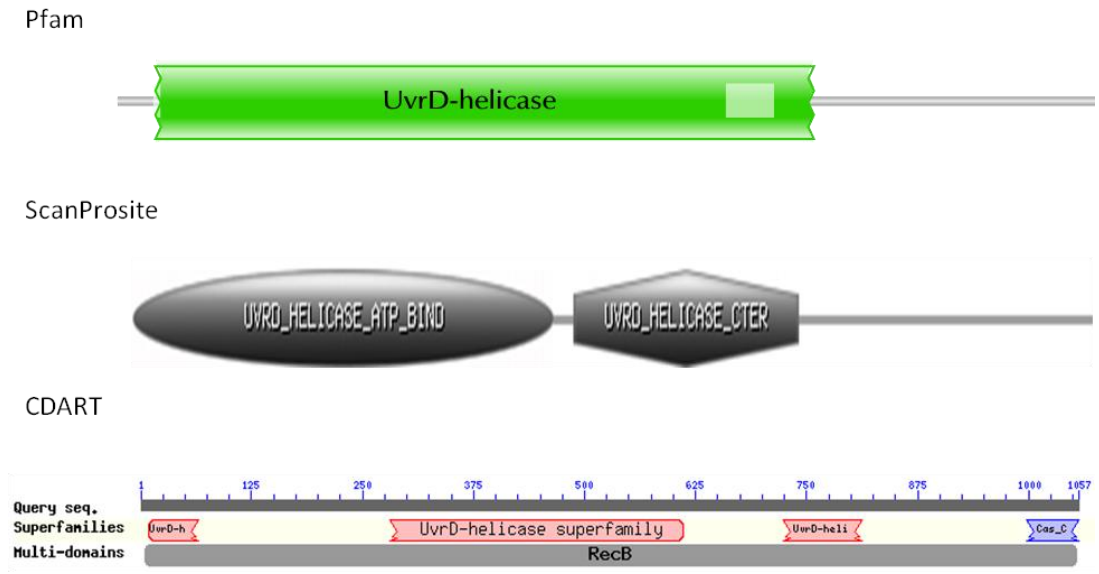
shows multi-domain matches to both RecD1 and D2, but it is not the size associated with the D2 group, rather its size associates it with RecD1. In addition, it does not have an N-terminal extension with a RuvA-like domain that is indicative of the RecD2 type. The original NCTC9343 annotation gave BF0679 as a protein with 37% identity to *E. coli* RecD. Together, these data and the above criteria suggest that *B. fragilis* RecD may be of the D1 type.

3.5.2 *In Silico* Analysis - AddAB

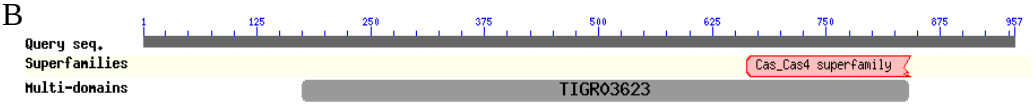
To further examine the role of the *B. fragilis addAB* protein products, searches for protein homology and presence of domains were performed using Pfam, Prosite and CDART. Domain scans of AddA using Pfam and Prosite both gave UvrD as the best match; UvrD helicase (Pfam E value = 3.2e-57) and UvrD ATP binding domain (Prosite score of hits against a profile 26.5) and the UvrD C-terminal domain (Prosite score 9.4) (figure 3.5.2.1A). The CDART scan of AddA found several domain matches, but the highest was to UvrD (E value = 6.74e-22). The AddA CDART multi-domain scan best match was to RecB (E value = 4.59e-61) (figure 5.10.1A). AddB had no significant domain matches using Pfam or Prosite, however CDART found a match to the Cas_cas4 super family (E value = 7.64e-07) (figure 5.10.1B). This domain as annotated by CDART as having no known function but is classified by the presence of 3 conserved cysteines at its C-terminus. Yeeles *et al*., (2009) have shown the presence of an Fe-S cluster in *B. subtilis* AddB consisting of 4 Cysteines (3 are at the C-terminus while a fourth flanks the other side of the nuclease motif). This structure is essential for DNA end binding, since mutation of any of the cysteines within the iron sulphur cluster caused loss of AddAB binding (Yeeles *et al.,* 2009). The presence of an iron cluster in *B. fragilis* AddB was confirmed by Reuter *et al*., (2010); consisting of three C-terminal cysteine residues (C946, C949 and C955) and a fourth cysteine (C674), flanking a nuclease motif. The AddB CDART multi-domain best match was to TIGRO3623 (E value = 1.53e-09) a protein family whose members have homology to the AddB protein family. The RecB, AddA and AddB sequences were aligned using the program COBALT. The nuclease consensus sequences for *B. subtilis* (RexA - G-i-i-D-x(12)-D-Y-K-t-d) and

RexB (G-r-i-D-R-i-D-x(9–12)-v-D-Y-K-S-s) were used to identify the AddA and AddB nuclease sites (Quiberoni *et al.*, 2001). The Walker A motif consensus sequence (GXXXXGKT/S) was used to highlight the ATPase essential for helicase activity. Both AddA and AddB contain nuclease motifs but only AddA contains a predicted Walker A motif for binding of ATP (figure 3.5.2.1C and D). The amino acid sequence analysis in conjunction with the domain hits suggests that both *B. fragilis* AddA and AddB are nucleases but only AddA is a helicase.

A



B



C

```
RecB  1    MSDVAETLDPLRLPLQGERLIEASAGTGKTFTIAALYLRLLLGLGGSAAFPRPLTVEELLVVTFTEAATAELRGRIRSNI  80
AddA  1    MSELI--------------VYKASAGSGKTFTLAVEYIKLLIR--------NPRAYRQILAVTFTNKATAEMKERILSQL  58
AddB  1    --------------------------------------------------------------------------------
```

D

```
RecB  1011 QLSARNKQVEMEFYLPISEPLIASQLDTLIRQFDPLSAGCPPLEFMQVRGMLKGFIDLVFRHEGRYYLLDYKSNWLGEDS  1090
AddA  954  -----EKQIRSLTVKAFSLPEVQEWYSGEWRLFNECAIIYKDKGVLQTRRP-----DRVMMKNEQVVVVDFK---FGKAN  1020
AddB  786  -------------YAPFRMMGMEQEVVEDIRIEGPV-------GKLSLRIG—G TIDRMDSKEGTLRIVDYKTG--GSPK  841
```

Figure 3.5.2.1: sequence analysis of AddA and AddB

A. Pfam, Prosite and CDART graphical representation of proposed domains of AddA. B. CDART graphical representation of proposed domains and multi domain family match for AddB. C Portions of alignments of RecB, AddA and AddB using COBALT. Highlighted in bold and underlined is the proposed Walker A motif of AddA. No Walker A motif was found in AddB. D. Alignments of RecB, AddA and AddB using COBALT. Highlighted in bold and underlined are the proposed nuclease motifs of AddA and AddB

3.6 Summary

The main aim of this work was to identify the *B. fragilis* NCTC9343 genes involved in the pre-synapsis stage of homologous recombination. Through inter-species complementation, production of deletion strains and phenotypic characterization, the *B. fragilis* BF2192 and BF2191 genes have been identified as encoding AddA and AddB homologues, respectively.

It was demonstrated that their presence in an *E. coli* Δ *recBCD* strain is sufficient to alleviate the growth defect of the strain and enable repair of DNA DSBs caused by UV exposure. Expression of AddAB rescued the recombination deficient phenotype to the same level as wild type *E. coli*. This suggests that the mode of action of the proteins on DNA is similar to that of *E. coli* RecBCD complex and must generate a 3′ overhang, the substrate required for synapsis by RecA. It shows that no accessory proteins are required to direct the *B. fragilis* complex to DNA ends and suggests that *B. fragilis* chi sites are present in sufficient numbers to allow recombination to occur. It may also be possible that the *B. fragilis* AddAB chi site sequence is the same as *E. coli* or similar enough to allow for a level of degeneracy and therefore recognition of the site by the *B. fragilis* machinery.

The *B. fragilis* work (section 3.4) has shown that AddAB is required for the repair of DSBs caused by endogenous damage. Deletion of the genes was deleterious to the cells, observed as a drastic reduction of growth rate. It was also observed that the *B. fragilis addAB* deletion is more detrimental to cell survival than the equivalent *E. coli* Δ *recBCD* mutation. This could imply that *B. fragilis* has more endogenous damage, or that the genes have another role in protection of the cell. Alternatively, *E. coli* may have other systems that can compensate for the loss of RecBCD.

The AddAB complex is also required for the repair of DSBs caused by exogenous damage. The proteins are essential for UV and metronidazole resistance in *B. fragilis* NCTC9343; deletion of the genes made the strain more sensitive to UV and metronidazole exposure. In addition to this the wild type *B. fragilis* NCTC9343 was

more sensitive to UV exposure than wild type *E. coli* MG1655. At 80J/m2 of UV light MG1655 was unaffected whereas NCTC9343 had a 4 $\log_{10}$ or more reduction in viability.

In this chapter other roles for the *B. fragilis* genes were examined. In section 3.4.2.4 it was found that the genes were not involved in the oxygen tolerance of *B. fragilis*. In section 3.4.2.5 it was shown that NCTC9343 has an SOS system and that the *B. fragilis addAB* genes are involved in activation of the SOS response.

Most importantly it has been demonstrated that *B. fragilis* contains a recombination system that was originally believed to be present only in Gram positive bacteria. The *B. fragilis* proteins share homology, though low, with the *Lactococcus lactis* RexAB proteins, and therefore is a two component system, with both proteins required to work in concert for DSB repair. At present it appears there are three different recombination systems used throughout the eubacteria; the AddAB/RexAB, AdnAB and the RecBCD system. All systems function in a similar manner but differ in the physical composition of the machinery. At the time of starting this project the paradigm for DSB repair was the Gram negative *E. coli recBCD* system; this project has shown that assumption that RecBCD is the representative system for Gram negative bacteria to be untrue because the Gram negative *B. fragilis* has an AddAB type system.

All of the above points will be discussed in chapter 5. To more fully understand the nature of the *B. fragilis* recombination machinery the following chapter looks at the purification of the BF AddAB proteins and their biochemical action on DNA.

Chapter 4 – Purification and Functional analysis of AddAB and associated proteins

4.1 Introduction

In chapter 3 the BF2192 and BF2191 genes were identified as encoding putative AddA and AddB homologues, respectively, and their role in DSB repair in relation to endogenous and exogenous DNA damage in *B. fragilis* NCTC9343 was demonstrated.

This chapter describes the purification of the AddAB complex, the AddA and AddB proteins separately and examines the biochemical action of the proteins on DNA. In the early stages of this project it was believed that the *B. fragilis* proteins involved in pre-synaptic DSB repair were encoded by BF0679 (*recD*) and BF2192 (*addA*). Initial work focused in part on the putative RecD protein; although it seemed likely that the putative RecD is not involved in *B. fragilis* pre-synaptic processing. The role of the putative RecD is puzzling, therefore purification was attempted with the aim of determining biochemical function. In addition, this helped in development of protocols that were useful for further work; therefore the opening sections of this chapter will examine the purification of this protein. The expression and purification of the AddAB complex and the AddA and AddB proteins were the main focus of this work. In this chapter, the problems encountered during purification will be described and finally how the purification has enhanced our understanding of the protein complex. The biochemical actions of the AddAB complex, and the AddA and AddB proteins, will be described with a focus on the preferred DNA substrate for the complex. The requirement for co-factors for function of the complex will also be investigated. Since this chapter follows the experimental determination of protocols for purification, it will include descriptions of these protocols. Ultimately the purification strategy developed here, led to the purification of AddAB for use in the first published analysis of this group of proteins at single molecule resolution (Reuter *et al*., 2010).

4.2 Purification

This section will look at the purification of the putative RecD and the AddAB proteins. Problems were encountered with low expression of proteins; therefore different methods of expression were tried to improve yields. Different techniques for purification will be described, with discussion of their effectiveness at producing pure and active protein samples for use in biochemical assays. The outcomes were important for determining the final choice of method for purification.

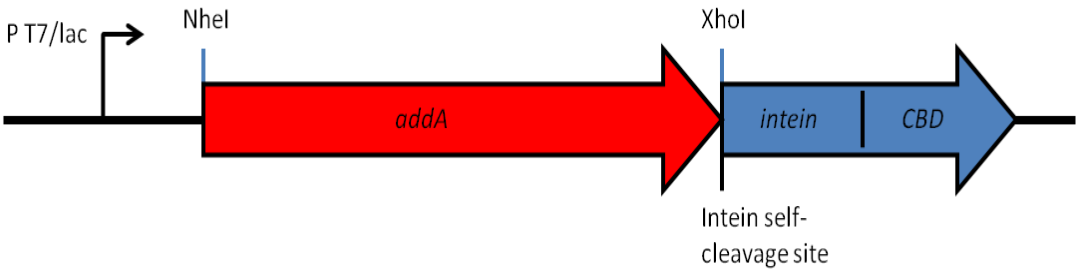
4.2.1 Purification of proteins using the IMPACT system

At the start of this project it was proposed that the protein products the *B. fragilis* genes BF0679 and BF2192 were involved in the pre-synaptic processing of DSBs in *B. fragilis* NCTC9343, therefore this first section will describe the purification of putative RecD and AddA. To ensure purification of active protein a method that produced a high yield of proteins with no tag was required. The IMPACT (Intein Mediated Purification with an Affinity Chitin-binding Tag) system from NEB was selected for a one step purification method. The target protein is fused to an intein tag that contains a chitin binding domain (CBD); the intein protein is self-cleavable in the presence of thiols. Once the fusion protein is bound to a chitin column, the target protein is eluted by adding thiols to the elution buffer, thus freeing the desired protein while leaving the intein bound to the chitin column.

4.2.1.2 Expression

The pTYB2 plasmids provide expression of fusion proteins consisting of an intein tag on the C-terminus of the target protein. The BF0679 and BF2192 genes were cloned into pTYB2 to produce pFP111 and pFP110, respectively (figure 4.2.1.2.1 and 4.2.1.2.2). The expressed fusion proteins of pFP111 and pFP110 were named RecD-CBD and AddA-CBD, respectively.

A.



B.

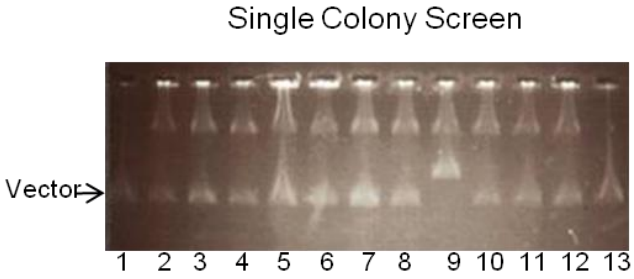


Figure 4.2.1.2.1: Production of pFP110

Construction of the AddA expression vector pFP110. A. The *addA* (BF2192) gene was amplified by PCR and ligated into the pTYB2 plasmid, such that it was fused to the intein-*CBD* ORF at the C-terminus of AddA. B. Lane 9 of the single colony screen shows a plasmid larger than the pTYB2 vector (lane 1 and 13) indicating the presence of an insert (PTYB2 plasmid = 7.5kb, Insert = 3.2kb Lanes 2,3,4,5,6,7,8,10,11 and 12 show plasmids with no insert.

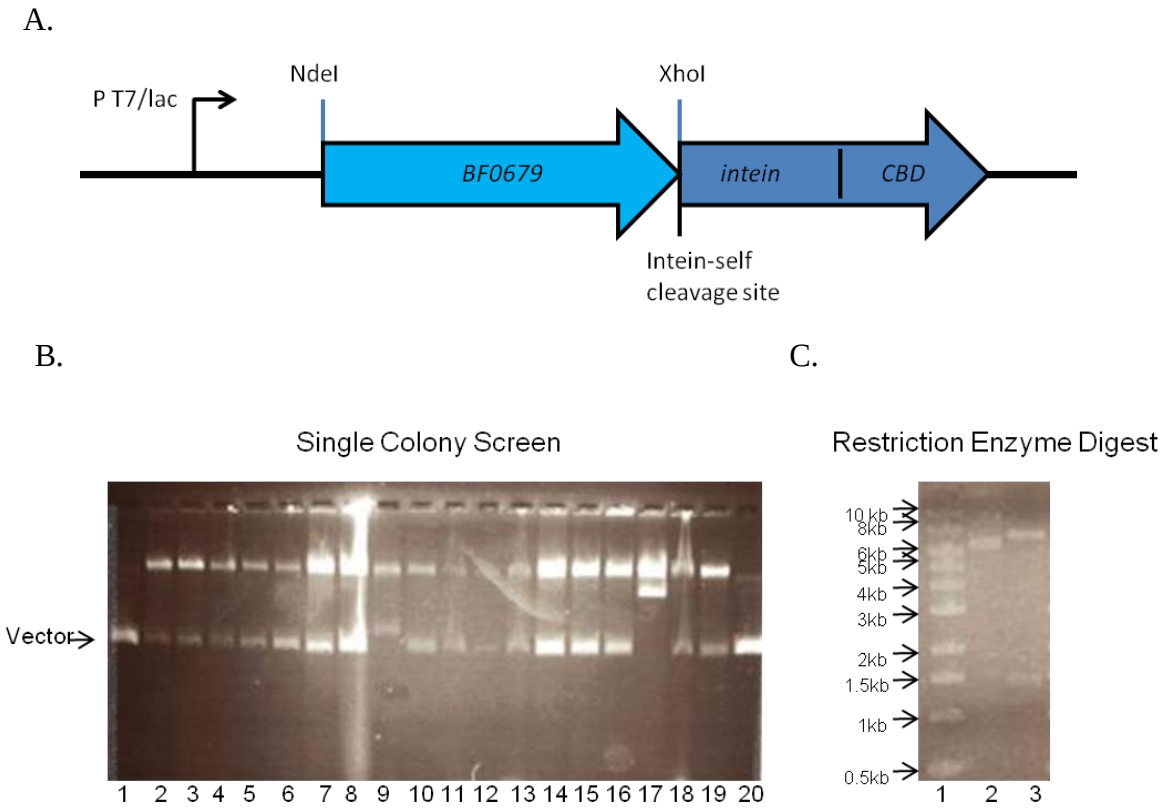


Figure 4.2.1.2.2: Production of pFP111

Construction of the BF0679 expression vector pFP111. A. The BF0679 gene (encoding a putative RecD) was amplified by PCR and ligated into pTYB2 producing a fusion protein containing intein to the C-terminus of the putative RecD. B. Lane 9 of the single colony screen shows a plasmid larger than pTYB2 (lane 1 and 20) indicating the presence of an insert. Lane 17 suggests dimerisation of the plasmid. Lanes 3-8,10-16, and 18-19 show plasmid with no insert. C. The restriction enzyme digests confirmed the presence of insert in the plasmid. Lane 1- 1kb DNA ladder, lane 2 – pFP110, lane 3 - pFP110 digested with NdeI/XhoI, releasing the 1.4kb insert and the 7.5kb plasmid.

Over-expression of both RecD-CBD and AddA-CBD proved problematic, so various expression methods were tried. Initial attempts used the host strain supplied by NEB, ER2566 and were performed at 37^{o}C. A 1/100 dilution of an overnight culture of each strain was used to inoculate 1l of fresh LB in a 5l flask. The culture was grown at 37°C with shaking to an OD_{600} of 0.4 before addition of IPTG to a final concentration of 1mM. The AddA protein has a predicted size of 117 kDa while the RecD protein has a predicted size of 52kDa. With the intein-CBD tag (55kDa)

BF0679-CBD would be predicted to be 107 kDa and AddA-CBD would be 170 kDa. Bands of the expected sizes were detected by SDS-PAGE, however the levels of expression of both the BF0679-CBD and AddA-CBD proteins were low (Fig. 4.2.1.2.3.). A lower temperature (20^{o}C overnight) was used for induction in an attempt to increase yield, however, there was no significant increase in induction for either of the proteins (Fig. 4.2.1.2.3).

The low levels of expression could have been caused if the proteins were deleterious to the cells. The ER2566 host uses the Lac I repressor to control the T7 RNA polymerase and is known to have leaky expression. To reduce background expression of T7 RNA polymerase another host was used; BL21 (DE3) plysS; this encodes T7 lysozyme that binds to T7 RNA polymerase, preventing transcription. No discernable increase in expression of BF0679-CBD or AddA-CBD was detected with this strain (Fig. 4.2.1.2.3.).

Another possible reason for low expression was proposed based on the codon usage differences between *B. fragilis* and *E. coli* (Table 4.2.1.2.1.). The codon usage of the putative *recD* and *addA* genes, when compared to the codon usage in the genome of *E. coli K12*, suggested that the variance was significantly different. Since the required codons are rare in *E. coli*, early termination of protein translation may occur. To combat this, a rare codon strain BL21 (DE3) codon plus – RIL was used as a host for expression of RecD-CBD and AddA-CBD. This strain encodes tRNAs for rare arginine codons AGA and AGG, isoleucine codon AUA, and leucine codon CUA; these codons showed the greatest difference in usage. However no significant increase in the expression level of RecD-CBD or AddA-CBD was found when this strain was used (Fig. 4.2.1.2.3.). This result could be due to the rare tRNAs not being supplied at a high enough level required for the over-expression of these proteins or may suggest that the tRNAs were not affecting the level of expression.

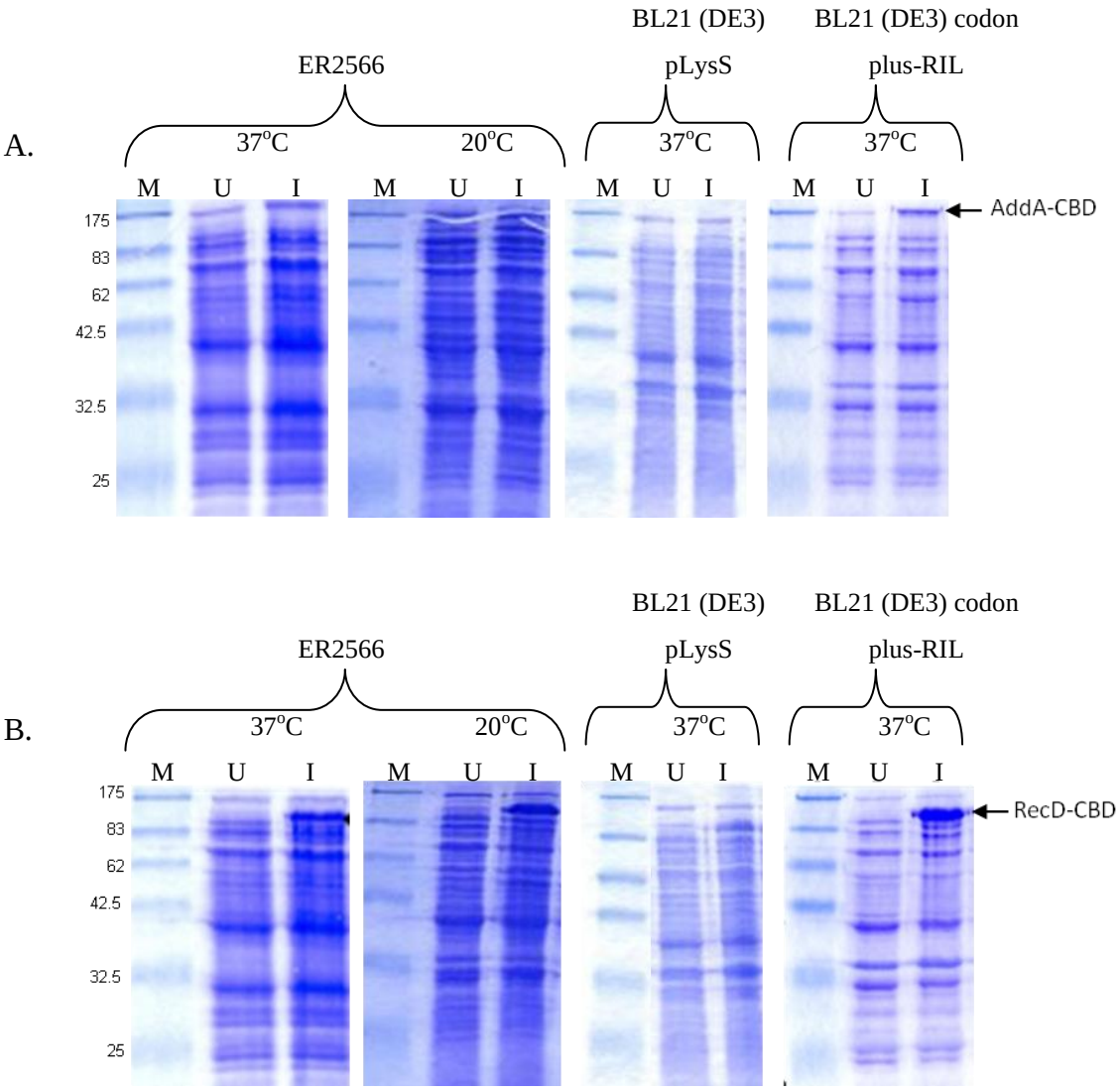


Figure 4.2.1.2.3: Expression conditions for RecD-CBD and AddA-CBD

A selection of various conditions used for expressing AddA-CBD and RecD-CBD. A. Shows SDS-PAGE analysis of AddA-CBD expression at 37°C or 20°C using either ER2566, BL21 (DE3) pLysS or BL21 (DE3) codon plus-RIL as host. B. Shows SDS-PAGE analysis of RecD-CBD expression at 37°C or 20°C using either ER2566, BL21 (DE3) pLysS or BL21 (DE3) codon plus-RIL as host.

M = 1 kDa marker, U = uniduced culture, I = Induced culture. Arrows highlighting bands of the expected size of expression products.

Codons	*E. coli K12*	*B. fragilis NCTC 9343*	*B. fragilis* NCTC 9343 *recD*	*B. fragilis* NCTC 9343 *addA*
AGG	1.2	3.3	8.5	2.1
AGA	2.1	6.5	2.1	3.1
AUA	4.3	17.1	21.1	20.9
CUA	3.9	4.8	4.2	10.4

Table 4.2.1.2.1: Codon usage comparison

Displayed values are in frequency per thousand codons of coding sequence. There is great variance in the codon usage between *E. coli* and *B. fragilis*; the table shows only a selection of codons, In this example the codons that would most impede expression of the AddA protein would be AUA and CUA, for the putative RecD protein it would be AGG and AUA. A full analysis of codon usage can be found at ://www.kazusa.or.jp/codon/ The Codon Usage Database (source GenBank)

To address whether the low levels of induced protein were due to degradation, proteins levels were observed over time (figure 4.2.1.2.4). Induced cultures were prepared as previously described (37^{o}C), it should be noted there was only one initial addition of IPTG. For AddA-CBD at 120 minutes the band appeared less intense than previous time points, which might suggest degradation of the polypeptide. For RecD-CBD there did not appear to be any significant degradation of the protein since band intensity remained constant after the initial induction.

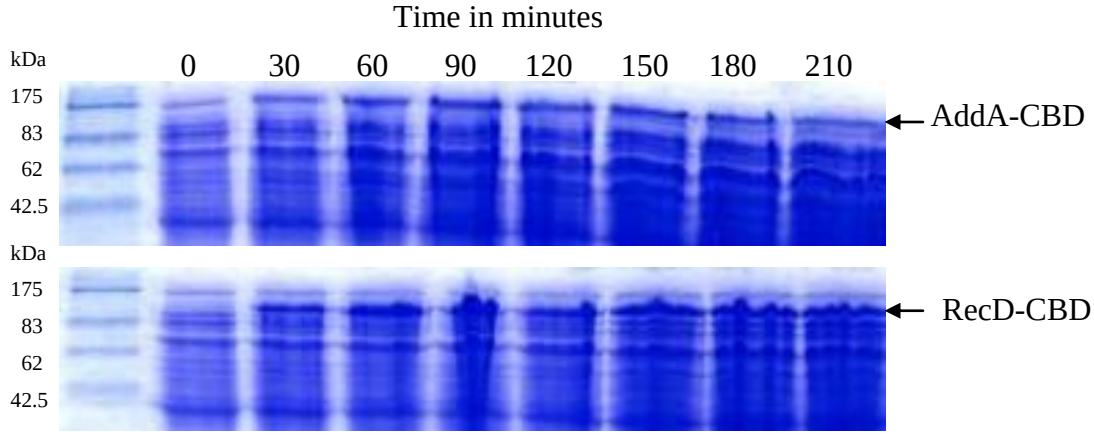


Figure 4.2.1.2.4: Levels of RecD-CBD and AddA-CBD over time

SDS-PAGE gel showing levels of banding over time of induced cultures (AddA-CBD and RecD-CBD) in host ER2566 incubated at 37°C. AddA-CBD appeared to have degradation - noticeable at 120 minutes, whereas RecD-CBD levels remained constant after the initial induction. Arrows highlight bands of the expected size of expression products.

In summary, alterations to the induction method did not produce significant increases in expression levels. The use of the rare codon strain did not improve induction of either of the proteins, suggesting that codon usage was not the reason for low expression. Lower temperatures gave only a slight increase in expression. Closely regulating the induction by use of the BL21 (DE3) plysS strain did not improve protein levels, implying that expression was not deleterious to the cell. The RecD-CBD induction was always greater than AddA-CBD; this difference suggested that there may be degradation of the AddA protein. This was supported by the apparent degradation of the AddA-CBD protein after 210 minutes following induction.

4.2.1.3 IMPACT Purification

Despite low levels of induction, purification was attempted. Preparation of the culture and purification of the proteins using the IMPACT system is described in materials and methods (section 2.8.5); the low levels of induction were reflected in the poor yields of purified proteins (figure 4.2.1.3.1). Analysis of the AddA elutions showed that the AddA bands were of low intensity. The RecD elutions produced more protein, but there was a large loss of protein when compared to the original induced band. In both cases there were large amounts of contaminating proteins; this can be seen more clearly in the concentrated elutions. This suggested that there was degradation of target protein (particularly AddA) and that other proteins were binding to the column and preventing the intein-CBD tag binding. The quantities and purity of the samples were not useful therefore other methods of purification of AddA were attempted.

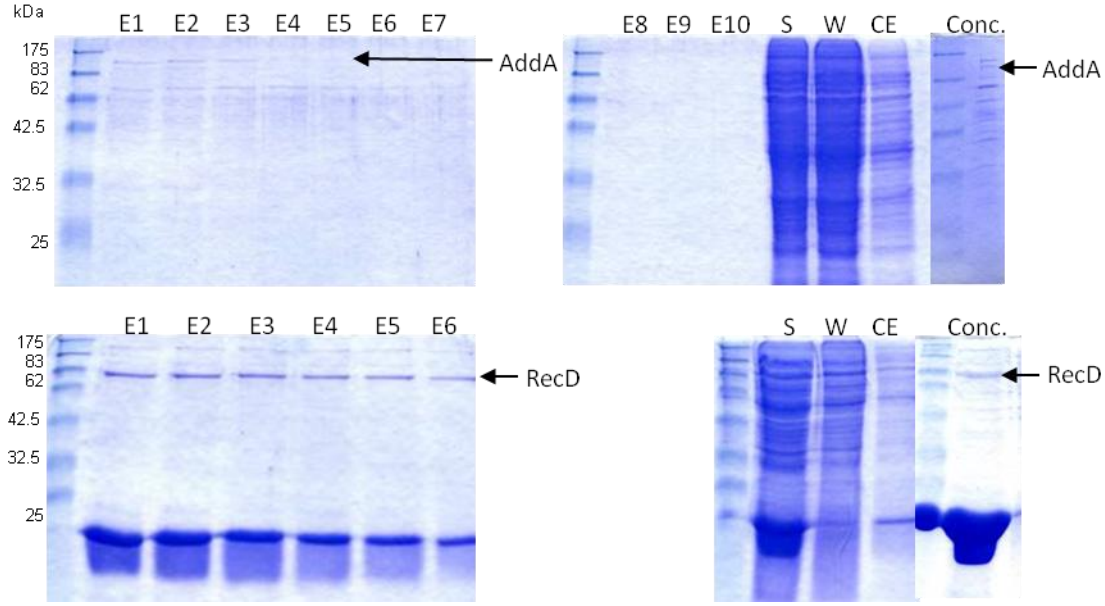


Figure 4.2.1.3.1: IMPACT purification

IMPACT purification of RecD and AddA. AddA protein band was faint, whereas the RecD band is stronger. Both purifications showed loss of target protein in the cleavage elute (CE) and wash (W). Both samples showed background contamination once concentrated (Conc.). E- elution, S- supernatant, W- wash, CE- cleavage elute wash, Conc.- concentrated elutions.

4.2.2.1 AddAB Purification Method 1

Genetic tests (chapter 3) had shown that BF0679 was not required for pre-synaptic processing, but had identified AddA and AddB as functional partners. This meant that work now focussed on purifying these proteins. Work in the previous section suggested there was an issue with degradation of AddA. It was proposed that co-expression of AddA and AddB in *E. coli* could provide structural protection from degradation by proteases, since sites which would be open to proteases could be hidden when the two proteins are complexed. In RecB the nuclease domain is connected to the rest of the protein by a linker region that is sensitive to proteolytic cleavage (Yu et al., 1998). Perhaps either, or both of the *B. fragilis* proteins have

sites that are sensitive to proteases, therefore the proteins were expressed together prior to purification.

Plasmid pTrc99a was used to construct the expression vector; this would produce the proteins without a tag. Absence of a tag might prevent any obstructions to the proteins forming a complex. The *B. fragilis* NCTC9343 BF2192 and BF2192 genes were amplified by PCR and cloned into pTrc99a giving plasmid pFP101 (Chapter 3, section 3.3.3.2). To prevent possible contamination with RecBCD a Δ *recBCD E. coli* strain (FP101) was used for expression.

A mid-logarithmic 2L culture of *E. coli* Δ *recBCD* FP101 was induced with 1mM IPTG for 16 hours at 18°C to over express the AddAB proteins (Figure 4.2.2.1.1). The SDS-PAGE analysis showed bands that were the appropriate size for AddA and AddB. Crude cell lysates were prepared as described in chapter 2 (section 2.8.2); cells were lysed using Brij to improve lysis and solubility of proteins; Brij is a non-ionic detergent that lyses cells more slowly than Triton x and is less likely to denature proteins, and so will increase yields. The Brij lysis was followed by an ammonium sulphate precipitation, performed as described in chapter 2 (section 2.8.2). The sample was resuspended in buffer (20mM tris-Cl, pH7.5, 0.1mM EDTA, 0.1mM DTT) + 100mM NaCl. The sample was dialysed to remove salts and Brij58, described in chapter 2 (section 2.8.9).

The protein extract was loaded onto a Q sepharose column (GE Amersham) in buffer (20mM tris-Cl, pH7.5, 0.1mM EDTA, 0.1mM DTT) + 100mM NaCl, then washed with buffer +300mM NaCl and eluted with a gradient to 600mM NaCl. Of the 20 1ml samples collected, 7 showed enrichment of AddAB (Figure 4.2.2.1.1). Elutions E1-7 were concentrated and resuspended in 50mM potassium phosphate buffer, pH7.5, followed by dialysis against the same buffer. The sample was loaded onto a Hydroxyappatite column and eluted with a gradient to 300mM potassium phosphate buffer, pH7.5. The Hydroxyappatite elutions indicated that the column had cleaned the sample as shown by the loss of background proteins (Figure 4.2.2.1.2). The samples appeared blurry on the gel; probably due to the high salt concentration of

potassium phosphate buffers. Elutions 1-4 and 5-8 were concentrated (Vivaspin) (Figure 4.2.2.1.2). Concentrated sample C1 showed enrichment for AddA and AddB; C1 was dialysed against buffer (20mM tris-Cl, pH7.5, 0.1mM EDTA, 0.1mM DTT) + 100mM NaCl, before loading onto a heparin column. This column was used since this should bind AddB but not AddA (Chedin et al., 2001), allowing the proteins to be eluted as a complex. The sample was eluted with a gradient to buffer (20mM tris-Cl, pH7.5, 0.1mM EDTA, 0.1mM DTT) + 500mM NaCl. Elution E4 from the heparin column showed the presence of AddA and B and loss of an amount of contamination (Figure 4.2.2.1.3). It is interesting to note that the concentration of AddB appeared to be higher than AddA. This could suggest that AddB was binding to the column and there was a certain amount of dissociation of AddA from AddB resulting in loss of AddA, or that AddA was more sensitive to degradation. Following elution from the heparin column, E4 was dialysed with buffer (20mM tris-Cl, pH7.5, 0.1mM EDTA, 0.1mM DTT) + 100mM NaCl before loading on to a mono Q column. The sample was washed with (20mM tris-Cl, pH7.5, 0.1mM EDTA, 0.1mM DTT) + 200mM NaCl and eluted with a gradient to (20mM tris-Cl, pH7.5, 0.1mM EDTA, 0.1mM DTT) + 600mM NaCl. The mono Q column elutions showed bands of a size appropriate for AddA and AddB (Figure 4.2.2.1.4) and indicated the sample was relatively pure.

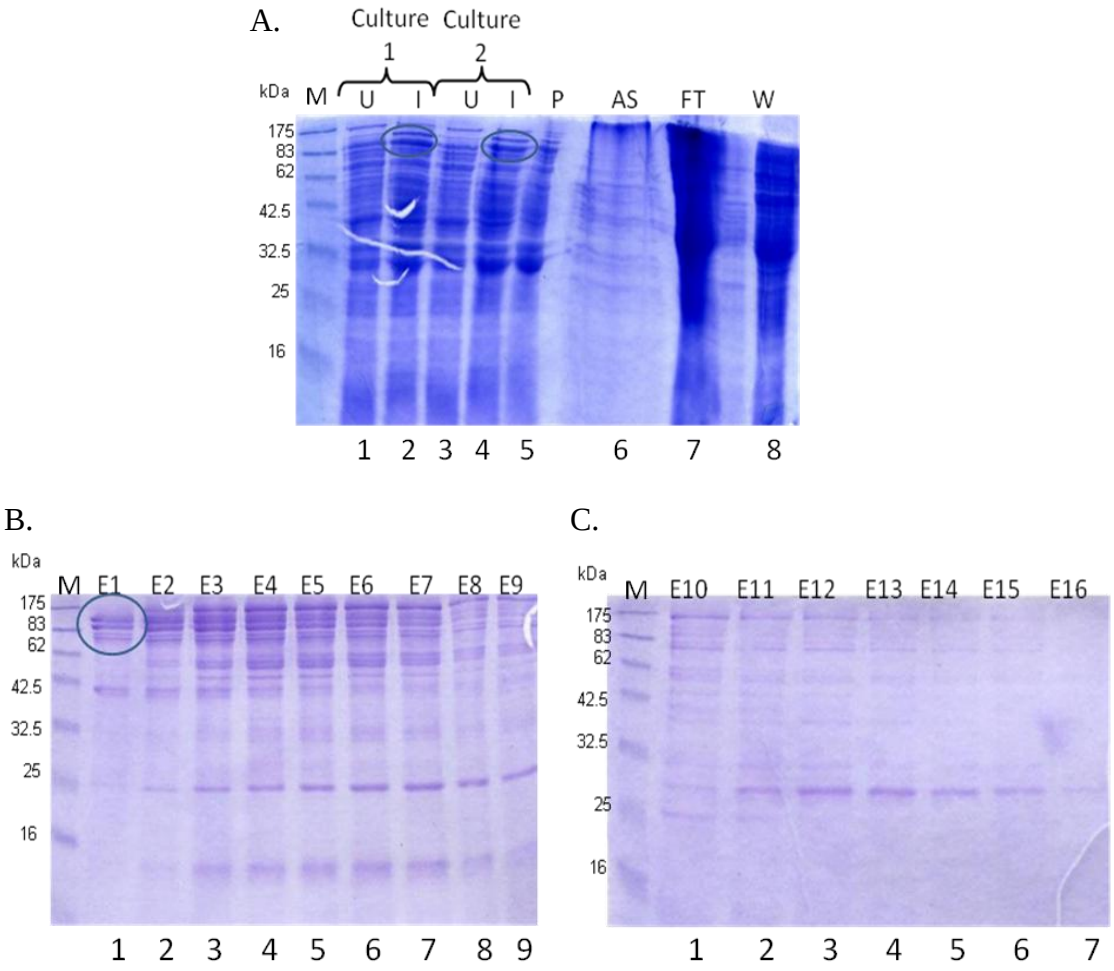
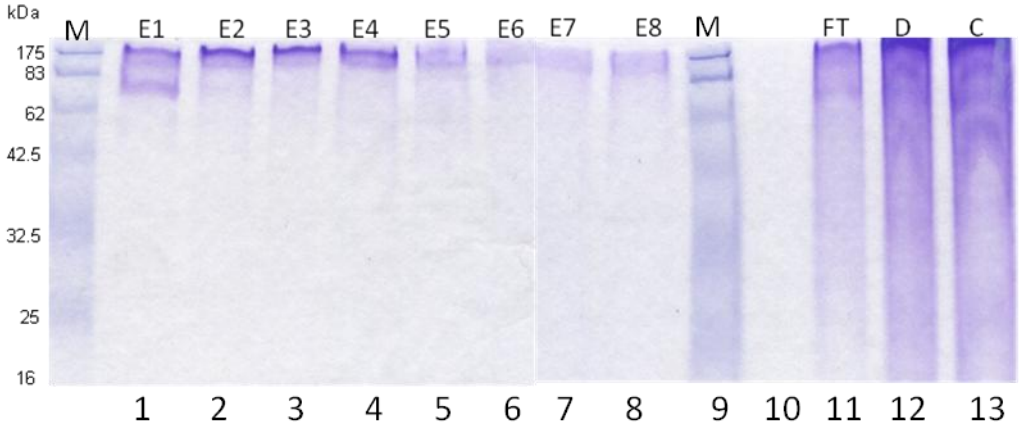


Figure 4.2.2.1.1: SDS-PAGE Gel of Induction of AddAB and Q sepharose column elutions

A. SDS-PAGE analysis showing uninduced and induced cultures for expression of AddAB. Both induced cultures showed bands of sizes appropriate for AddA and AddB (encircled in lanes 2 and 4; AddA is the higher of the two bands). B. SDS-Page analysis of elutions from the Q sepharose column - many of the contaminating lower molecular weight proteins had been lost. Elutions E1- 7 (lanes 1-7) show enrichment of AddAB. C. SDS-Page analysis of elutions from the Q sepharose column – as column progressed there is loss of bands of appropriate size for AddAB (lanes 1-7). M – marker U – uninduced, I – induced, P- pellet AS - ammonium sulphate, supernatant, FT- flow through, W- wash

A.



B.

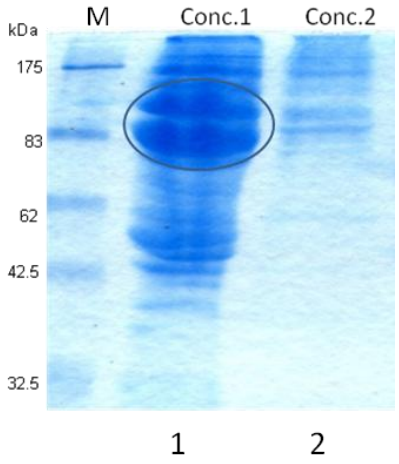


Figure 4.2.2.1.2: SDS-PAGE Gels of Hydroxyappatite column elutions

A. Samples of the collected elutions E1-8 from the hydroxyappatite column were run on a SDS-Page gel (lanes 1-8). The low level of spurious banding in the hydroxyappatite elutions indicated that this step had removed many background contaminating proteins. (M – marker, FT- flow through, D- dialysed sample for loading onto column, C- concentrated Q sepharose elutions 1-7 from previous column used for loading onto hydroxyappatite column)

B. The hydroxyappatite elutions were pooled and concentrated (hydroxyappatite elutions E1-4 = Conc.1 and E5-8 = Conc.2) and run on an SDS-PAGE gel. The concentrated sample – Conc.1 showed the presence of bands of the appropriate size for AddA and AddB (encircled). (M – marker, Conc.1 – concentrated hydroxyappatite elutions E1-4 , Conc.2- concentrated hydroxyappatite elutions E5-8)

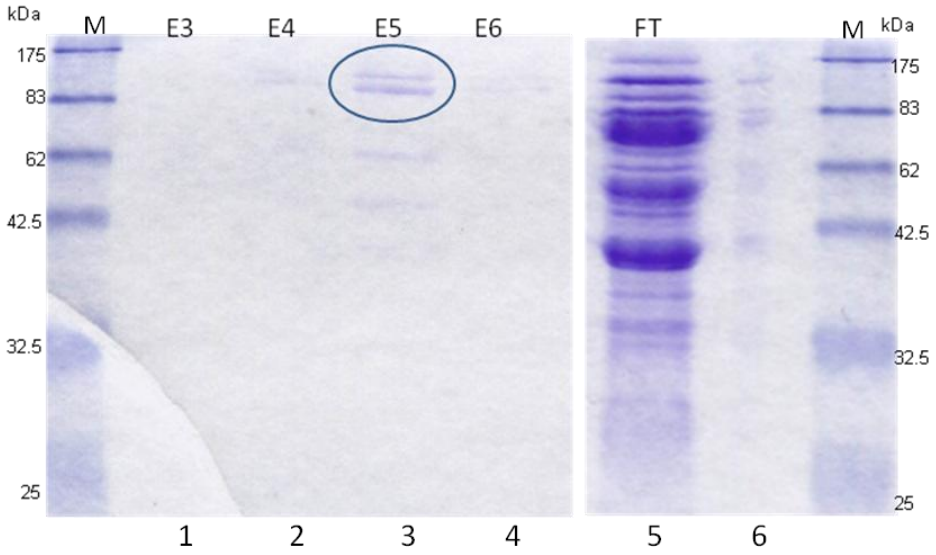


Figure 4.2.2.1.3: SDS-PAGE Gel of Heparin column elutions

SDS-PAGE analysis of samples from the Heparin column elutes. Elution E5 (lane 3) displayed the strongest enrichment for banding of appropriate size for AddAB proteins (encircled). The heparin column had further purified AddAB proteins and removed a large number of background proteins (seen in the flow through – Lane 5).M- marker, FT- flow through from heparin column.

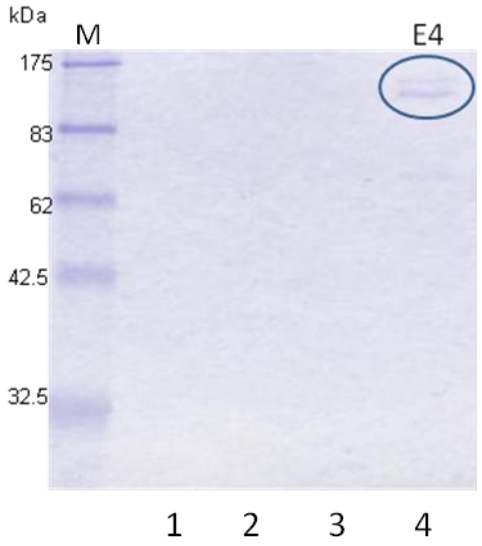


Figure 4.2.2.1.4: SDS-PAGE Gel of mono Q column elutions

Samples of the elutions from the mono Q column were run on a SDS-PAGE gel (lanes 1-4). The E4 elution (lane 4) displayed two bands of appropriate size for the AddAB proteins (encircled). The gel analysis showed that the mono Q column had polished the protein sample and removed the majority of contaminating background proteins, but there was a low yield of AddAB. M- marker.

4.2.2.2 AB Purification Method 2

The method in the previous section involved numerous columns and concentration steps, making the purification lengthy and resulting in loss of proteins and reduction in activity. Therefore another method was devised that would reduce the time involved but increase the yield. Many different trials of which steps to keep and which to remove where considered; below is a description of the final method for purification of AddAB.

The induction and lysis was performed as previously described in section 4.2.2.1, with the substitution of cOmplete™ mini protease cocktail inhibitor (Roche) in place of PMSF. This sample was dialysed as described (section 4.2.2.1) for preparation before loading onto a Q sepharose columnn, and eluted as previously described for the Q sepharose columnn (section 4.2.2.1.). Bands of the appropriate size for AddAB were seen in all elutions (figure 4.2.2.2.1), with enrichment in elutes 7-14. Elutions 7-11 were concentrated as these showed the best enrichment of AddAB; bands 12-14 had higher levels of background protein contamination so were not used. The concentrated Q sepharose column elutes showed enrichment of AddAB with bands of the appropriate size (figure 4.2.2.2.2). The 500µl concentrated sample was loaded on to a Tricorn Gel Filtration column (Tricorn Superose 12 10/300 GL). The gel filtration step removed background proteins (figure 4.2.2.2.2). Elutions 17-19 were concentrated (final concentration 0.12 mg/ml), as were elutions 20-23(final concentration 1.2 mg/ml), and proteins bands from the sample were prepared for mass spectrometry. The results of the mass spectrometry showed that the bands were predominantly AddAB; band 1 – BF2192 AddA, band 2 – BF2191 AddB, band 3 - *E. coli* catalase, band 4 - *E. coli* polyribonucleotide nucleotidyltransferase, band 5 - E. coli isocitrate lyase (figure 4.2.2.2.2). This adaptation of the previous method (section 4.2.2.1) allowed a rapid purification with higher yield of AddAB .

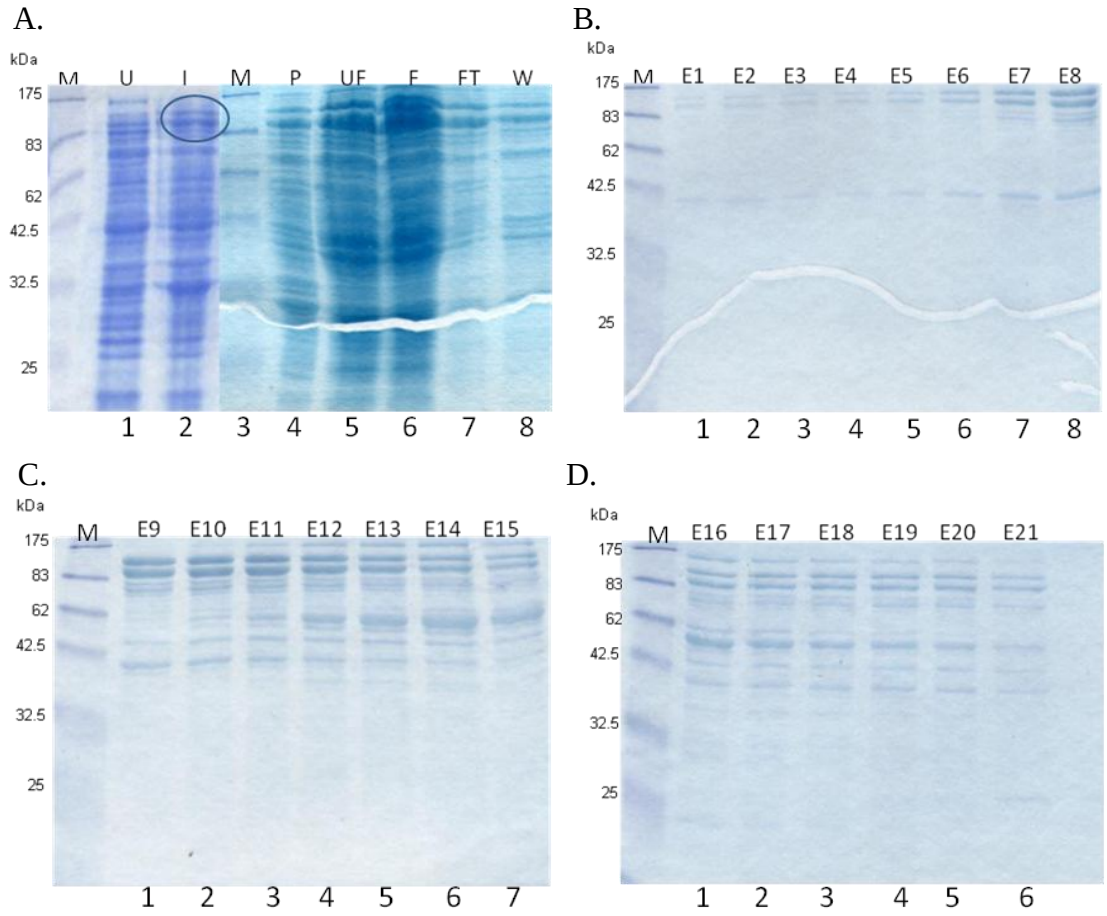


Figure 4.2.2.2.1: SDS-PAGE - Induction of AddAB and Q Sepharose column elutions

A. The induction for AddAB did not appear particularly strong. The induced sample I (lane 2) shows an intenser banding pattern of proteins at the appropriate size for AddA and B (encircled), with the AddB appropriate size band showing the best induction. B, C and D. The Q sepharose column elutions showed bands of appropriate size for AddA and B in all elutions, with strong enrichment in elutes E7-14 (panel B - lanes 7 and 8, panel C – lanes 1 to 6). M – marker, U – uninduced, I – induced, P-pellet, UF- unfiltered sample to be loaded on column, F- filtered sample to be loaded on column, FT-flow through, W- wash

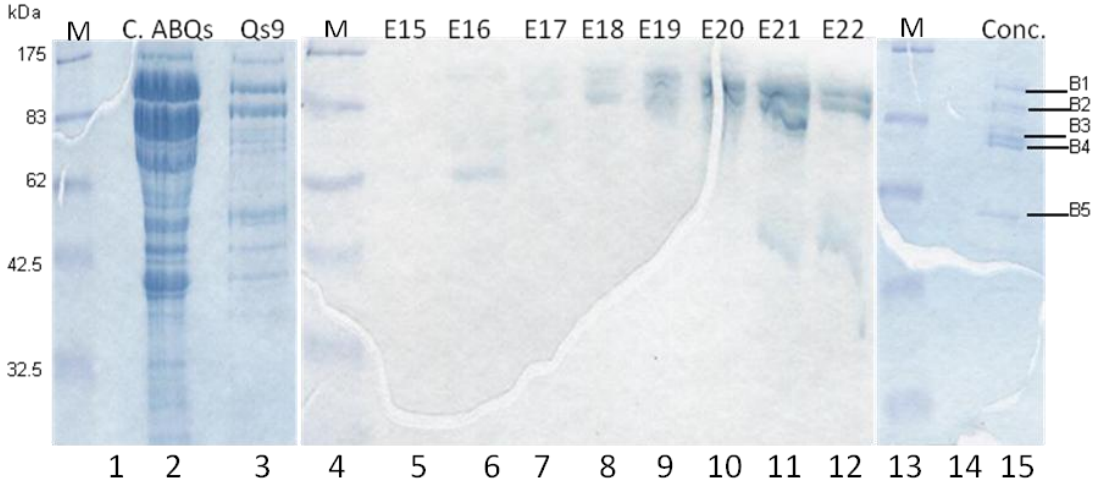


Figure 4.2.2.2.2: SDS-PAGE analysis of AddAB Gel Filtration column elutions

The Q sepharose column elutions E7 -11 (Figure 4.2.2.2.1) were concentrated (C.ABQs) and were loaded on to a gel filtration column (left panel). Only elutions E16 – E22 (lanes 6-12) of the gel filtration column showed appropriate bands for AddA and AddB. The final gel (right panel) showed the concentrated GF column elution (marked as Conc. in lane 15), highlighted bands, B1-B5 (lane 15) were sent for mass spectrometry. M – marker, C.ABQs – concentrated AB elutions from Q sepharose column, Qs9 – elution 9 from Q sepharose column, Conc. – concentration of GF column elutions, B1 – band 1, B2- band 2, B3- band 3, B4, band 4, B5- band 5.

4.2.3.1 Purification of AddA and AddB separately.

The above method used to purify the AddA and AddB proteins separately without tags. The Δ *recBCD E. coli* strain (FP101) harbouring the pFP102 (*addA*) or pFP103 (*addB*) plasmids (see chapter 3, section 3.3.3.4) were used for expression of the AddA or AddB protein, respectively. The cultures were induced and lysed as previously described (section 4.2.2.2). The SDS-PAGE analysis (figure 4.2.3.1.1 and 4.2.3.1.2) showed low levels of induction with bands that were the appropriate size for AddA and AddB; this would support the assumption that expression of the proteins separately allows degradation. The lysates were prepared for the Q sepharose column as previously described (4.2.2.2) with one alteration, at the dialysis step Triton x was added to buffers to a final concentration of 0.1%, to prevent protein aggregation. As a precautionary measure all flow-throughs from columns were kept in case proteins did not bind to the column. Figures 4.2.3.1.1 and 4.2.3.1.2 show elutions from the Q sepharose columns for the AddA and AddB purifications; neither protein was enriched. On examining the flow-through from both the AddA and AddB Q sepharose columns, there was a band of appropriate size for the respective proteins (figure 4.2.3.1.1 and 4.2.3.1.2), therefore the elutions were discarded and the flow-through from the Q sepharose columns was loaded onto a gel filtration column. Elutions from the gel filtration columns showed appropriate size bands for both AddA and AddB in elutes 3-6, but both samples had backgrounds of contaminating proteins (figure 4.2.3.1.3). To reduce the background protein contamination and increase yields of target protein a tagged system was used as this that might offer some protection against degradation as it could occlude proteolytic sensitive sites on the proteins while allowing quick purification of the proteins.

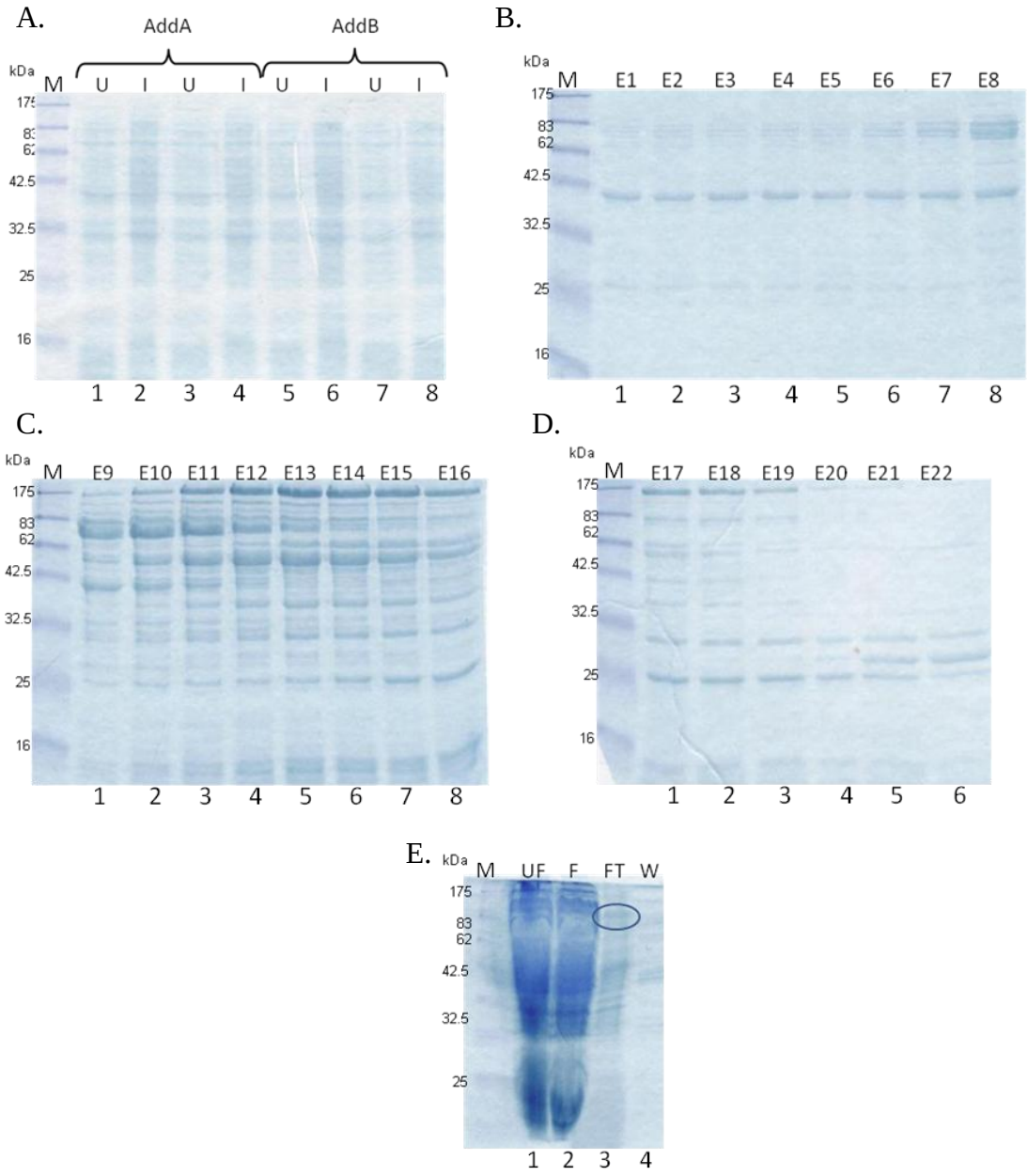


Figure 4.2.3.1.1: SDS-PAGE - Induction of AddA and AddA Q Sepharose column elutions

A. SDS-PAGE showing the induced samples, I, from each of the 1L cultures for AddA (lanes 2,4,6 and 8). The banding overall within the gels were faint, and may in part be due to the quality of the gel. Panels B, C and D show the elutions from the Q sepharose column for the AddA sample. There appeared to be no enrichment for bands of the expected size of AddA. E. SDS-PAGE of unfiltered (UF) and filtered (F) sample for loading onto column, flow through (FT) and wash (W) from the column. In the flow through (lane3) there is a band of appropriate size for AddA (encircled). M – marker, U – uninduced, I – induced, UF- unfiltered sample to be loaded on column, F- filtered sample to be loaded on column, FT- flow through, W- wash.

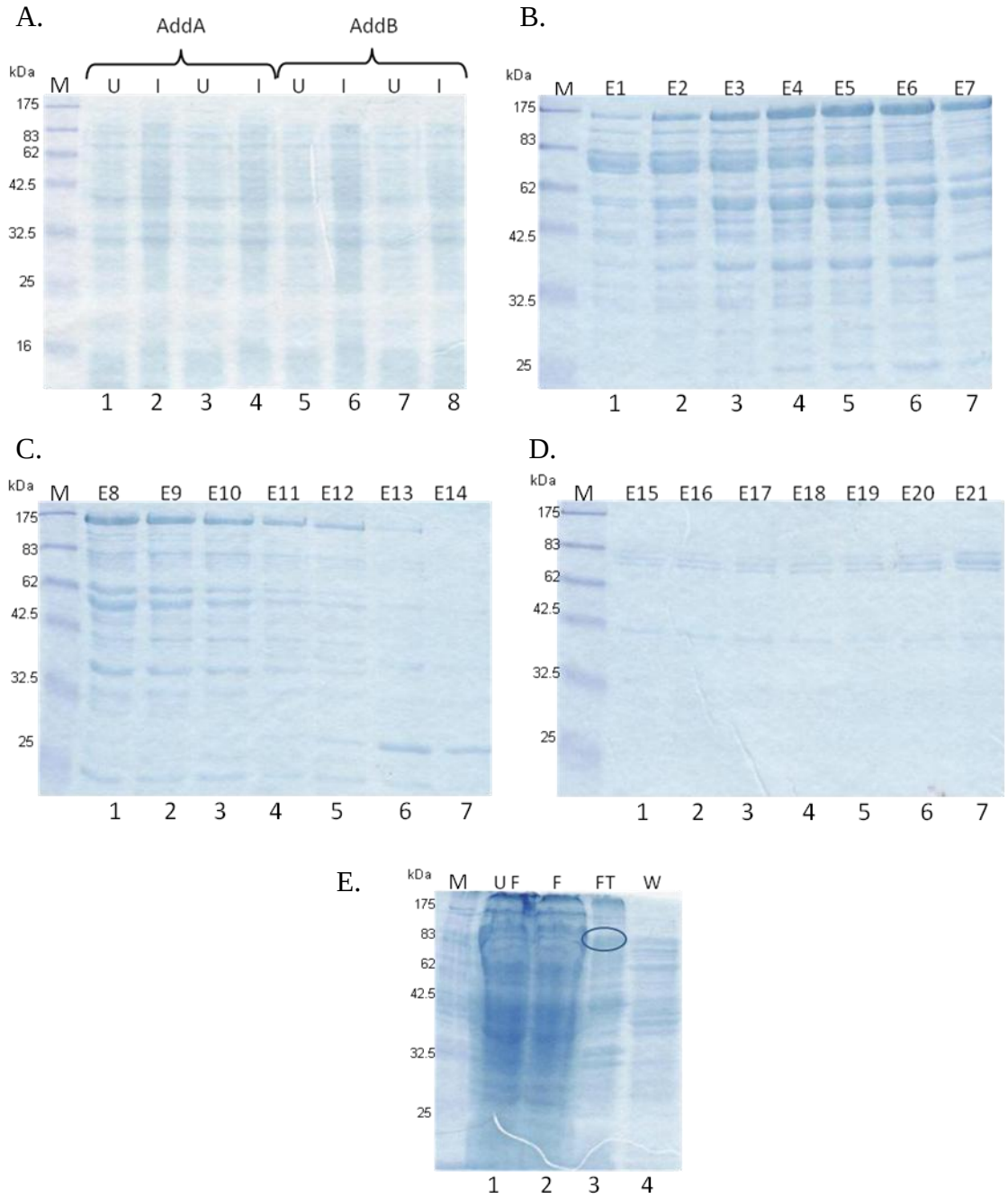


Figure 4.2.3.1.2: SDS-PAGE - Induction of AddB and AddB Q Sepharose column

A. SDS-PAGE showing the induced samples, I, from each of the 1L cultures for AddB (lanes 2,4,6 and 8). The banding overall within the gels were faint, and may in part be due to the quality of the gel. Panels B, C and D show the elutions from the Q sepharose column for the AddB sample. There appeared to be no enrichment for bands of the expected size of AddB. E. SDS-PAGE of unfiltered (UF) and filtered (F) sample for loading onto column, flow through (FT) and wash (W) from the column. In the flow through (lane3) there is a band of appropriate size for AddB (encircled). M – marker, U – uninduced, I – induced, UF- unfiltered sample to be loaded on column, F- filtered sample to be loaded on column, FT- flow through, W- wash.

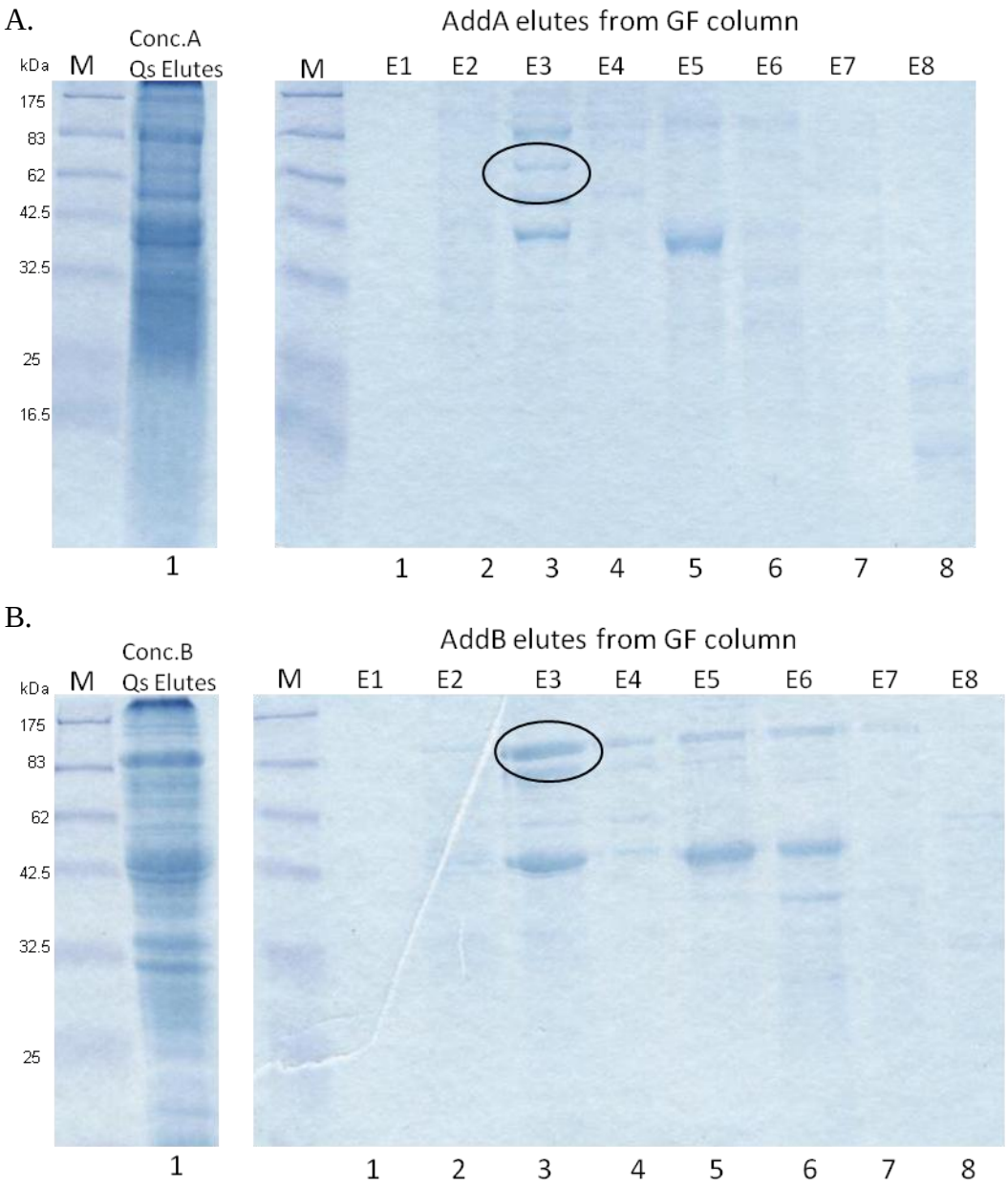


Figure 4.2.3.1.3: SDS-PAGE - AddA and AddB GF column elutions

SDS-PAGE showing the elutions from the gel filtration columns for the AddA (panel A RHS) and AddB (panel B RHS) purification and the concentrated Q sepharose column elutions that were loaded onto the GF column (panel A and B LHS gel). Though there were bands present at the appropriate size for AddA in elutions E3-E6 (panel A RHS gel lanes 3-6) and AddB elutions E2-E8 (panel B RHS gel lanes 2-8) in the each of the respective columns there was a large amount of contaminating background proteins. M – marker, Conc. A Qs elutes- concentrated flow through from Q sepharose AddA column, Conc. B Qs elutes- concentrated flow through from Q sepharose AddA column.

4.2.4.1 Purification of AddA-MBP and AddB-MBP

To try to enhance the final yield of AddA and AddB, maltose binding protein (MBP) fusions were used. It was hoped that the tag would confer protection from degradation. Once the proteins had been purified using an amylose column, the tag would be removed and the proteins purified useing a Q sepharose column followed by a gel filtration column to further purify the proteins. To produce N-terminal MBP fusions, the *addA* and *addB* genes were ligated into pMAL-c2x (NEB) giving plasmids pFP107 and pFP108, respectively (figures 4.2.4.1.1 and 4.2.4.1.2). The AddA-MBP fusion protein had a predicted size of 161kDa and the AddB-MBP fusion protein had a predicted size of 149kDa.

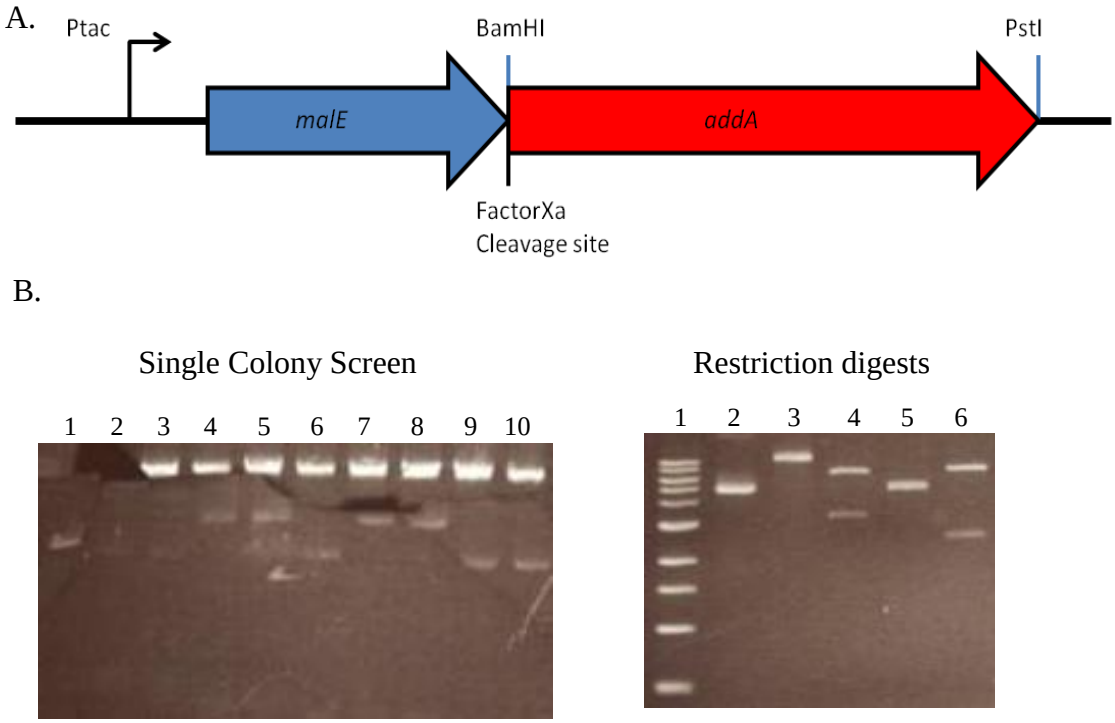
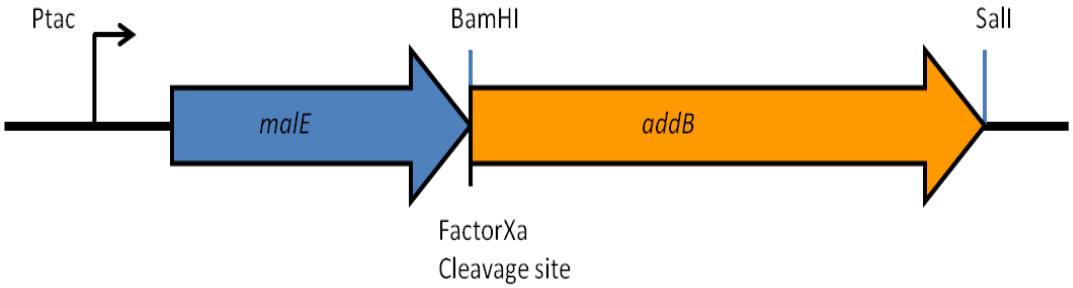


Figure 4.2.4.1.1: Production of pFP107

Construction of the AddA-MBP expression vector pFP107. A. The *addA* gene was amplified by PCR and ligated to the pMAL-c2x plasmid to produce a fusion protein containing MBP fused to the N terminus of AddA. B. Agarose gel showing single colony screen (LHS panel), lanes 4, 5, 7 and 8 showed a plasmid of larger size than the pMAL-c2x control (lane 1) indicating the presence of an insert. The restriction enzyme digests (RHS panel) confirmed the presence of insert in the plasmid. Lane 1 – 1kb ladder, Lane 2 – pMal plasmid, lane 3 - pFP107 linearised by NcoI, lane 4 – pFP107 BamHI/PstI digest releases 3.2kb *addA* insert and 6.6kb plasmid, lane 5 – pFP107, lane 6 – pFP107 NdeI digest releases two fragments 2.5kb and 7.3kb.

A.



B.

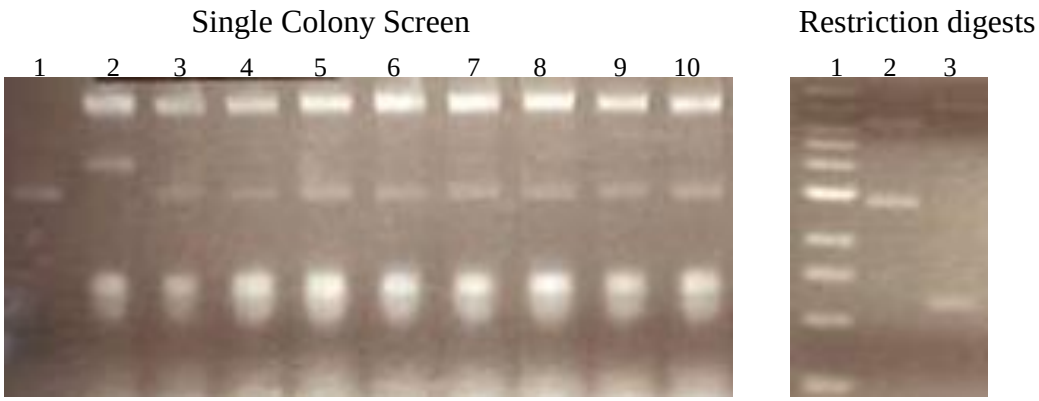


Figure 4.2.4.1.2: Production of pFP108

Construction of the AddB-MBP expression vector pFP108. A. The *addB* gene was amplified by PCR and ligated to the pMAL-c2x plasmid to produce a fusion protein containing MBP fused to the N terminus of the AddB protein. B. Agarose gel showing single colony screen (RHS panel), lane 2 of the single colony screen showed a plasmid of larger size than the pMAL-c2x control (lane 1) indicating the presence of an insert (lanes 3-10 show plasmids with no integration of insert). The restriction enzyme digests (RHS panel) confirmed the presence of insert in the plasmid; Lane 1 – 1kb lader, lane 2 – BamHI/SalI digest of pFP108 releasing the 2.8kb *addB* insert and the 6.6kb plasmid, lane 3 – HindIII digest of pFP108 giving 1.2kb and 2.84kb fragments.

The *E. coli* Δ *recBCD* strain harbouring either the pFP107 or pFP108 plasmid was used for expression of the AddA-MBP or AddB-MBP proteins, respectively. The cultures were induced as previously described (section 4.2.2.1) and protein lysates prepared and purified as described in chapter 2 (section 2.8.6). Twenty one, 1ml fractions were collected and analysed by SDS PAGE. Appropriate fractions were concentrated and passed through a gel filtration column as previously described.

Induction of AddA-MBP and AddB-MBP was low (figure 4.2.4.1.3 and 4.2.4.1.4) but appeared higher when compared to the respective proteins without a tag (4.2.3.1). The AddA-MBP and AddB-MBP lysates were loaded onto amylose columns. The amounts of protein detected in the elutions were low (Figure 4.2.4.1.3 and 4.2.4.1.4) but higher than previously found with using the Impact system (section 4.2.1.3) or purifying proteins without a tag (section 4.2.3.1). The yield of purified AddB-MBP compared to AddA-MBP was higher, even though the induced levels for both proteins appeared the same, so perhaps loss of AddA-MBP was due to the protein being more sensitive to degradation. In both purifications, it was observed that the best enrichment of the desired proteins was in the earlier elutes; AddA-MBP elutes 2-7 and AddB-MBP elutes 3-8 were concentrated from 6ml to 0.5ml. To remove MBP, factor Xa was used, but it did not cleave either protein elutes even when used with different concentrations of SDS. Unfortunately the lack of cleavage meant that the purified AddA and AddB had an additional polypeptide of 42kDa attached. This could impede the activity of either or both of the proteins in helicase/nuclease assays, but also could offer protection from proteolytic degradation. Each concentrated sample was then loaded on to a gel filtration column, however the elutes for both samples showed no proteins of the appropriate sizes. One possibility was that the proteins had dimerised or oligomerised and had eluted in the void volume. To check for this, the void elutes were analysed by SDS PAGE and were found to contain proteins of appropriate size for AddA-MBP and AddB-MBP (Figure 4.2.4.1.5). These fractions were concentrated using vivaspin columns with the addition of 0.1% triton to prevent aggregation and loss of protein on the filter. The concentrated samples were loaded onto a gel filtration column, elutes were collected including the void sample (Figure 4.2.4.1.6). As with the previous gel filtration column, the proteins only eluted in the void – suggesting they were oligomerising. This resulted in pure samples of AddA-MBP and AddB-MBP.

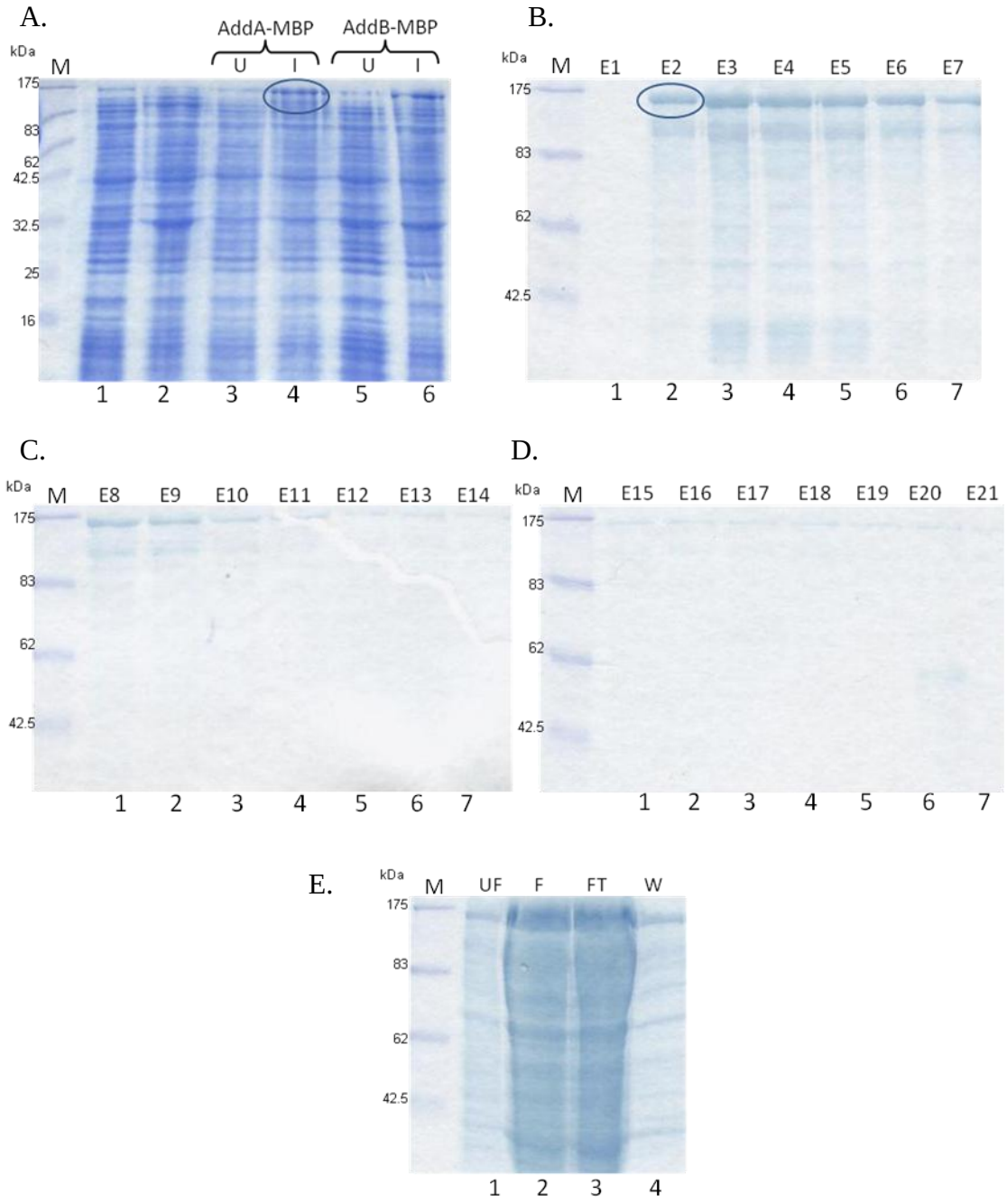


Figure 4.2.4.1.3: SDS-PAGE - Induction of AddA-MBP and Amylose Column elutions

A. SDS-PAGE showing uninduced (U-lane 3) and induced (I-lane4) culture for expression of AddA-MBP (appropriate sized band encircled). B, C and D. The amylose column elutions E1 to 21 (panel B, lanes 1-7, panel C lanes 1-7 and panel D lanes1-7) all show an appropriate size band for AddA-MBP. E. SDS-PAGE of unfiltered (UF) and filtered (F) sample for loading on to amylose column, also shown flow through (FT) and wash (W) from amylose column. M- marker, U – uninduced, I – induced, UF- unfiltered sample to be loaded on column, F- filtered sample to be loaded on column, FT- flow through, W- wash.

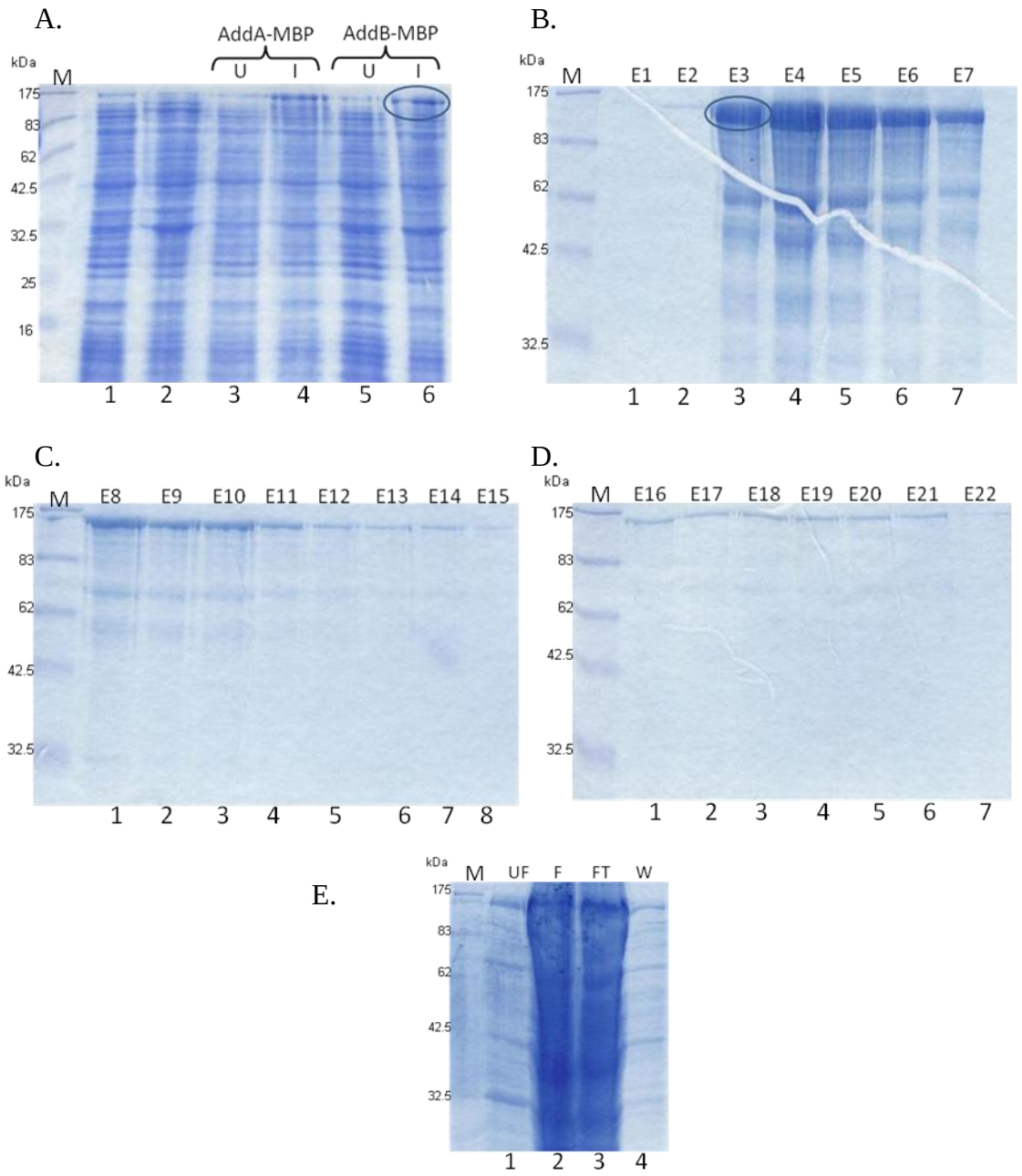


Figure 4.2.4.1.4: Induction of AddB-MBP - Amylose Column

A. SDS-PAGE showing uninduced (U) and induced (I) cultures of expression of AddB-MBP (appropriate sized band encircled). B, C and D. The amylose column elutions 1 to 22 (panel B, lanes 1-7, panel C lanes 1-8 and panel D lanes1-7) all show appropriate size bands for AddB-MBP. E. SDS-PAGE of unfiltered (UF) and filtered (F) sample for loading on to amylose column. M –marker, FT - flow through from amylose column, W - wash W from amylose column U – uninduced, I – induced, UF- unfiltered sample to be loaded on column, F- filtered sample to be loaded on column, FT- flow through, W- wash.

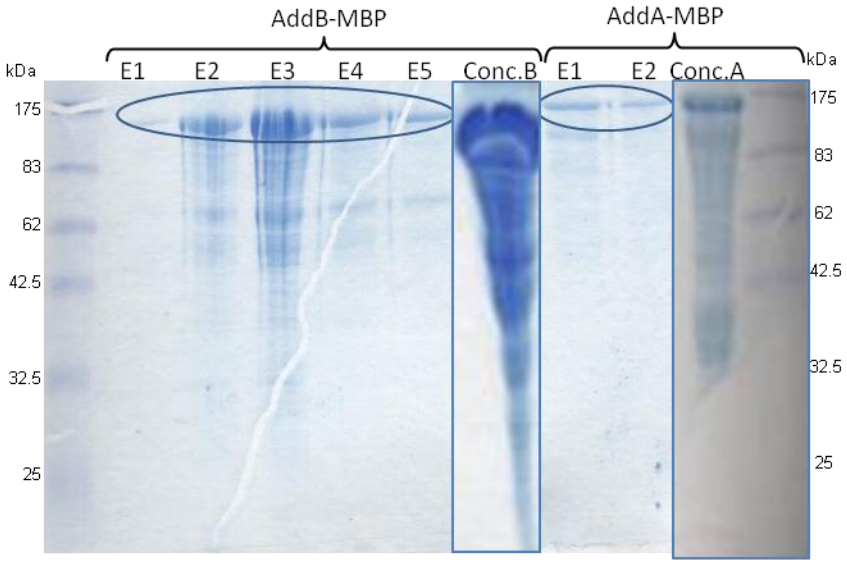


Figure 4.2.4.1.5: 1st GF column for AddB-MBP and AddA-MBP

SDS-PAGE analysis showing the concentrated amylose column elutions for AddB-MBP and AddA-MBP (Conc.B and Conc.A) that were loaded onto a GF column. Proteins of appropriate size for AddB-MBP and AddA-MBP and were seen only in the void elutes (encircled E1-5 for AddB-MBP and encircled E2-3 for AddA-MBP). Conc.A – concentrated AddA-MBP amylose column elutions Conc.B – concentrated AddB-MBP amylose column elutions

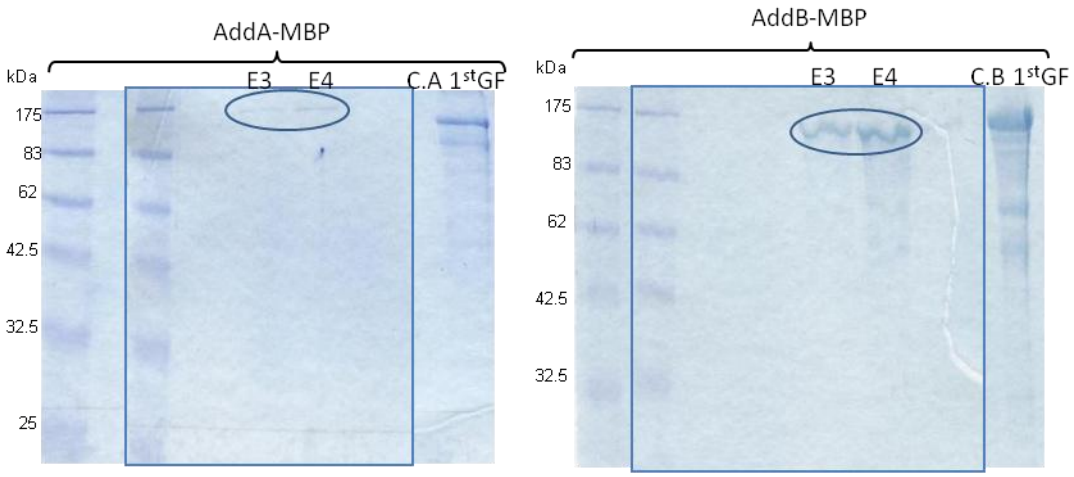


Figure 4.2.4.1.6: 2nd GF column for AddB-MBP and AddA-MBP

SDS-PAGE analysis showing the AddA-MBP and AddB-MBP concentrated elutions from the first GF columns (CA.1stGF and CB.1stGF) that were loaded onto the second GF column. Proteins of appropriate size for AddA-MBP and AddB-MBP were only seen in the elutes (E3 and E4) from the void.

4.3 Biochemical analysis of AddAB

The purification work in the previous sections provided 3 protein samples for further functional analysis: AddAB, AddA-MBP and AddB-MBP. This section examines the biochemical action of the proteins to address the questions of whether the proteins only act in concert or can act independently, and which proteins have helicase and/or nuclease functions. The preferred substrate of AddAB was investigated, to establish if the complex acted on linear or circular DNA and if there was an end preference. The requirement for ATP and metal cofactors was also investigated.

4.3.1 AddAB action on Linear and Super-coiled DNA

To determine if *B. fragilis* AddAB has exonuclease or endonuclease activity, two DNA substrates were examined; super-coiled ds DNA and linear ds DNA (Figure 4.3.1.1). The super-coiled DNA used was pUC19 and the linear DNA was pUC19 digested with SmaI to give blunt-ended ds DNA. Since helicases require an energy source and nucleases require metal ions, all reactions contained ATP and magnesium.

Reactions were performed in nuclease buffer (25mM Tris-acetate pH7.5, 1mM DTT, 1mM ATP, 2mM $Mg(Oac)_2$, 10nM DNA, made up to 20µl volume with dH_2O). Reactions were pre-warmed at 37^{o}C before addition of AddAB to a final concentration of 0.5nM. Reactions were stopped at various time points by adding an equal volume of nuclease termination buffer (100mM EDTA, 2.5% SDS w/v, made up to volume with 6x DNA loading Buffer). The samples were electrophoresed through a 1.2% agarose gel.

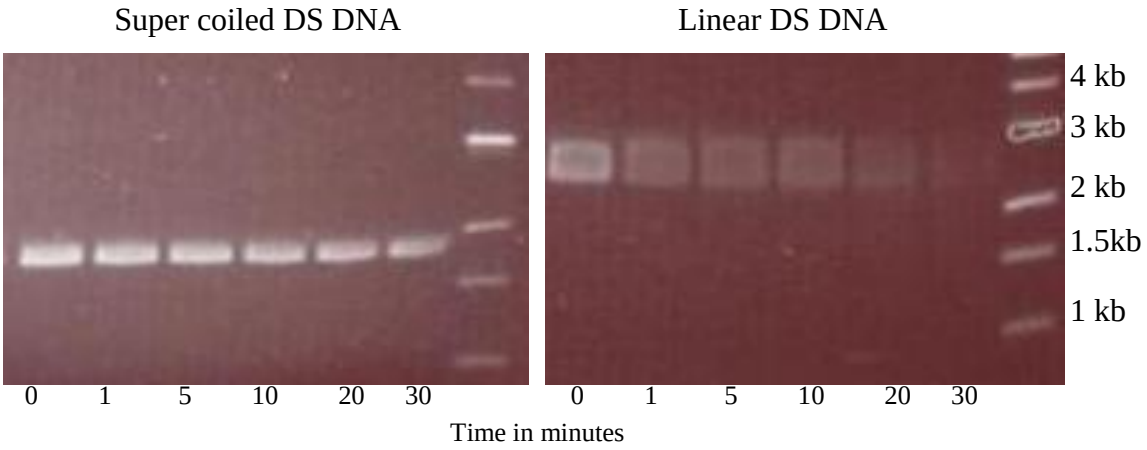


Figure 4.3.1.1: Action of AddAB on super-coiled or linear double-stranded DNA

Agarose gels showing the action of AddAB on two different substrates; super-coiled ds DNA or linear ds DNA. The RHS panle shows that there was no degradation of the super-coiled ds DNA by AddAB at any time point. The LHS panel shows that linear ds DNA was degraded by AddAB as time progressed, shown by the reduction in intensity of the bands over time.

In the presence of AddAB there was no degradation of the super-coiled DNA within 30 minutes however, the linear DNA was degraded almost completely by the end of the time course (Figure 4.3.1.1). This result shows that AddAB is not an endonuclease but does have exonuclease activity, and can bind to blunt dsDNA ends.

4.3.2 AddAB – co-factor requirements

E. coli RecBCD and *B. subtilis* AddAB require ATP and magnesium for efficient exonuclease and helicase activity. To determine if helicase activity was essential for DNA degradation by *B. fragilis* AddAB the ATP requirement of the complex was tested (Figure 4.3.2.1). Reactions were set up with or without ATP and with different concentrations of magnesium. It should be noted that the reaction buffer contained single-stranded DNA binding protein (SSB) to prevent reannealing of strands. The DNA substrate for each reaction was pUC19 cut with EcoRI, giving a 4 base overhang. Reactions were performed in nuclease buffer as previously described in section 4.2.1. The samples were electrophoresed through a 1.2% agarose gel.

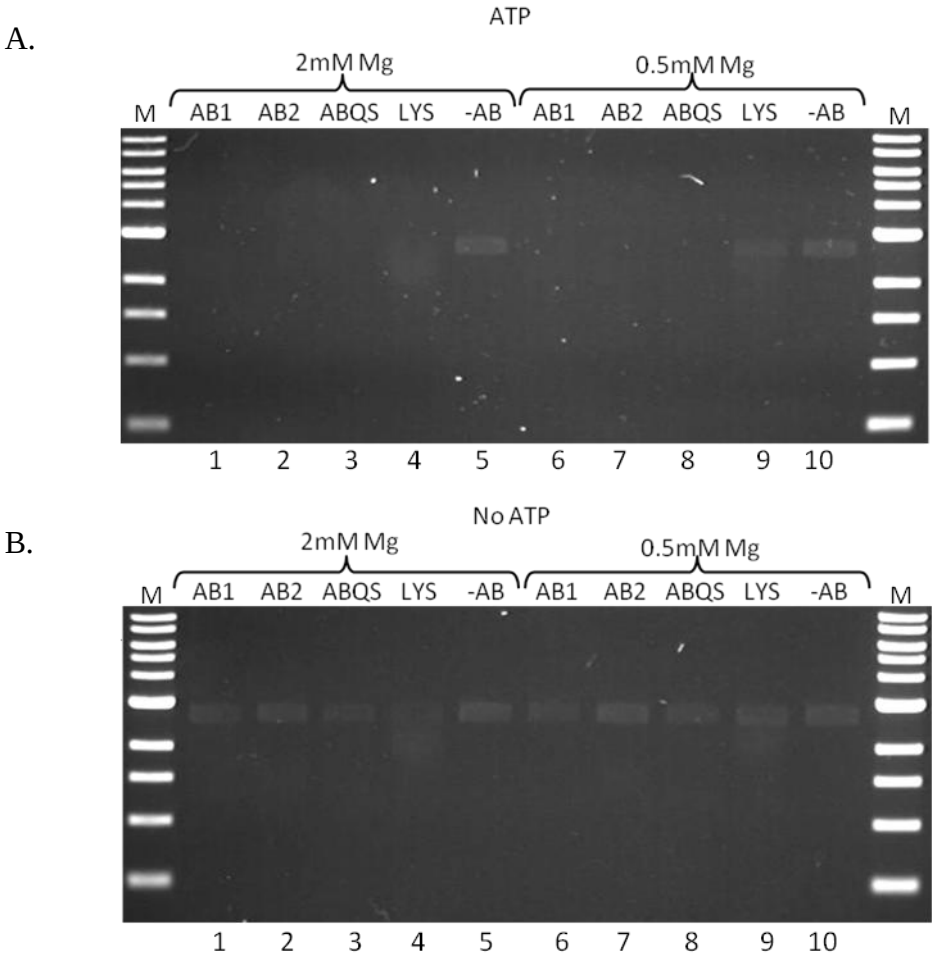


Figure 4.3.2.1: Exonuclease assay of AddAB

Agarose gel showing exonuclease activity of AddAB with (Panel A) or without ATP (Panel B). A. When ATP is present all AddAB samples (AB1, AB2 and ABQs) showed degradation of the DNA substrate regardless of the magnesium level (lanes 1, 2, 3, 6, 7, and 8). The lysate samples (lanes 4 and 9) showed degradation but not at as great a level as purified samples of AddAB. The no AddAB controls (lanes 5 and6) have no degradation of the DNA. B. The absence of ATP in a sample inhibits the exonuclease activity of AddAB (lanes 1, 2, 3, 6, 7 and 8), except in lysate samples (lanes 4 and 9) where some degradationof DNA is present.. AB1 and AB2 – AddAB samples from the gel filtration purification (section 4.2.2.2), ABQS – elution 9 from the Q sepharose column (section 4.2.2.2), Lys – lysate (section 4.2.2.2), -AB – control with no AddAB present, M – 1kb DNA ladder.

DNA degradation was not detected in the reactions without ATP, except in the lysate reactions; this may be due to the lysates containing ATP or other nucleases. Reactions containing ATP and AddAB showed degradation of DNA regardless of the level of Mg^{2+}. The lysate reactions in both conditions whether ATP was present or not, showed a greater degradation with the higher level of Mg^{2+}. Despite the presence

of SSB, there was no evidence of ss-DNA on the gels. These experiments demonstrated that *B. fragilis* AddAB can bind and degrade DNA with a short 4bp overhang in an ATP dependant manner, indicating that helicase activity is required for exonuclease function.

4.3.3 AddAB – Magnesium co-factor

In previous experiments degradation of linear DNA by AddAB occurred at both high and low concentrations of magnesium. To further examine the effects of magnesium concentration on AddAB activity a protein titration was performed (Figure 4.3.3.1). Reactions were prepared in nuclease buffer. Reactions were pre-warmed at 37^{o}C before addition of AddAB at different concentrations. The samples were incubated for 10 minutes before addition of nuclease termination buffer (Section 4.3.1). Samples were electrophoresed through a 1.2% agarose gel.

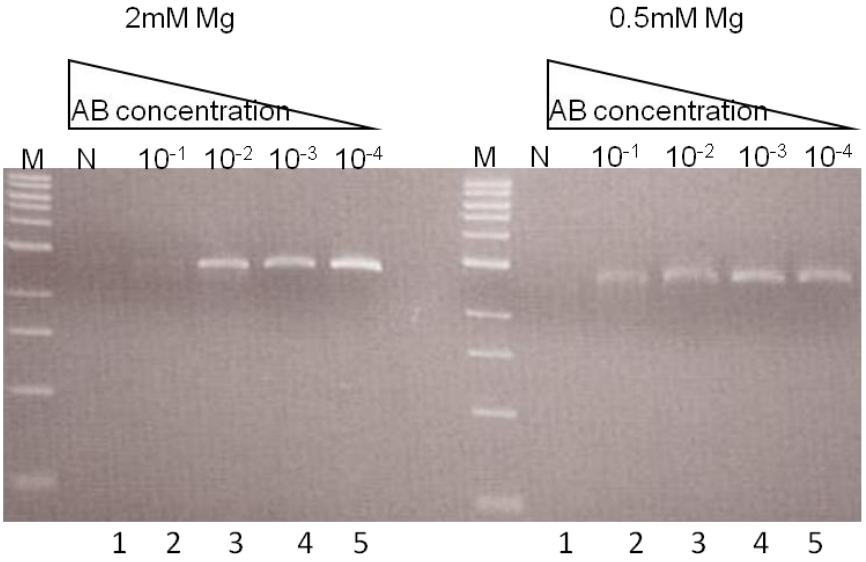


Figure 4.3.3.1: Titration of AddAB

Agarose gel showing titration of AddAB and its activity on linear DNA at two concentrations of Mg^{2}. The undiluted sample (N – lane 1 of both the RHS and LHS panels) contained AddAB to a final concentration of 0.25nM, each subsequent reaction was a 1/10 dilution of AddAB concentration (lanes 2-5 of both the RHS and LHs panels). At both the high and low concentrations of Mg^{2+} degradation of DNA was apparent in the reactions containing the neat samples of AddAB (lanes 1 of both gels). At 10^{-1} dilution of AddAB (lane 2 of both gels) the degradation appears stronger in the reaction with the higher level of magnesium (RHS gel). M- 1kb lDNA ladder, N-neat undiluted sample.

Both concentrations of magnesium allowed degradation of DNA to proceed, however the higher concentration of Mg^{2+} gave significant degradation at 0.25nM AddAB, whereas the lower concentration of Mg^{2+} did not show a similar level of degradation. This shows that while both magnesium concentrations allow degradation, the higher concentration increases the processivity of AddAB.

4.3.4 Effect of Calcium on degradation

Calcium has been shown to inhibit RecBC exonuclease function (Rosamond *et al*., 1979). To test the effect of calcium on AddAB function, $CaCl_2$ was added to a final concentration of 1mM. This concentration was reported to have the greatest effect on RecBC nuclease activity while still allowing ATPase activity (Rosamond *et al*., 1979).

Reactions were performed in nuclease buffer, with or without Ca^{2+} at 1mM. The DNA substrate was pUC19 cut with EcoRI. Reactions were pre-warmed at 37^{o}C before addition of AddAB to a final concentration of 0.5nM. The samples were incubated for 10 minutes before addition of nuclease termination buffer (Section 4.3.1). The samples were electrophoresed through a 1.2% agarose gel (figure 4.3.4.1).

Lane	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Ca^{2+} (1mM)		-	-	-	-		-	-	-	-		+	+	+	+
Mg^{2+}(0.5mM)		-	-	-	-		+	+	+	+		-	-	-	-
Mg^{2+}(2mM)		+	+	+	+		-	-	-	-		+	+	+	+
Linear DNA		+	+	+	-		+	+	+	-		+	+	+	-
Supercoiled DNA		-	-	-	+		-	-	-	+		-	-	-	+
ATP		+	-	+	+		+	-	+	+		+	-	+	+
AddAB		-	+	+	+		-	+	+	+		-	+	+	+

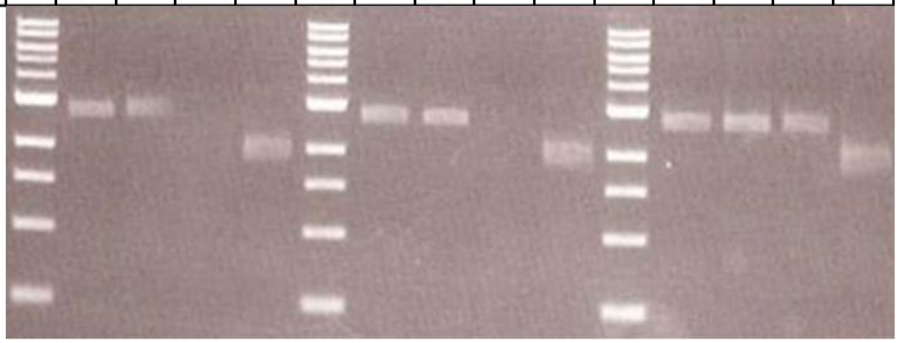


Figure 4.3.4.1: Action of AddAB in the presence of calcium

Agarose gel showing the action of AddAB on linear ds DNA is dependent on the metal co-factor present. At both the high and low level of Mg^{2+} there was degradation of linear ds DNA (lanes 4 and 9). The presence of calcium inhibited AddAB activity on linear ds DNA (lane 14). Lanes 1, 6, and 11 contain 1kb DNA ladder, lanes 2,7 and 12 were controls where AddAB was absent, lanes 3,8 and 13 were controls where ATP was absent. Lanes 5, 10 and 15 were controls containing supercoiled DNA.

In reactions with both concentrations of Mg^{2+} AddAB mediated degradation was only observed when ATP was present and the substrate was linear DNA (lanes 4 and 9). When calcium was present in the reaction no degradation of linear substrate was observed. This showed that Calcium inhibited the exonuclease activity of AddAB.

4.3.5 Determining exonuclease activities of AddA-MBP and AddB-MBP

To test whether AddA or AddB could act independently, the nuclease activity of each protein was tested (Figure 4.3.5.1). The MBP fusion proteins AddA-MBP and AddB-MBP were used for these assays (section 4.2.4.1). Reactions were performed as previously described (section 4.3.3). High Mg^{2+} concentration (final concentration 2mM) was used since this would ensure optimal nuclease activity. The DNA

substrate was pUC19 linearised with EcoRI. Reactions were pre-warmed at 37°C before addition of AddA-MBP or AddB-MBP to different final concentrations. The samples were incubated for 10 minutes before addition of nuclease termination buffer. Samples were electrophoresed through a 1.2% agarose gel (figure 4.3.5.1).

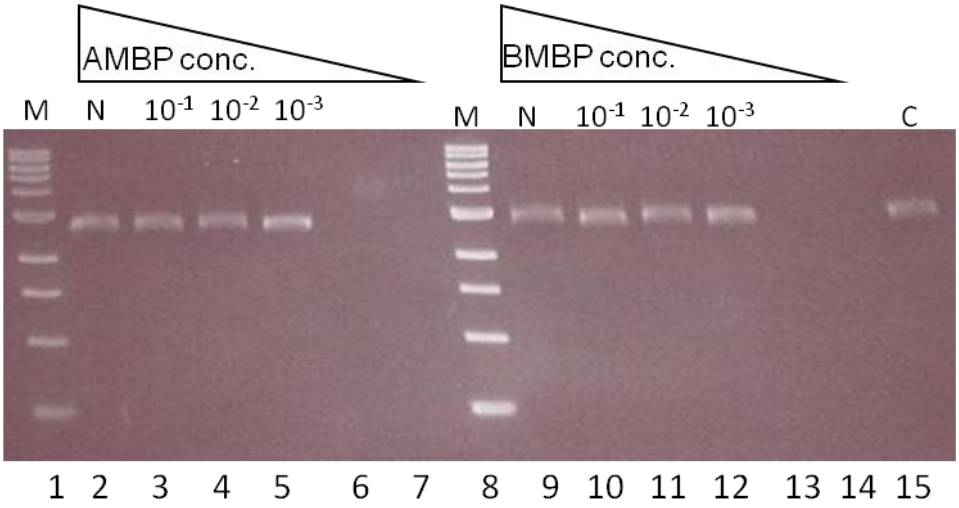


Figure 4.3.5.1: AMBP and BMBP Protein nuclease titration assay

Agarose gel showing effect of titrating AddA-MBP or AddB-MBP on linear DNA. The undiluted samples (N) contained AddA-MBP (lane 2) and AddB-MBP (lane 9) to a final concentration of 0.25nM, each subsequent reaction was a 1/10 dilution of concentration (lanes 3-5 for AMBP and lanes 10-12 for BMBP). At all concentrations of AddA-MBP or AddB-MBP no degradation of DNA was apparent. Lanes 1 and 8 contain 1kb DNA ladder, lanes 6, 7, 13 and 14 are empty, lane 15 is a control containing reaction without enzyme present.

No degradation of plasmid DNA was observed under any of the reaction conditions, which indicated that the individual proteins did not display nuclease activity. The proteins may therefore require each other for formation of a functional nuclease. Other possibilities, however, are that the purification process may have rendered the proteins inactive or the MBP tag may have inhibited activity by occluding the nuclease site or by causing oligomerisation of the proteins and therefore preventing activity. An attempt to reconstitute the AddAB complex by mixing the MBP fusion proteins did not yield an active exonuclease (data not shown), which suggests the MBP tags were inhibiting enzyme activity.

4.3.6 AddAB helicase Assay

To examine the helicase action of the complex and its separate protein components, the activities of AddAB, AddA-MBP and AddB-MBP were tested on radio-labelled plasmid DNA (Figure 4.3.6.1). This was designed to detect single-stranded DNA generated by the helicase actions of AddAB. The addition of SSB to the reactions should stabilise the unwound product and allow detection of the radiolabelled DNA strands. The pUC19 plasmid was cut with HindIII and then 5′ end labelled with ^{32}P (chapter 2, section 2.7.14). The nuclease reactions were performed as previously described (section 4.3.1) with one concentration of Mg^{2+} (2mM). The reactions were stopped at 5 minute intervals and samples were electrophoresed through a 1.2% agarose gel. The gel was dried onto nitrocellulose and exposed to film (Amersham Biosciences Hyperfilm-HP) overnight at -70^{o}C with an intensifying screen (Dupont Cronex Lightning-Plus), the film was developed in a Konica SRX-101A developer.

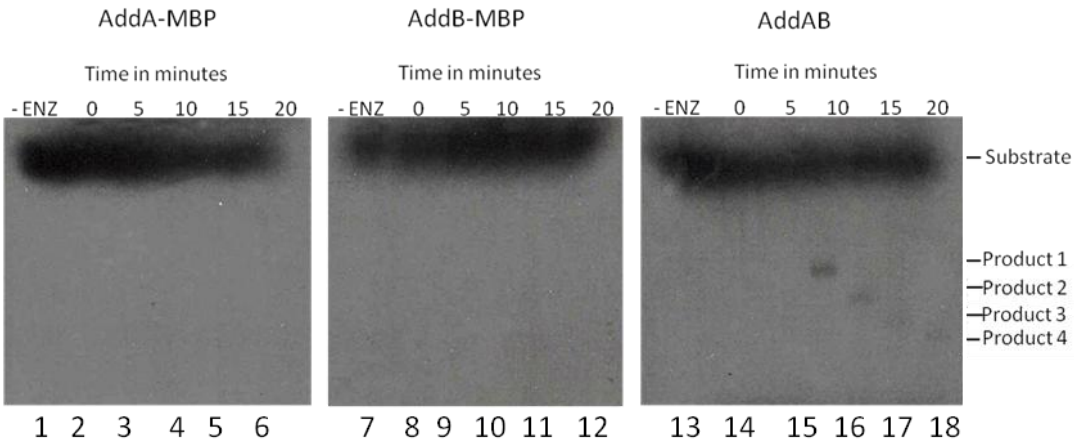


Figure 4.3.6.1: Nuclease assay

Autoradiographs of radiolabelled linear pUC19 digested with AddA-MBP, AddB-MBP or AddAB. Only the AddAB samples showed degradation of DNA as time progressed (lanes 15-18). Lanes 1, 7 and 13 contain no enzyme controls. There is no degradation of DNA in the AddA-MBP or the AddB-MBP reactions (lanes 2-6 and lanes 8-12 respectively).

The autoradiographs for reactions containing AddA-MBP and AddB-MBP showed no evidence of degradation of the DNA substrate. One interpretation of these data would be that AddA and AddB cannot act independently of each other. Another possibility is that because AddA and AddB have a MBP tag attached, this could be

inhibiting the reaction, either by preventing access to the DNA or occluding the active sites on the proteins. The AddAB autoradiograph displayed very faint smearing, indicative of degradation; interestingly there were also four faint bands of distinct sizes detected at different times (at 5, 10, 15, and 20 minutes). This was unexpected since only single-stranded DNA or a smear were anticipated. A possible explanation of this result is that pUC19 may contain a chi site that would prevent or reduce degradation of the DNA. The reduction in size of the fragment as time progressed was perhaps due to degradation from the other end of the DNA.

4.3.7 AddAB is an exonuclease and helicase

To further elucidate the activity of the AddAB complex on ds DNA breaks, radiolabelled substrates were used (Figure 4.3.7.1). Two types of ds DNA substrates were produced. Oligonucleotide pairs were used so that one substrate contained a 40bp 5′ overhang while the other substrate had a 40bp 3′ overhang. The substrates were labelled with 32P and the nuclease/helicase reaction was performed as previously described (section 4.3.6). The reactions were electrophoresed through a 6% polyacrylamide gel for analysis; the gel was dried filter paper prior to exposure to film.

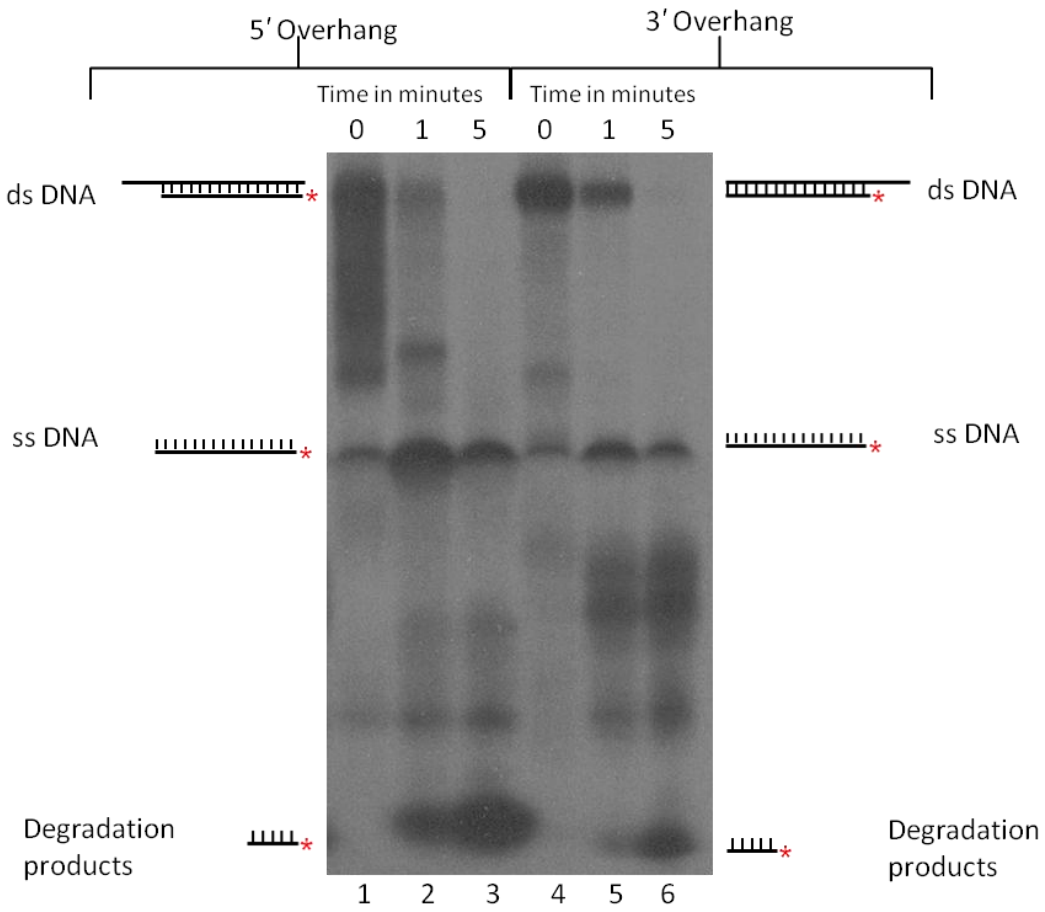


Figure 4.3.7.1: AddAB – Assay of exonuclease and helicase activities

Autoradiograph showing AddAB activity on 5′ overhang and 3′ overhang substrates; Oligonucleotide pairs (Chapter 2: table 2.5 - 5′ overhang pair = RD5overhangHelicase3/RD5overhangHelicase5 and 3′ overhang pair = RD3overhangHelicase3/RD3overhangHelicase5) were used so that one substrate contained a 40bp 5′ overhang (lanes 1, 2 and 3) while the other substrate had a 40bp 3′ overhang (lanes 4, 5 and 6) (Oligonucleotides taken from Dillingham *et al*., 2003). As time progresses both DNA substrates are unwound and degraded by AddAB - Experiment performed in conjunction with G. Blakely.

Two ds DNA substrates were used, each with a strand specific single-strand DNA overhang at one end and a blunt break at the other. AddAB was able to unwind and degrade both substrates. At 1 minute both reactions showed a product equivalent to single-stranded DNA, that implied unwinding of ds DNA was occurring. By 5 minutes the majority of substrate had been unwound and degraded, with a complete loss of the double-stranded DNA and an accumulation of small radiolabelled

products. These results indicated that at least one of the nucleases had 3-5′ activity, since only products with an intact 5′ would be detected (figure 4.3.7.2)

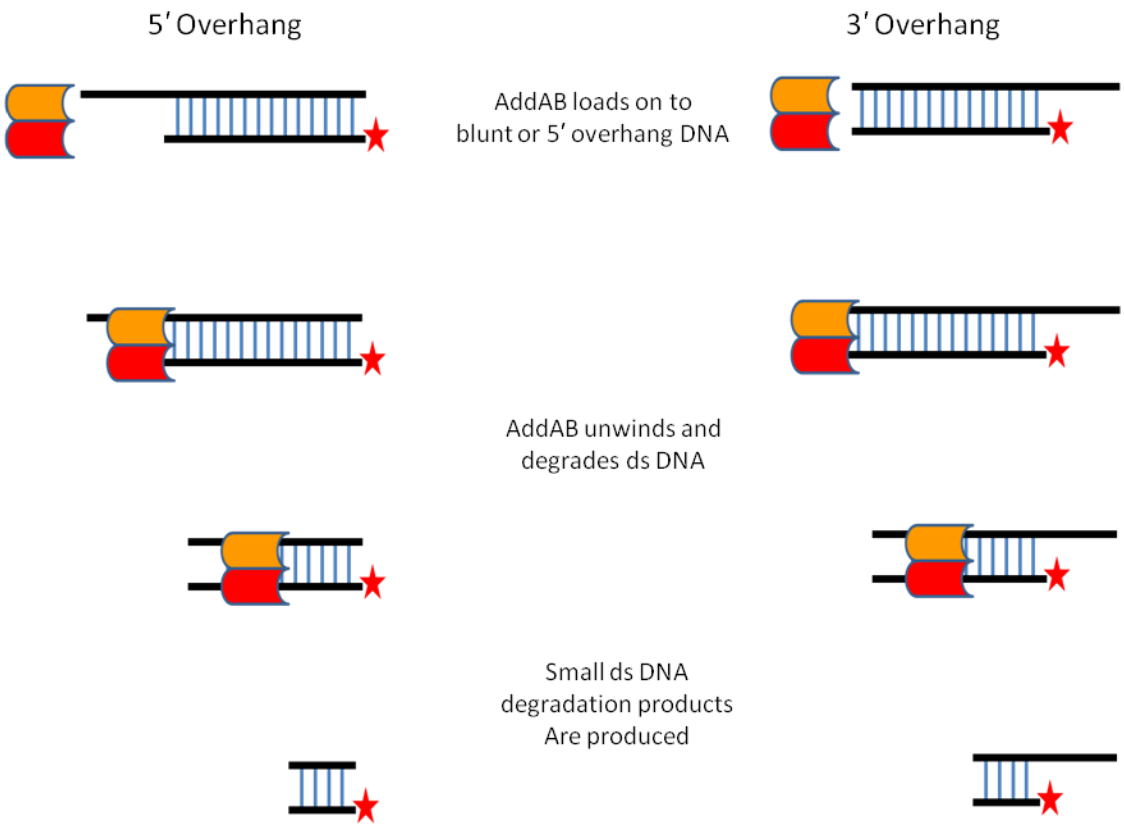


Figure 4.3.7.2: Representation of AddAB action on DNA substrates

AddAB loads onto 5′ overhang and/or blunt ended DNA. The complex translocates along the DNA unwinding and degrading the strands, resulting in small 5′ labelled products.

4.4 Summary

This work developed a rapid method for purification of active AddAB. Low expression may have been due to degradation of the proteins, however co-purification of the complex provided protection from degradation. Using N terminal MBP fusion proteins provided a method for purification of the separate components of the complex, while at the same time overcoming the degradation issues.

Purified AddAB had helicase and exonuclease activities but had no detectable endonuclease properties. AddAB can act on different ds DNA substrates including 5′ overhangs, 3′ overhangs and blunt ends. The complex requires ATP for exonuclease and helicase activity.

Chapter 5: Discussion

5.1: Introduction

The main aim of this work was to identify the genes in *B. fragilis* NCTC9343 that encode the pre-synapsis proteins involved in processing of DSBs during homologous recombination. Through inter-species complementation, production of deletion strains and phenotypic characterization of the deletion strains, the *B. fragilis* genes (BF2192 and BF2191) have been identified as encoding AddA and AddB homologues, respectively. *In vivo*, these proteins have a role in DSB repair in response to DNA damage, whether from exogenous or endogenous sources. The AddAB system requires both polypeptides to work in concert for pre-synaptic processing of DSBs. A method for rapid purification of the AddAB complex was used; *in vitro* experiments showed that the complex was an ATP-dependant helicase and nuclease that acted on DSBs.

5.2: The process of pre-synaptic DSB repair is conserved in *E. coli* and *B. fragilis*

Expression of *B. fragilis* AddAB in an *E. coli* Δ *recBCD* strain was sufficient to alleviate the growth defect of the strain and enable repair of DNA DSBs caused by UV exposure; in both instances the phenotype of the *E. coli* Δ *recBCD* was rescued to the same level as wild type. This result is in keeping with Kooistra *et al*., (1993) who found that expression of the *B. subtilis* AddAB system in *E. coli* Δ *recBCD* strains alleviated the UV sensitivity and rescued the cell viability defect.

These results imply that the processing of DSBs by *B. fragilis* AddAB is similar to that of the *E. coli* RecBCD complex and hence must generate a 3′ overhang, the substrate required for synapsis by RecA. A further implication is that no accessory proteins are required by the *B. fragilis* AddAB complex to process DSBs.

This work also suggested that *B. fragilis* chi sites must be present in *E. coli* in sufficient numbers to stimulate recombination. If *E. coli* did not contain *B. fragilis* chi sites there would have been complete degradation of the chromosome and

viability would not have been restored. The identified chi sites in *E. coli* (5′-GCTGGTGG-3′), *L. lactis* (5′-GCGCGTG-3′) and *B. subtilis* AddAB (5′ AGCGG 3′) are all short (5-8bp) sequences (Smith *et al*., 1981; Biswas *et al.,* 1995; El Karoui *et al.,* 1998; Chedin *et al.,* 1998). If it is assumed that the *B. fragilis* chi site sequence is short, then there is a high probability that it will be present in the *E. coli* genome in sufficient numbers to ensure recombinogenic function of AddAB. In *E. coli,* RecC is essential for correct recognition of chi (Handa *et al*., 1997; Arnold *et al*., 2000). Our work is consistent with data for *B. subtilis* AddAB that shows the function of chi recognition must be executed by either or both of the AddAB proteins (Chedin *et al*., 2000).

5.3: Metronidazole ultimately causes DSBs

The presence of functional *addAB* genes in *B. fragilis* is required for the repair of DSBs caused by exogenous damage. The proteins are therefore essential for UV and metronidazole resistance in *B. fragilis* NCTC9343. The exact mode of action of metronidazole is unknown; it is believed that the nitro group of metronidazole is chemically reduced by anaerobic respiration, producing a cytotoxic nitro-radical anion intermediate, that causes ss nicks in DNA, which ultimately become DSBs following replication. Deletion of *addAB* made the strain more sensitive to metronidazole than the wild type, implying that metabolism of metronidazole causes DNA DSBs. It should be noted, that the majority of work on metronidazole sensitivity in recombination deficient mutants used *E. coli* as a model and to our knowledge, our work is the first demonstration of metronidazole sensitivity in an obligate anaerobe that cannot process DSBs. A *B. fragilis recA* mutant is also sensitive to metronidazole, however, this mutation abolishes all strand-exchange whether arising from single strand nicks or DSBs (Steffens *et al*., 2010).

5.4: Comparison of recombination deficient strains: *E. coli* and *B. fragilis*

When examining the phenotype of the *B. fragilis* Δ *addAB* strain, it was observed that UV sensitivity and the growth defect of the strain was more severe than in the

equivalent *E. coli* Δ *recBCD* strain. This result could imply that *B. fragilis* has more endogenous DNA damage than *E. coli*. This may be due to anaerobic growth, yet current thought suggests that aerobic growth, due to the presence of oxygen and its harmful by-products is more damaging to DNA than anaerobic growth. To test this, it would be interesting to characterise the phenotype of the *E. coli* Δ *recBCD* strain under anaerobic conditions and compare the results to aerobic conditions. Another cause of higher endogenous DNA damage levels in *B. fragilis* could be that replication forks stall more frequently than in *E. coli*. This could simply be due to the *B. fragilis* genome being larger (an additional 0.5 Mega bases) than *E. coli* or perhaps the unique genomic feature of site-specific inversions of the NCTC9343 genome could impede replication (Cerdeno-Tarraga *et al*., 2005). One could imagine that if a replication fork was to encounter an on-going inversion event this could disrupt the replication fork and ultimately cause DNA damage.

5.5: AddAB is not involved in short term oxygen exposure tolerance

The AddAB proteins did not appear to be involved in the oxygen tolerance of *B. fragilis*. Wild type and Δ *addAB B. fragilis*, whether grown anaerobically or with an oxygen exposure, showed no difference in viability. This implied there was no DNA damage due to short term oxygen exposure in *B. fragilis*, therefore oxygen and its harmful by-products must have been eliminated. This indicates that other proteins are involved in the removal of DNA damaging products following oxygen exposure.

Bacteroides is known to survive short periods of oxygen exposure. It can also be the sole infecting agent during abscess formation, which shows an ability to tolerate oxygen. Baughn and Malamy, (2004) showed that *B. fragilis* could survive nanomolar levels of oxygen by use of cytochrome bd oxidase, while Rocha *et al*., (1996) found that the global oxidative stress response (OSR) system induced genes that included catalase, in response to oxygen or hydrogen peroxide exposure.

5.6: SOS Response in *B. fragilis*

B. fragilis NCTC9343 appears to have an SOS system and the *addAB* genes are required for activation of this response. Interestingly *B. fragilis* NCTC9343 has no annotated *lexA* or *sfiA* genes. In *E. coli* DNA damage causes LexA to self-cleave causing lifting of repression of the SOS regulon; SfiA is induced and inhibits polymerisation of Ftz, leading to inhibition of septation and filament formation. This project has shown that septation is inhibited following metronidazole treatment (chapter 3), which indicates the *B. fragilis* SOS response must have a different pathway for induction of the SOS system, and the genome must encode novel inhibitors of septation.

5.7: *B. fragilis* putative RecD

The *in vivo* work in *E. coli* provided evidence that *B. fragilis* putative *recD* was not required as part of *B. fragilis* pre-synaptic complex for repair of DSBs. It was also demonstrated that the *B. fragilis* putative *recD* could not rescue the phenotype of an *E. coli recD* null strain (NM1035). To further illucidate the role of the *B. fragilis* putative RecD an in silico analysis was performed based the criteria provided by the work of Rocha *et al*., (2005) and Montague *et al*., (2009). The conclusion of the analysis was that that *B. fragilis* putative RecD may be of the D1 type, even though it is in an AddAB containing bacterium; to further elucidate its role/type it would be interesting to characterise the phenotype of a *B. fragilis recD* deletion mutant. The other possibility is that the criteria used to allocate RecD as either 1 or 2 are incorrect and that these authors have made a generic assumption without examining the genomes of *Bacteroides* or related genera.

5.8: Expression and purification – implications for structure

Throughout the expression and purification of AddAB, induced protein levels were low. This may have been due to degradation of the proteins. To overcome this

problem two methods were used; co-purification of the proteins and N-terminal MBP tags. The *E. coli* RecBCD structure may provide insight into why it was so difficult to express and purify the separate subunits of the AddAB complex. The crystal structure of RecBCD shows that the RecB nuclease domain has a linker region of 70 amino acids (Singleton *et al.*, 2004) (figure 5.8.1 A). Yu *et al.*, (1998) showed that within the linker region (residues 928-933) there was a site sensitive to proteolytic degradation. Within the RecBCD complex, RecB and RecC are tightly bound to each other in a manner that occludes many sites that are susceptible to proteolysis (Singleton *et al*., 2004) (figure 5.8.1 A). An NCBI DART multi-domain search assigns RecB as the closest match to AddA (E value = 4.59e-61) indicatinging that they may have similar folds or structure. It is therefore possible that AddA has sensitive proteolytic sites in a similar position as RecB and that expression and co-purification or the MBP tag obscured this site.

A. B.

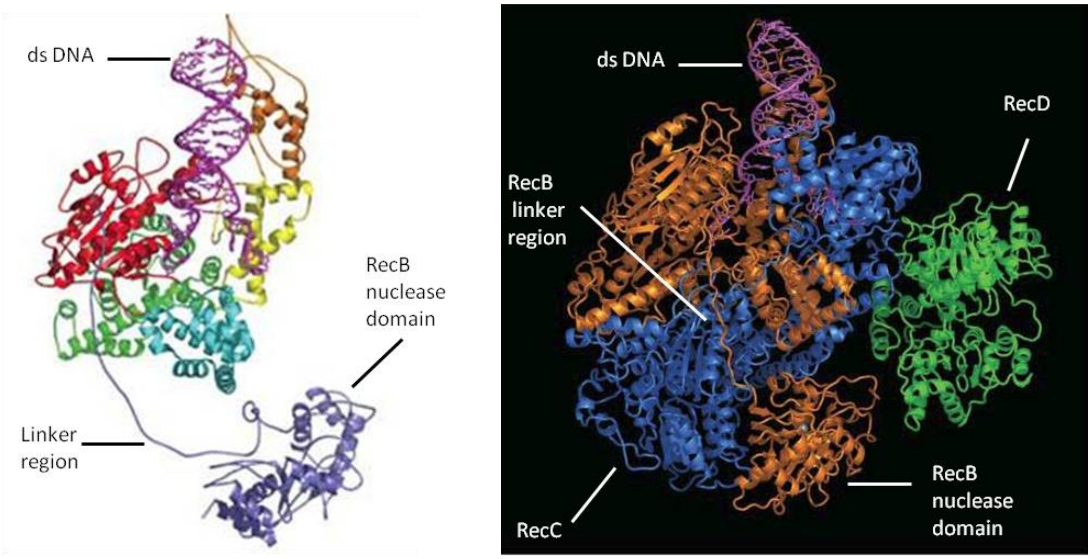


Figure 5.8.1: Crystal structure of RecB

A. The crystal structure of RecB showing the long linker region, that is susceptible to proteolysis, connecting the nuclease domain to the rest of the protein. B. The RecC and RecB subunits associate closely when forming the RecBCD complex, this protects the linker region. Figure adapted from Singleton *et al*., (2004)

5.9: AddAB is an ATP-dependant helicase and exonuclease with a requirement for magnesium

In vitro, B. fragilis AddAB can unwind and degrade blunt-ended and nearly blunt (4 nucleotide overhangs) ds DNA substrates this is consistent with findings for the activity of RecBCD on DNA substrates (Roman and Kowlaczykowski, 1989; Taylor and Smith, 1995). *B. subtilis* AddAB can also degrade blunt or nearly blunt ds DNA but cannot act on ds DNA with 50 nucleotide 5′ single-strand overhangs (Chedin *et al*., 2000). In contrast our results for *B. fragilis* AddAB showed that the complex can act on a ds DNA substrate with 40 nucleotide 5′ single-strand overhangs (figure 4.3.7.1) indicating that the complex can load onto a 5′ single-strand.

In chapter 4 (section 4.3.2) it was shown that ATP was essential for AddAB helicase and nuclease activity, which is consistent with the ATP requirement of *E. coli* RecBCD and *B. subtilis* AddAB (Dillingham *et al*., 2003; Kooistra and Venema., 1991). It was also observed (section 4.3.3) that high concentrations of magnesium (2mM) enhanced the degradative activity of AddAB, which is consistent with the findings that varying the magnesium concentration affected activity of *E. coli* RecBCD (Roman and Kowalczykowski,1989) Chedin *et al*., (2000) and Yeeles and Dillingham, (2007) both found that different concentrations of magnesium (1mM-2mM) enhanced *B. subtilis* AddAB nuclease activity.

5.10: Proposed helicase and exonuclease functions of the AddAB subunits

The biochemical characterisation of purified *B. fragilis* AddAB established the ATP-dependant degradation of ds DNA, but nuclease or helicase functions could not be assigned to the specific subunits of the complex. Yeeles and Dillingham, (2007) assigned nuclease and helicase functions to *B. subtilis* AddA and AddB by analysis of mutants. A helicase mutant containing a K36A mutation in the conserved lysine of the Walker A motif of the AddA subunit could not unwind DNA and therefore could not degrade DNA. Altered complexes that contained a mutation (aspartate to alanine) in the nuclease motifs of both AddA and AddB were constructed; indicated as

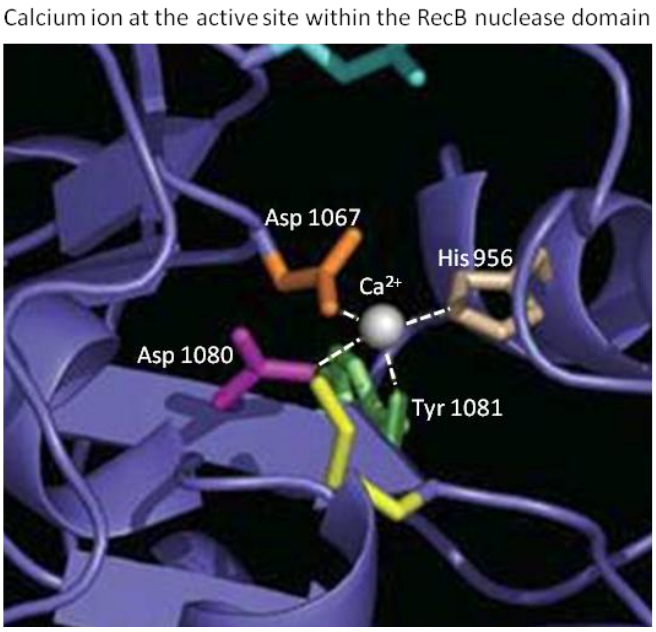
$AddA^{N}B$, $AddAB^{N}$ and $AddA^{N}B^{N}$. All the altered complexes displayed wild type helicase activity, but nuclease activity was severely reduced in the $AddA^{N}B^{N}$ complex. Complexes with either altered AddA or AddB nuclease ($AddA^{N}B$, $AddAB^{N}$) showed an intermediate nuclease activity compared to the $AddA^{N}B^{N}$ and wild type complexes. This demonstrated that AddA and AddB are both nucleases, while only AddA is a helicase (Yeeles and Dillingham, 2007). During this project the *B. fragilis* AddA and AddB subunit were not purified individually, therefore only suppositions could be made as to which of the *B. fragilis* AddAB subunits contained helicase or nuclease functions. In chapter 3 an *in silico* analysis of the protein sequences of AddA and AddB showed that the AddA subunit contained a heliacase and nuclease domain while the AddB subunit only contained a nuclease domain.

5.11: AddA and AddB sequences suggest binding of magnesium in a manner similar to RecB

In section 4.3.4 it was observed that calcium abolished nuclease activity of *B. fragilis* AddAB. Calcium has been shown to inhibit *E. coli* RecBC exonuclease function and *B. subtilis* AddAB exonuclease function (Rosamond *et al*., 1979; Kooistra *et al.,* 1997). The crystal structure of RecBCD showed a calcium ion (from the crystalisation buffer) bound at the region believed to be the active nuclease site of RecB, where magnesium is thought to bind (Figure 5.11.1A). The RecB amino acid residues of His 956, Asp 1067, Asp 1080 and the main chain carbonyl of Tyr 1081 were coordinated with the calcium ion (Singleton *et al*., 2004). Examination of the amino acid sequence of the AddA and AddB nuclease sites shows a similar arrangement to the RecB nuclease sequence (highlighted by red in figure 5.11.1B). The amino acids from the RecB nuclease site that coordinate with the calcium ion are shown in larger font; a similar sequence arrangement is found in AddA (His 888, Asp 1000, Asp 1013 and Phe 1014, shown in larger font) and AddB (His 706, Asp 820, Asp 833 and Tyr 834, shown in larger font). AddA has the substitution of phenylalanine for tyrosine at position 834. Tyrosine and phenylalanine have very similar properties only differing in their side chain structure by an OH group; as the calcium ion only made contact with the carbonyl of the main chain of tyrosine, it

would seem phenylalanine is an acceptable substitution. The alignments also show that AddA and AddB both have a histidine at a similar position to RecB, relative to the nuclease site. Our results combined with analysis of the AddA and AddB protein sequences suggest that *B. fragilis* AddAB may bind magnesium and calcium in a similar manner to RecBCD.

A.



B.

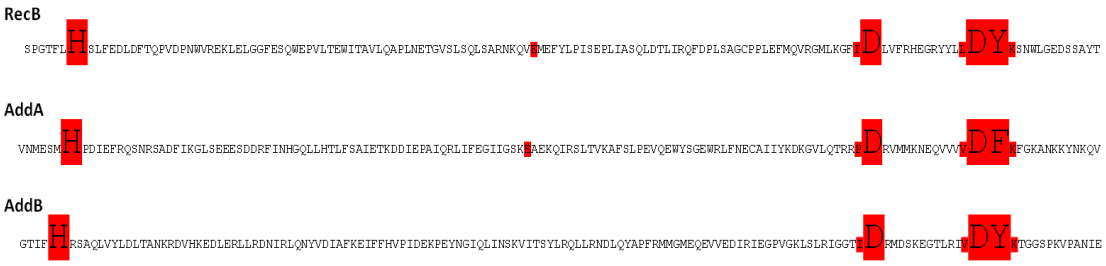


Figure 5.11.1: AddA and AddB sequence suggest binding of magnesium in a manner similar to RecB

A. Calcium ion bound at the active nuclease site of RecB. Figure adapted from Singleton *et al*., (2004) B. Sequence analysis showing AddA and AddB amino acids proposed to be involved in metal cofactor binding (highlighted by larger font).

5.12: Summary

In conclusion, the pre-synaptic processing proteins involved in homologous recombination in *B. fragilis* have been identified. The function of these proteins was demonstrated *in vivo* by rescue of a repair-deficient strain of *E. coli*. To further examine the role of the AddAB proteins in DSB repair, a *B. fragilis* strain with a deletion of *addAB* was constructed and shown to be sensitive to DNA damaging agents. The identified genes were BF2192 and BF2191, and have been renamed *addA* and *addB*, respectively. This project has demonstrated that *B. fragilis* encodes a two component system, where both genes products are required to work in concert for pre-synaptic processing of DSBs. A method for rapid purification of the AddAB complex was determined. This method was used in the first published paper that characterised the biochemical properties of *B. fragilis* AddAB through single molecule observations (Reuter *et al*., 2010). The AddAB complex was purified and shown to be an ATP-dependant helicase and exonuclease that acted on double-stranded DNA ends.

At the time of starting this project only two types of pre-synaptic systems had been studied and were believed to be phylogenetically distinct; the Gram negative RecBCD system and the Gram positive AddAB system. Recent work by Sinha *et al.*, (2009) has identified a third pre-synaptic processing system, AdnAB from *Mycobacterium*. Our work has shown that the Gram negative bacterium *B. fragilis* contains an AddAB-type system and demonstrates that AddAB is not as phylogentically restricted as previously thought. This suggests that the original view of the *E. coli* RecBCD system being the paradigm for DSB repair in Gram negative bacteria was incorrect. These observations are supported by genomic analyses that show AddAB homologues are wide-spread in the eubacteria (Cromie, 2009).

5.13: Future Work

Though the aims of this project were addressed the results have raised new questions and possible new directions for research.

Further biochemical assays of *B. fragilis* AddAB have been undertaken (Reuter *et al*., 2010). It would appear prudent to further this line of research so that physical assignation of function of the domains and subunits can be made. To this end assays performed with only one of the subunits may shed light on their roles. Production of mutant AddAB complexes that are deficient for helicase or nuclease activity and chi recognition/activation would allow further elucidation of the complexes role.

It would be diligent if further characterisation of the *B. fragilis* mutants produced in this project (*addAB, addB sbcCD* null strains) was undertaken. This could provide further insight into the mode of action of the pre-synaptic machinery and homologous recombination in *B. fragilis*. In addition it would be desirable to complete the work that had begun on producing other *B. fragilis* mutants (*addA* and *recF*) and determine if they have a role in recombination.

The results of the initial work that focussed on the putative *B. fragilis recD* (BF0679) posed an interesting question. It would appear that the proposed gene product of *B. fragilis* BF0679 is of a RecD type, but which type and what is its function? It would seem appropriate that the putative *B. fragilis* RecD protein should be functionally characterised to determine if it is RecD1 type or RecD2. Work had begun on producing a *B. fragilis recD* mutant, completion of this mutant and characterisation of the mutant would allow determination of the type assignation of the putative RecD protein.

The SOS response in *B. fragilis* is an interesting result particularly as yet there is no annotated LexA. This could mean that *B. fragilis* has a novel mode of repression/induction of the SOS system and would prove an interesting direction for study.

As *B. fragilis* is a medically relevant pathogen, potential use of the results of this project may provide a starting point for further work in understanding the role of metronidazole. A further use of this work would be to determine if a drug could be produced that would make *B. fragilis* more sensitive to metronidazole or one that could make it more sensitive to endogenous damage i.e. unable to repair DSBs. As chi sites appear to be unique to each species – the recognition site for *chi* within the AddAB complex could be a target for a drug. One could speculate that an inhibitor of AddAB could be one that would interact with the chi recognition site in a manner that would prevent loading onto DNA. This could provide an interesting line of investigation for possible new antimicrobials, which is necessary due to rising rates of resistance in bacteria to current antimicrobials.

References

Amundsen, S. K., A. F. Taylor, M. Reddy, and G. R. Smith. (2007) Intersubunit signaling in RecBCD enzyme, a complex protein machine regulated by Chi hot spots. *Genes & Development* **21**: 3296–3307.

Amundsen, S. K., J. Fero, L. M. Hansen, G. A. Cromie, J V. Solnick, G. R. Smith, and N. R. Salama. (2008) *Helicobacter pylori* AddAB helicase-nuclease and RecA promote recombination-related DNA repair and survival during stomach colonization. *Molecular Microbiology* **69**(4): 994–1007.

Anderson, D. G., and S. C. Kowalczykowski (a). (1997) The recombination hot spot χ is a regulatory element that switches the polarity of DNA degradation by the RecBCD enzyme. *Genes & Development* **11**: 571-581.

Anderson, D. G., and S. C. Kowalczykowski (b). (1997) The Translocating RecBCD Enzyme Stimulates Recombination by Directing RecA Protein onto ssDNA in a χ -Regulated Manner. *Cell* **90**: 77–86.

Arakawa, K., R. Uno, Y. Nakayama, and M. Tomita. (2007) Validating the significance of genomic properties of Chi sites from the distribution of all octamers in *Escherichia coli*. *Gene* **392**: 239–246.

Arnold, D. A., N. Handa, I. Kobayashi and S. C. Kowalczykowski. (2000) A Novel, 11 Nucleotide Variant of χ, χ *: One of a Class of Sequences Defining the *Escherichia coli* Recombination Hotspot χ. *J. Mol. Biol.* **300**: 469-479.

Baughn, A. D., and M. H. Malamy. (2004) The strict anaerobe *Bacteroides fragilis* grows in and benefits from nanomolar concentrations of oxygen. *Nature* **427**: 441-444.

Backhed, F., H. Ding, T.Wang, L. V. Hooper, G. Y. Koh, A. Nagy, C. F. Semenkovich, and J. I. Gordon. (2004) The gut microbiota as an environmental factor that regulates fat storage. *PNAS* **101**(44): 15718–15723.

Bianco, P. R., L. R. Brewer, M. Corzett, R.Balhorn, Y. Yehk, S. C. Kowalczykowski & R. J. Baskin. (2001) Processive translocation and DNA unwinding by individual RecBCD enzyme molecules. *Nature* **409**: 374-378.

Biswas, I., E. Maguin, S. D. Ehrlich and A. Gruss. (1995) A 7-base-pair sequence protects DNA from exonucleolytic degradation in *Lactococcus lactis. Natl. Acad. Sci.* **92**: 2244-2248.

Blattner F.R., G. 3rd. Plunkett, C.A. Bloch, N.T. Perna, V. Burland, M. Riley, J.Collado-Vides, J.D. Glasner, C.K. Rode, G.F. Mayhew, J. Gregor, Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y. (1997) The complete genome sequence of *Escherichia coli* K-12. *Science* **277**(5331):1453-74.

Brazier J.S., S.L. Stubbs, and B.I. Duerden. (1999) Metronidazole resistance among clinical isolates belonging to the *Bacteroides fragilis* group: time to be concerned? *J Antimicrob Chemother* **44**(4):580-1.

Britz M. L., and R. G. Wilkinson. (1979) Isolation and Properties of Metronidazole-Resistant Mutants of *Bacteroides fragilis*. *Antimicrobial Agents and Chemotherapy* **16**(1): 19-27.

Bry L., P. G. Falk, T. Midvedt and J. I. Gordon. (1996) A model of host-microbial interactions in an open mammalian ecosystem. *Science* **273**: 1380-1383.

Casanueva, A. I., L. Paul, S.Patrick & V. R. Abratt. (2008) An AraC/XylS family transcriptional regulator homologue from *Bacteroides fragilis* is associated with cell survival following DNA damage. *FEMS Microbiol Lett* **278**: 249–256.

Castellia, M., M. Malagolia, A. I. Rubertoa, A. Baggioa, C. Casolarib, C. Cermellib, M. R. Bossac, T. Rossid, F. Paoluccie and S. Roffiae. (1997) In-vitro studies of two 5-nitroimidazole derivatives. *Journal of Antimicrobial Chemotherapy* **40**: 19–25.

Cerdeno-Tarraga A.M., S. Patrick, L. C. Crossman, G. Blakely, V. Abratt, N. Lennard, I. Poxton, B. Duerden, B. Harris,1 M. A. Quail, A. Barron, L. Clark, C. Corton, J. Doggett, M. T. G. Holden, N. Larke, A. Line, A. Lord, H. Norbertczak, D. Ormond, C. Price, E. Rabbinowitsch, J. Woodward, B. Barrell, J. Parkhill (2005) Extensive DNA Inversions in the *B. fragilis* Genome Control Variable Gene Expression. *Science* **307:** 1463- 1465

Chang, K., S. Ho, J. Yang, and J. Wang. (1997) Isolation of a Genetic Locus Associated with Metronidazole Resistance in *Helicobacter pylori*. *Biochemical and Biophysical Research Communications* **236**: 785–788.

Chaudhury, A. M., and G. R. Smith. (1984) *Escherichia coli recBC* deletion mutants. *J Bacteriol* **160**(2):788-91.

Chedin, F., P. Noirot, V. Biaudet, and S. D. Ehrlich. (1998) A five-nucleotide sequence protects DNA from exonucleolytic degradation by AddAB, the RecBCD analogue of *Bacillus subtilis. Mol Microbio*. **29**: 1369-1377.

Chedin, F., S. D. Ehrlich, **and S. C. Kowalczykowski.** (2000) The *Bacillus subtilis* AddAB helicase/nuclease is regulated by its cognate Chi sequence in vitro. *J Mol Biol* **298**(1):7-20.

Chedin, F., N. Handa, M. S. Dillingham, and S. C. Kowalczykowski. (2006) The AddAB Helicase/Nuclease Forms a Stable Complex with Its Cognate χ Sequence During Translocation. *The Journal of Biological Chemistry* **281**(27): 18610-18617.

Chen, H., B. Ruan, M. Yu, J.Wang, and D. A. Julin. (1997) The RecD Subunit of the RecBCD Enzyme from *Escherichia coli* is a Single-stranded DNA-dependent ATPase. *The Journal of Biological Chemistry* **272**(15): 10072–10079.

Chen, J., K. P. Tan, W. E. Ward, and L. U. Thompson. (2003) Exposure to Flaxseed or Its Purified Lignan during Suckling Inhibits Chemically Induced Rat Mammary. Tumorigenesis. *Exp Biol Med* **228**(8):951-8.

Church, D. L., R. D. Bryant, H. R. Rabin, and E. J. Laishley. (1991) Physiological effects of metronidazole *on Clostridium pasteurianum. Journal of Antimicrobial Chemotherapy* **28**: 221-228.

Churchill, J. J., D. G. Anderson, and S. C. Kowalczykowski. (1999)The RecBC enzyme loads RecA protein onto ssDNA asymmetrically and independently of χ, resulting in constitutive recombination activation. *Genes and Development* **13**: 901–911.

Churchill, J. J., and S. C. Kowalczykowski. (2000) Identification of the RecA Protein-loading Domain of RecBCD Enzyme. *J Mol Biol* **297**: 537-542

Clavel, T., G. Henderson, W. Engst, J.l Dore and M. Blaut. (2006) Phylogeny of human intestinal bacteria that activate the dietary Lignan secoisolariciresinol diglucoside. *FEMS Microbiol Ecol* **55**: 471–478.

Cobb, B., A., Q. Wang, A. O. Tzianabos and D. L. Kasper. (2004) Polysaccharide Processing and Presentation by the MHCII Pathway. *Cell* **117**(5): 677-687.

Cox M. M., M. F. Goodman, K. N. Kreuzer, D. J. Sherratt , S. J. Sandler, K. J. Marians. (2000) The importance of repairing stalled replication forks. *Nature* **404:**37–41

Cromie, G. A. (2009) Phylogenetic ubiquity and shuffling of the bacterial RecBCD and AddAB recombination complexes. *J Bacteriol* **191:**5076-5084

Cromie, G. A., J. C. Connelly, and D. R. F. Leach. (2001) Recombination at Double-Strand Breaks and DNA Ends: Conserved Mechanisms from Phage to Humans. *Molecular Cell* **8:** 1163–1174

Dabert, P. S., D. Erlich and A. Gruss. (1992) χ sequence protects against RecBCD degradation of DNA in vivo. PNAS Genetics **89**: 12073-12077.

Dajkovic, A., and J. Lutkenhaus. (2006) Z ring as executor of bacterial cell division. *J Mol Microbiol Biotechnol* **11**(3-5):140-51.

Della M., P. L. Palmbos, H. M. Tseng, L. M. Tonkin, J. M. Daley, L. M. Topper, R. S. Pitcher, A. E. Tomkinson, T. E. Wilson, and A. J. Doherty. (2004) Mycobacterial Ku and ligase proteins constitute a two-component NHEJ repair machine. *Science* **306**: 683-685.

Dillingham, M. S., M. Spies, and S. C. Kowalczykowski. (2003) RecBCD enzyme is a bipolar DNA helicase. *Nature* **423**: 893- 897.

Dillingham, M. S., M. R. Webb, and S. C. Kowalczykowski. (2005) Bipolar DNA Translocation Contributes to Highly Processive DNA Unwinding by RecBCD Enzyme. *The Journal of Biological Chemistry* **280**(44): 37069- 37077.

Diniz, C. G., S. G. Santos, A. Cristina N.R. Pestana, L. M. Farias and M. Auxiliadora R. Carvalho. (2000) Chromosomal Breakage in the *B. fragilis* Group Induced by Metronidazole Treatment. *Anaerobe* **6**: 149-153

Dixon, D. A., and S. C. Kowalczykowski. (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.

Dudáš, A., and M. Chovanec. (2004) DNA double-strand break repair by homologous recombination. *Mutation Research* **566**: 131-167.

Eckburg, P. B., E. M. Bik, C. N. Bernstein, E. Purdom, L. Dethlefsen, M. Sargent, S. R. Gill, K. E. Nelson and D. A. Relman. (2005) Diversity of the Human Intestinal Microbial Flora. *Science* **308**: 1635-1638

El Karoui, M., D. Ehrlich, and A. Gruss. (1998) Identification of the *lactococcal* exonuclease recombinase and its modulation by the putative Chi sequence. *Proc Natl Acad Sci USA Genetics* **95**: 626–631.

Eftimiadi C., E. Buzzi, M. Tonetti, P. Buffa, D. Buffa, M. T. J. van Steenbergen, J. de Graaff and G. A. Bottag (1987) Short-chain fatty acids produced by anaerobic bacteria alter the physiological responses of human neutrophils to chemotactic peptide. *Journal of Infection* **14**: 43-53

Eggleston, A., T. O'neill, E. M. Bradbury and S. C. Kowalczykowski (1995) Unwinding of nucleosomal DNA by a DNA helicase. *The Journal of Biological Chemistry* **270**(5): 2024-2031

Erridge, C., A. Pridmore, A. Eley, J. Stewart and I. R. Poxton. (2004) Lipopolysaccharides of *Bacteroides fragilis*, *Chlamydia trachomatis* and *Pseudomonas aeruginosa* signal via Toll-like receptor 2. *Journal of Medical Microbiology* **53**: 735–740.

Falagas, M. E., E. Siakavellas. (2000) *Bacteroides*, *Prevotella*, and *Porphyromonas* species: a review of antibiotic resistance and therapeutic options. *Int J Antimicrob Agents*. **15**(1):1-9.

Fang, H., C. Edlund, M. Hedberg, C. E. Nord. (2002) New findings in beta-lactam and metronidazole resistant *Bacteroides fragilis* group. Int J Antimicrob Agents **19**(5):361-70.

Finch et al. 1986a Finch, P. W., A. Storey, K. E. Chapman, K. Brown, I. D. Hickson and P. T. Emmerson. (1986) Complete nucleotide sequence of the *Escherichia coli recB* gene. *Nucleic Acids Research* **14**: 8573-8582.

Finch et al. 1986b Finch, P. W., R. E. Wilson, K. Brown, I. D. Hickson, A. E. Tomkinson and P. T. Emmerson. (1986) Complete nucleotide sequence of the *Escherichia coli recC* gene and of the *thyA-recC* intergenic region. *Nucleic Acids Research* **14**(11): 4437-4451.

Finch et al. 1986c Finch, P. W., A. Storey, K. Brown, I. D. Hickson, and P. T. Emmerson. (1986) Complete nucleotide sequence of *recD*, the structural gene for the α subunit of Exonuclease V of Escherichia coli. *Nucleic Acids Res.* **14**(21): 8583 - 8594.

Finegold, S. M., V. L. Sutter, P. T. Sugihara, H. A. Elder, S. M. Lehmann, and R. L. Phillips. (1977) Fecal microbial flora in Seventh Day Adventist populations and control subjects. *The Journal of American clinical Nutrition* **30**: 1781-1792

Ganesan, S., and G. R. Smith. (1993) Strand-specific binding to duplex DNA ends by the subunits of the *Escherichia coli* RecBCD enzyme. *J Mol Biol* **229**(1):67-78.

Garcia-Rodriguez, J. A., J. E. Garcia-Sanchez and J. L. Munoz-Bellido. (1995) Antimicrobial resistance in anaerobic bacteria: current situation. *Anaerobe* **1**: 69-80.

Gibson, F. C., A. B. Onderdonk, D. L. Kasper, and A. O. Tzianabos. (1998) Cellular Mechanism of Intraabdominal Abscess Formation by *Bacteroides fragilis*. *The Journal of Immunology* **160**(10):5000-5006.

Guzman, C. A., F. Biavasco, and C. Pruzzo. (1997) Adhesiveness of *Bacteroides fragilis* Strains Isolated from Feces of Healthy Donors, Abscesses, and Blood. *Current Microbiology* **34**: 332–334.

Gupta, R. S. (2004) The Phylogeny and Signature Sequences Characteristics of *Fibrobacteres*, *Chlorobi*, and *Bacteroidetes Critical Reviews in Microbiology* **30**(2):123 -143

Gupta, R. S. and Lorenzini E. (2007) Phylogeny and molecular signatures (conserved proteins and indels) that are specific for the *Bacteroidetes* and *Chlorobi* species. *BMC Evolutionary Biology* **7**:71.

Haijema et al 1996a: Haijema, B. J., M. Noback, A. Hesseling, J. Kooistra, G. Venema, and R. Meima (1996) Replacement of the lysine residue in the consensus ATP-binding sequence of the AddA subunit of AddAB drastically affects chromosomal recombination in transformation and transduction of *Bacillus subtillis. Molecular Microbiology* **21**(5): 989-999.

Haijema et al 1996b: Haijema, B. J., G. Venema, and J. Kooistra. (1996) The C Terminus of the AddA Subunit of the *Bacillus subtilis* ATP-Dependent DNase Is Required for the ATP-Dependent Exonuclease Activity but Not for the Helicase Activity. *Journal of Bacteriology* **178**: 5086–5091.

Halpern, D., A. Gruss, J. Claverys & M. El Karoui. (2004) *rexAB* mutants in *Streptococcus pneumoniae*. *Microbiology* **150**: 2409–2414.

Handa, N., S. Ohashi, K. Kusano and I. Kobayashi. (1997) χ*, a χ-related 11-mer sequence partially active in an *E. coli recC** strain. *Genes to Cells* **2**: 525-536.

Handa, N., P. R. Bianco, R. J. Baskin, and S. C. Kowalczykowski. (2005) Direct Visualization of RecBCD Movement Reveals Cotranslocation of the RecD Motor after chi Recognition. *Molecular Cell* **17**: 745-750.

Hefferin, M. L. and A. E. Tomkinson. (2005) Mechanism of DNA double-strand break repair by non-homologous end joining. *DNA Repair* **4**: 639-648.

Hoekstra, W. P. M., J. E. N. Bergmans, and E. M. Zuidweg. (1980) Role of RecBC Nuclease in *Escherichia coli* Transformation. *Journal of Bacteriology* **143**(2):1031-1032

Hooper, L. V., M. H. Wong, A. Thelin, L. Hansson, P. G. Falk and J. I. Gordon. (2001) Molecular Analysis of Commensal Host-Microbial Relationships in the Intestine. *Science* **291**: 881-884.

Hofstad T., (1984) Pathogenicity of Anaerobic Gram-Negative Rods: Possible Mechanisms. *Reviews of Infectious Diseases* **6**(2):189-199

Ingham, H. R, S. Eaton, C. W. Venables, and P. C. Adams. (1978) *Bacteroides fragilis* resistant to metronidazole after long-term therapy. *Lancet* **1(8057)**: 214.

Ings, R. M . J., J. A . McFadzean, and W. E. Ormerod (1974) The mode of action of Metronidazole in *Trichomonas vaginalis* and other micro-organisms. *Biochemical Pharmocology* **23**: 1421-1429.

Jackson, D., A. Salem and G. H. Coombs. (1984) The in-vitro activity of metronidazole against strains of *Escherichia coli* with impaired DNA repair systems. *Journal of Antimicrobial Chemotherapy* **13**: 227-236.

Janion, C. (2008) Inducible SOS response system of DNA repair and mutagenesis in *Escherichia coli*. *Int J Biol Sci* **4**(6):338-44.

Kinder, S. A., J. L. Badger, G.O. Bryant, J. C. Pepe and V. L. Miller. (1993) Cloning of the YenI restriction endonuclease and methyltransferase from *Yersinia enterocolitica* serotype O8 and construction of a transformable R-M+ mutant. *Gene* **136**(1-2):271-5.

Kooistra, J. and G. Venema. (1991) Cloning, Sequencing, and Expression of *Bacillus subtilis* Genes Involved in ATP-Dependent Nuclease Synthesis. *Journal of Bacteriology* **173**(12): 3644-3655

Kooistra J., B. J. Haijema and G. Venema. (1993) The *Bacillus subtilis addAB* genes are fully functional in *Escherichia coli*. *Molecular Microbiology* **7**(6): 915-923.

Kooistra J., B. J. Haijema, A. Hesseling-Meinders and G. Venema. (1997) A conserved helicase motif of the AddA subunit of the *Bacillus subtilis* ATP-dependant nuclease (AddAB) is essential for DNA repair and recombination. *Molecular Microbiology* **23**(1): 137-149.

Kowalczykowski, S. C. (2000) Initiation of genetic recombination and recombination-dependent replication. *Trends Biochem Sci* **25**(4):156-65.

Kowalczykowski, S. C., D. A. Dixon, A. K. Eggleston, S. D. Lauder, and W. M. Rehrauer. (1994) Biochemistry of Homologous Recombination in *Escherichia coli*. *Microbiological Reviews* **58**(3) 401-465

Kuzminov, A. (1999) Recombinational Repair of DNA Damage in *Escherichia coli* and Bacteriophage λ. *Microbiology and molecular reviews* **63**(4): 751-813.

Lam, S. T., M. M. Stahl, K. D. McMilin, and F. W. Stahl. (1974). Rec-mediated recombinational hot spot activity in bacteriophage lambda. II. A mutation which causes hot spot activity. *Genetics* **77**: 425-433.

Lassmann, B., D. R. Gustafson, C. M. Wood, and J. E. Rosenblatt. (2007) Re-emergence of Anaerobic Bacteremia. *CID* **44**: 895-900.

Mazmanian, S. K., C. H. Liu, A. O. Tzianabos, and D. L. Kasper. (2005) An Immunomodulatory Molecule of Symbiotic Bacteria Directs Maturation of the Host Immune System. *Cell* **12**2: 107-118.

Mazmanian S. K., J. L. Round and D. L. Kasper. (2008) A microbial symbiosis factor prevents intestinal inflammatory disease. *Nature* **453**: 620-625.

McPartland, A., L. Green, H. Echols. (1980) Control of *recA* Gene RNA in *E. coli*: Regulatory and Signal Genes. *Cell* **20**(3): 731-737.

Mertens, K., L. Lantsheer, D. G. Ennis and J. E. Samuel. (2008) Constitutive SOS expression and damage-inducible AddAB-mediated recombinational repair systems for *Coxiella burnetii* as potential adaptations for survival within macrophages. *Molecular Microbiology* **69**(6): 1411–1426.

Michel, B. (2005) After 30 years of study, the bacterial SOS response still surprises us. *PLoS Biol* **(7)**:1174-1176.

Moeller, R., E. Stackebrandt, G. Reitz, T. Berger, P. Rettberg, A. J. Doherty, G. Horneck, and W. L. Nicholson. (2007) Role of DNA Repair by Nonhomologous-End Joining in *Bacillus subtilis* Spore Resistance to Extreme Dryness, Mono- and Polychromatic UV, and Ionizing Radiation. *J Bacteriol* **189**(8): 3306–3311.

Moncrief, J. S., R. J. Obiso, L. A. Barroso, J. J. Kling, R. L. Wright, R. L.Van Tassell, D. M. Lyerly, and T. D. Wilkins. (1995) The enterotoxin of *Bacteroides fragilis* is a metalloprotease. *Infect Immun* **63**(1):175-81.

Montague, M., C. Barnes, H.O. Smith, R. Chuang and S. Vashee. (2009) The Evolution of RecD Outside of the RecBCD Complex. *J Mol Evol* **69**:360–371

Moore, W. E. C. and L. V. Holdeman. (1974) Human Fecal Flora: The Normal Flora of 20 Japanese-Hawaiians. *Applied Microbiology* **27**(5): 961-979.

Myers, L. L., B. D. Firehammer, D. S. Shoop, and M. M. Border. (1984) *Bacteroides fragilis*: a Possible Cause of Acute Diarrheal Disease in Newborn Lambs. *Infection and Immunity* **44**(2):241-244

Nakano, V., R. M. F. Piazzab, and M. J. Avila-Camposa. (2006) A rapid assay of the sialidase activity in species of the *Bacteroides fragilis* group by using peanut lectin hemagglutination. *Clinical Microbiology Anaerobe* **12**: 238–241.

Namavar, F., E. B. M. Theunissen, A. M. J. J. Verweij-Van Vught, P. G. H. Peerbooms, M. Bal, H. F. W. Hoitsma and D. M. Maclaren. (1989) Epidemiology of the *Bacteroides fragilis* group in the colonic flora in 10 patients with colonic cancer. *J. Med. Microbiol*. **29**: 171-176

Nobles, E. R. (1973) *Bacteroides* infections. *Ann Surg* **177**(5):601-6.

Nowosielska, A. (2007) Bacterial DNA repair genes and their eukaryotic homologues: The role of recombination in DNA repair and genome stability. *Acta Biochemica Polonica* **54** (3): 483–494.

Obiso, R. J., A. O. Azghani, and T. D. Wilkins. (1997) The *Bacteroides fragilis* toxin fragilysin disrupts the paracellular barrier of epithelial cells. *Infection and Immunity* **65**(4):1431-9.

Oishi, M., and S. D. Cosloy. (1972) The genetic and biochemical basis of the transformability of *Escherichia coli* K-12. *Biochem Biophys Res Commun* **49**:1568-1572.

Onderdonk, A. B., D. L. Kasper, B. J. Mansheim, T. J. Louie, S. L. Gorbach, J. G. Bartlett. Experimental Animal Models for Anaerobic Infections *Reviews of Infectious Diseases* (1979) **1**(2):291-301

Otto, B. R., S. J. M. van Dooren, C. M. Dozois, J. Luirink, and B. Oudega. (2002) *Escherichia coli* Hemoglobin Protease Autotransporter Contributes to Synergistic Abscess Formation and Heme-Dependent Growth of *Bacteroides fragilis*. *Infection and Immunity* **70**(1): 5–10.

Oyston, P. C. F. and P. S. Handley. (1991) Surface components of *Bacteroides fragilis* involved in adhesion and haemagglutination. *J. Med. Microbiol*. **34**:51-55.

Patrick, S. (2001) *Bacteroides*. Chapter 91. pp 1921-1948. In: Molecular Medical Microbiology B. Boulnois, G. Griffin, C. Hormaeche, G. Keusch, M. Levine, H. Smith, P. Williams, M. Sussman (Eds). *London: Academic Press.* Chapter 91. pp 1921-1948.

Patrick, S. and B. I. Duerden. (2006) Gram-negative non-spore forming obligate anaerobes. Chapter 45, pp 541-556. In: Principles and Practice of Clinical Bacteriology 2nd Edn. SH Gillespie, P Hawkey (Eds). *London:Wiley* Chapter 45, pp 541-556.

Patrick S., J. Reid and A. Coffey. (1986) Capsulation of *invitro* and *invivo* Grown *Bacteroides* Species. *Journal of General Microbiology* **132: 1099-1109.**

Patrick S., S. Houston, Z. Thacker and G. W. Blakely. (2009) Mutational analysis of genes implicated in LPS and capsular polysaccharide biosynthesis in the opportunistic pathogen *Bacteroides fragilis*. *Microbiology* **155**: 1039–1049

Phillips, R. J., D. C. Hickleton, P. E. Boehmer, and P. T. Emmerson. (1997) The RecB protein of *Escherichia coli* translocates along single-stranded DNA in the 3' to 5' direction: a proposed ratchet mechanism. *Mol Gen Genet* **254**(3):319-329.

Pitcher, R. S., N. C. Brissett, and A. J. Doherty. (2007) Nonhomologous End-Joining in Bacteria: A Microbial Perspective. *Annu Rev Microbiol* **61**:259–82

Prasad, K. (1999) Reduction of Serum Cholesterol and Hypercholesterolemic Atherosclerosis in Rabbits by Secoisolariciresinol Diglucoside Isolated From Flaxseed. *Circulation.* **99**: 1355-1362.

Prasad, K. (2000) Oxidative stress as a mechanism of diabetes in diabetic BB prone rats: Effect of secoisolariciresinol diglucoside (SDG). *Molecular and Cellular Biochemistry* **209**: 89–96, 2000.

Prasad, K. (2001) Secoisolariciresinol diglucoside from flaxseed delays the development of type 2 diabetes in Zucker rat. *J Lab Clin Med* **138**(1): 32-39.

Quiberoni, A., I. Biswas, M. El Karoui, L. Rezaiki, P. Tailliez, and A. Gruss. (2001) In Vivo Evidence for Two Active Nuclease Motifs in the Double-Strand Break Repair Enzyme RexAB of *Lactococcus lactis*. *Journal of Bacteriology* **183**(13): 4071-4078.

Redondo, M. C., M. D. J. Arbo, J. Grindlinger, and D. R. Snydman. (1995) Attributable Mortality of Bacteremia Associated with the *Bacteroides fragilis* Group. *Clinical Infectious Diseases* **20**(6): 1492-1496.

Reid, J. H., and S. Patrick. (1984) Phagocytic and serum killing of capsulate and non-capsulate *Bacteroides fragilis J. Med. Microbiol.* **17:**247-257

Reuter, M., F. Parry, D. T. F. Dryden, and G. W. Blakely. (2010) Single-molecule imaging of *Bacteroides fragilis* AddAB reveals the highly processive translocation of a single motor helicase *Nucleic Acids Res* **38**(11):3721-31.

Rhee, K., P. Sethupathi, A. Driks, D. K. Lanning, and K. L. Knight. (2004) Role of Commensal Bacteria in Development of Gut-Associated Lymphoid Tissues and Preimmune Antibody Repertoire. *Journal of Immunology* **172**: 1118-1124.

Rocha, E. R., T. Selby, J.P. Coleman, and C. J. Smith. (1996). Oxidative stress response in an anaerobe, *Bacteroides fragilis*: a role for catalase in protection against hydrogen peroxide. *J. Bacteriol.* **178**: 6895–6903.

Rocha, E. P. C., E. Cornet, and B. Michel. (2005) Comparative and Evolutionary Analysis of the Bacterial Homologous Recombination Systems. *PLoS Genetics* **1**(2): 0247- 0259.

Rogemond, V., R. M. Guinet. (1986) Lectin like adhesins in the *Bacteroides fragilis* group. *Infect. Immun.* **53**(1): 99-102.

Roman, L. J., and S. C. Kowalczykowski. (1989) Characterization of the Helicase Activity of the *Escherichia coli* RecBCD Enzyme Using a Novel Helicase Assay. *Biochemistry* **28**: 2863-2873.

Rosamond, J., K. M. Telander, and S. Linn. (1979) Modulation of the Action of the RecBC Enzyme of *Escherichia coli* K-12 by $Ca^{2+.}$ *The Journal of Biological Chemistry* **254**(17): 8646-8652.

Rotstein, O. D., T. Vittorini, J. Kao, M. I. McBurney, P. E. Nasmith, and S. Grinstein. (1989) A Soluble *Bacteroides* By-Product Impairs Phagocytic Killing of *Escherichia coli* by Neutrophils. *Infection And Immunity* **57**(3): 745-753

Salyers, A. A. (1984) *Bacteroides* of the human lower intestinal tract. *Ann Rev Microbiol* **38:**293-313

Salyers, A. A., N. B. Shoemaker, and L. Y. Li. (1995) In the driver's seat: the *Bacteroides* conjugative transposons and the elements they mobilize. *J Bacteriol* **177**(20): 5727–5731.

Sears, C.L., L.L.Myers, A.Lazenby, and R. L. Van Tassell. (1995) Enterotoxigenic *Bacteroides fragilis. Clin Infect Dis* **20**(S2): 142-148.

Simms-Cendan, J. S. (1996) Metronidazole. *Infectious Diseases Update* **3**(5): 153-156.

Simon, G. L., M. S. Klemper, D. L. Kasper and S. L. Gorbach. Alterations in Opsonophagocytic Killing by Neutrophils of *Bacteroides fragilis* Associated with Animal and Laboratory Passage: Effect of Capsular Polysaccharide *J Infect Dis.* (1982) **145**(1):72-77.

Simon, R., U. Priefer and A. Pühler. (1983) A broad host-range mobilization system for in vivo genetic engineering: transposon mutagenesis in Gram negative bacteria. *Bio/Technology* **1**:784–791.

Singleton M. R. and D. B. Wigley (2002) Modularity and Specialization in Superfamily 1 and 2 Helicases. *Journal of Bacteriology* **184** (7): 1819–1826.

Singleton, M. R., M. S. Dillingham, M. Gaudier, S. C. Kowalczykowski and D. B. Wigley. (2004) Crystal structure of RecBCD enzyme reveals a machine for processing DNA breaks. *Nature* **432**:187-193.

Sinha, K. M., M-C. Unciuleac, M. S. Glickman, and S. Schuman. (2009) AdnAB: a new DSB-resecting motor-nuclease from mycobacteria *Genes Dev* **23:** 1423-1437.

Shoop, D.S., L.L. Myers, and J.B. LeFever. (1990) Enumeration of enterotoxigenic *Bacteroides fragilis* in municipal sewage. *Appl Environ Microbiol* **56**(7): 2243-2244.

Smith, **G. R., S. M. Kunes**, **D. W. Schultz**, **A. Taylor and K. L. Triman.** (1981) Structure of chi hotspots of generalized recombination. *Cell* **24**(2): 429-436.

Sourice, S., V. Biaudet, M. El Karoui, S. D. Ehrlich and A. Gruss. (1998) Identification of the Chi site of *Haemophilus influenzae* as several sequences related to the *Escherichia coli* Chi site. *Mol Microbiol* **27**: 1021-1029.

Spies, M., P. R. Bianco, M. S. Dillingham, N. Handa, R. J. Baskin, and S. C. Kowalczykowski. (2003) A Molecular Throttle: The Recombination Hotspot chi Controls DNA Translocation by the RecBCD Helicase. *Cell* **114**: 647–654.

Spies M., I. Amitani, R. J. Baskin, and S. C. Kowalczykowski. (2007) RecBCD enzyme switches lead motor subunits in response to χ- Recognition. *Cell* **131**(4): 694–705.

Spies, M., and S. C. Kowalczykowski. (2006) The RecA Binding Locus of RecBCD Is a General Domain for Recruitment of DNA Strand Exchange Proteins. *Molecular Cell* **21**: 573–580.

Stappenbeck, T. S., L. V. Hooper, and J. I. Gordon. (2002) Developmental regulation of intestinal angiogenesis by indigenous microbes via Paneth cells. *PNAS* **99** (24): 15451–15455.

Steffens, L. S., S . Nicholson, L. V. Paul, C. E. Nord, S. Patrick and V. R. Abratt. (2010) *Bacteroides fragilis* RecA protein overexpression causes resistance to metronidazole. *Res Microbiol* **161**(5):346-354.

Stingele, F., B. Corthésy, N. Kusy, S. A. Porcelli, D. L. Kasper and A. O. Tzianabos (2004) Zwitterionic Polysaccharides Stimulate T Cells with No Preferential V β Usage and Promote Anergy, Resulting in Protection against Experimental Abscess Formation. *J Immunol* **172**:1483-1490

Stukalin, E. B., H. Phillips III, and A. B. Kolomeisky. (2005) Coupling of Two Motor Proteins: A New Motor Can Move Faster. *PRL* **94**: 2381011-2381014.

Sun, J., D. A. Julin, and J. Hu. (2006) The Nuclease Domain of the *Escherichia Coli* RecBCD Enzyme Catalyzes Degradation of Linear and Circular Single-Stranded and Double-Stranded DNA. *Biochemistry* **45**: 131-140.

Sund, C. J., E. R.Rocha, A. O.Tzianabos, W.G. Wells, J. M. Gee, M. A. Reott, D. P. O'Rourke, and C. J. Smith. (2008) The *Bacteroides fragilis* transcriptome response to oxygen and H2O2: the role of OxyR and its effect on survival and virulence. *Mol Microbiol* **67(1)**:129-42.

Sutter, V. L., Y. Y. Kwok, and S. M. Finegold. (1972) Standardized antimicrobial disc susceptibility testing of anaerobic bacteria. I. Susceptibility of *Bacteroides fragilis* to tetracycline. *Appl Microbiol* **23**(2):268-75.

Tally, F. P., B. R. Goldin, N. Sullivan, J. Johnston, and S. L. Gorbach. (1978) Antimicrobial Activity of Metronidazole in Anaerobic Bacteria. *Antimicrobial Agents and Chemotherapy* **13**(3): 460-465.

Taylor, A. F., and G. R. Smith. (1995) Monomeric RecBCD Enzyme Binds and Unwinds DNA. *The Journal of Biological Chemistry* **270**(41): 24451-24458.

Taylor, A. F., and G. R. Smith. (2003) RecBCD enzyme is a DNA helicase with fast and slow motors of opposite polarity. *Nature* **423**: 889-893.

Teng, L., P. Hsueh, J. Tsai, S. Liaw, S. Ho, and K. Luh. (2002) High Incidence of Cefoxitin and Clindamycin Resistance among Anaerobes in Taiwan. *Antimicrobial Agents and Chemotherapy* **46**(9): 2908–2913.

Turner, P., R. Edwards, V. Weston, Gazis A., P. Ispahani, and D. Greenwood. (1995) Simultaneous resistance to metronidazole, co-amoxiclav, and imipenem in clinical isolate of *Bacteroides fragilis. Lancet.* **345**(8960):1275-1277.

Van Houte, J., and R. J. Gibbons. (1966) Studies of the cultivable flora of normal human feces. *Antonie vat~ Leelm,enhoek* **32**: 212-222.

Van Tassell, R.L. and T.D. Wilkins. (1978) Isolation of auxotrophs of *Bacteroides fragilis. Can J Microbiol* **24**: 1619–1621.

Wang, J. and D. A. Julin. (2004) DNA Helicase Activity of the RecD Protein from *Deinococcus radiodurans*. *The Journal of Biological Chemistry* **279**(50):52024–52032.

Wang, S. T., B. Setlow, E. M. Conlon, J. L. Lyon, D. Imamura, T. Sato, P. Setlow, R. Losick, and P. Eichenberger. The Forespore Line of Gene Expression in *Bacillus subtilis. J Mol Biol* (2006) **358**: 16–37.

Wasilauskas, B. L. (1989) Effectiveness of Cefotaxime Alone and in Combination with Desacetylcefotaxime Against *Bacteroides fragilis. Diagn Microbiol Infect Dis* **12**: 39-43.

Wareham, D. W., M. Wilks, D. Ahmed, J. S. Brazier, and M. Millar. (2005) Anaerobic Sepsis Due to Multidrug Resistant *Bacteroides fragilis*: Microbiological Cure and Clinical Response with Linezolid Therapy. *CID* **40**: e67-e68.

Weller, G. R., B. Kysela, R. Roy, L. M. Tonkin, E. Scanlan, M. Della, S. K. Devine, J. P. Day, A. Wilkinson, F. d'Adda di Fagagna, K. M. Devine, R. P. Bowater, P. A. Jeggo, S. P. Jackson,

and A. J. Doherty. (2002) Identification of a DNA nonhomologous end-joining complex in bacteria. *Science* **297**(5587):1686-9.

Wexler, H. M. (2007) *Bacteroides*: the Good, the Bad, and the Nitty-Gritty. *Clinical Microbiology Reviews* **20**(4):593–621

Wexler, H. M., E. K. Reada and T. J. Tomzynski. (2002) Characterization of omp200, a porin gene complex from *Bacteroides fragilis*: omp121 and omp71, gene sequence, deduced amino acid sequences and predictions of porin structure. *Gene* **283**:95–105

Wong, C. J., A. L. Lucius and T. M. Lohman. (2005) Energetics of DNA End Binding by *E. coli* RecBC and RecBCD Helicases Indicate Loop Formation in the 30-Single-stranded DNA Tail. *J Mol Biol* **352**: 765-782.

Wostmann, B. S., C. Larkin, A. Moriarty, and E. Bruckner-Kardoss. (1983) Dietary intake, energy metabolism, and excretory losses of adult male germfree Wistar rats. *Lab Anim Sci* **33**(1):46-50.

Xu J., and J. I. Gordon. (2003) Honor thy symbionts. *PNAS* **100** (18): 10452–10459.

Xu, J., M. K. Bjursell, J. Himrod, S. Deng, L. K. Carmichael, H. C. Chiang, L. V. Hooper, and J. I. Gordon. (2003) A Genomic View of the Human–*Bacteroides thetaiotaomicron* Symbiosis. *Science* **299**: 2074-2076.

Yeeles, J. T., and M. S. Dillingham. (2007) A dual-nuclease mechanism for DNA break processing by AddAB-type helicase-nucleases. *J Mol Biol* **371**(1): 66-78.

Yeeles, J. T., R. Cammack, and M. S. Dillingham. (2009) An Iron-Sulfur Cluster Is Essential for the Binding of Broken DNA by AddAB-type Helicase-Nucleases. *The Journal of Biological Chemistry* **284**(12): 7746–7755.

Yu, M., J. Souaya, and D. A. Julin. (1998) The 30-kDa C-terminal domain of the RecB protein is critical for the nuclease activity, but not the helicase activity, of the RecBCD enzyme from *Escherichia coli*. *Proc Natl Acad Sci USA Biochemistry* **95**: 981-986.

Zuniga-Castillo, J., D. Romero, and J. M. Martínez-Salazar. (2004) The Recombination Genes addAB Are Not Restricted to Gram-Positive Bacteria: Genetic Analysis of the Recombination Initiation Enzymes RecF and AddAB in *Rhizobium etli*. *Journal of Bacteriology* **186(23)**: 7905-7913.