Function of bacteriophage orf recombinasesin genetic exchange

Patricia Reed

Function of bacteriophage Orf recombinases in genetic exchange

by

Patricia Reed, B.Sc (Hons)
Ustinov College, Durham

Thesis submitted to the Department of Biological and Biomedical Sciences, University of Durham for the degree of Doctor of Philosophy

November 2006

Durham University

In loving memory of Quinn (2000-2004)

Research is endlessly seductive; writing is hard work. One has to sit down on that chair and think and transform thought into readable, consecutive, interesting sentences that both make sense and make the reader turn the page. It is laborious, slow, often painful, sometimes agony. It means rearrangement, revision, adding, cutting, rewriting. But it brings about a sense of excitement, almost a rapture; a moment on mount Olympus. In short it is an act of creation.

Barbara W Tuchman (1912-1989)

Contents

Tables and Figures

Declaration

The work presented in this thesis is my own, original research, except where indicated by statement or citation, and has not been submitted for any other degree.

Acknowledgements

I would like to thank my supervisor Gary for his guidance, advice and encouragement during the past 5 years, when not everything went according to plan. For his passion and enthusiasm for this work and his continued support throughout my PhD. I would also like to thank Fiona, Laura and all of the other lab members of the Wolfson research institute, for their friendship and support. A special mention must go to Diane for her excellent technical support and for keeping our lab running.

Many thanks to Dave Sherratt and his lab members for allowing me to take over their cold room at times during this study, and for their friendly conversations. Special thanks to James who helped me operate the FPLC machine and brought me extra jumpers to keep me warm during the long hours in the cold room.

Thanks to my mam and dad for their love and support throughout my whole life, and especially during the last few years when life has thrown a few hurdles in the way. For your love, support and endless encouragement I will always be truly grateful. To Rob and Seleena for welcoming me into your family and for giving me love and support, and of course for the relaxing windsurfing holidays. To all of my friends, too numerous to name individually, thank you for you support and encouragement throughout my PhD.

Finally, thanks to my boys. To Pippin and Odie for making me laugh and smile everyday, even through the hardest of times, I owe you my sanity. And to James, my love, my rock, I couldn't have done this without you. For your love, endless patience, encouragement and support I will always be truly thankful. I am blessed to have you in my life.

Function of bacteriophage Orf recombinases in genetic exchange

Patricia Reed

University of Durham

Recombination events in bacteriophages frequently occur by illegitimate exchange at short tracts of sequence homology, enabling these viruses to acquire novel genes and serve as vehicles for horizontal gene transfer. The emergence of new pathogenic organisms due to the acquisition of virulence determinants from bacterial viruses has stimulated considerable interest in the mechanisms of phage recombination.

Bacteriophage λ encodes its own recombination system, consisting of Exo, β and γ proteins. An additional λ recombinase, Orf, participates in the early stages of exchange, supplying a function equivalent to the *Escherichia coli* RecFOR complex. The host enzyme complex promotes the loading of the RecA strand exchange protein onto SSB-coated ssDNA. This thesis describes the purification and biochemical analysis of the λ Orf protein, in parallel with two distantly related homologs from *E. coli* cryptic prophage DLP12 and *Staphylococcus aureus* phage ϕETA.

λ Orf was found to belong to a family of proteins originating from diverse lambdoid phage and prophage sources. Members of this family reside within a conserved genetic module located between phage replication and cell lysis functions. Orf exists as a homodimer, arranged as a toroid with a shallow cleft running perpendicular to the central cavity. It binds preferentially to DNA containing single-stranded regions, and associates with *E. coli* SSB protein in the presence of ssDNA. The Orf homolog from *E. coli* DLP12 displayed similar properties. This work suggests that members of the Orf family function as recombination mediator proteins, stimulating the assembly of strand exchange proteins onto ssDNA, and highlights the importance of overcoming the barrier presented by SSB proteins during lambdoid phage recombination.

Chapter 1

Introduction

1.1 Summary

This thesis describes the purification and biochemical analysis of the Orf recombination protein from phage λ and its homologs, *E. coli* DLP12 Orf151 and *S. aureus* phage ϕETA Orf20. Background information relevant to the study of these recombinases is provided in this chapter. A brief description of the diversity of bacteriophages in nature is provided, their importance as contributors to the evolution of pathogenic bacteria and the ongoing studies aimed at unravelling the pathways of genetic recombination in phages. The λ Red recombination system and interplay between phage and host (*E. coli*) recombinases will also be discussed before introducing the main topic — the role of λ Orf during recombinational exchange.

1.2 Bacteriophage λ as a model system

Bacteriophages are a highly diverse group of viruses, capable of infecting a wide range of bacterial hosts. Their name comes from the Greek "phagein" meaning to eat, literally translated as "eaters of bacteria". More than 4500 different double-stranded (ds) DNA phages have been described, though few have been studied in detail (Hendrix *et al.*, 1999). They vary in morphology, genomic arrangement, mode of replication, lifecycle and host range. Even those with similar architecture may be completely unrelated at the nucleotide level. However, groups of phages with a

comparable gene layout and certain level of sequence relatedness have been documented. The best characterised family is the lambdoid or 'lambda-like' temperate phages, many of which are capable of productive genetic recombination with λ (Juhala *et al.*, 2000; Neely and Friedman, 1998; Miyamoto *et al.*, 1999; Plunkett *et al.*, 1999; Vander Byl and Kropinski, 2000). Lambdoid genomes display remarkable mosaicism, with regions of significant sequence homology interspersed with seemingly unrelated segments (Hendrix *et al.*, 1999; Juhala *et al.*, 2000; Highton *et al.*, 1990; Campbell, 1994; Hendrix, 2002). Their mosaic nature is evidence of the horizontal transfer of genetic material and that regular genome reorganisation is crucial for their evolutionary success.

Phages can exist as stable particles outside the bacterial cell, however they are metabolically inert in their extracellular form and the phage genetic material must re-enter the host in order to replicate. Upon injection, λ DNA can take up residence in the host chromosome in a dormant (lysogenic) state or develop lytically, yielding progeny phage and killing the host. The decision between these two developmental pathways is affected by the prevailing environmental conditions and the number of infecting particles per cell (Herskowitz and Hagen, 1980; Echols, 1983; Friedman, 1983; Weisberg, 1983). The lytic pathway is a productive cycle, during which the phage replicates its genome autonomously, expresses morphogenetic genes, assembles progeny phage particles and lyses the cell (Friedman, 1983). In contrast, the lysogenic lifestyle entails integration of the phage into the *E. coli* chromosome by site-specific recombination (Weisberg, 1983) and its genome is subsequently replicated along with its host. Expression of the lytic genes is repressed by the prophage and also confers immunity against further λ infection. The lytic pathway

can be restored by exposure to DNA-damaging agents, resulting in repressor cleavage (in the same way LexA repressor cleavage switches on the host SOS response) and induction of lytic gene expression. Phage λ development and the complex genetic networks ("the genetic switch") involved in the lysis-lysogeny decision have been studied in considerable depth (Herskowitz and Hagen, 1980; Hendrix, 1983; Ptashne, 2004; Dodd *et al.*, 2005; Oppenheim *et al.*, 2005) but will not be discussed further here.

The alternate lifestyles adopted by phages, allows them to serve as repositories of genetic information that can be readily shared among bacterial populations. During lytic growth, phage-encoded recombinases promote genetic rearrangements that occasionally result in acquisition of new genes, either from resident prophages or the host chromosome. Novel properties from an infecting phage can be expressed even in the prophage state, a process known as lysogenic conversion, which may confer an evolutionary advantage on the recipient bacterium. This ability to obtain new genetic material and mediate its transfer to other strains means that bacteriophages are vitally important for bacterial evolution, especially for the dissemination of virulence determinants. An assortment of virulence products are known to be encoded by phages, including toxins, antibiotic resistance factors, type III secretion systems, surface proteins and extracellular enzymes (Waldor, 1998; Miao and Miller, 1999; Boyd and Brussow, 2002; Wagner and Waldor, 2002). Numerous bacteria owe their pathogenicity to products expressed by prophages. Notable examples include the Cholera toxin of *Vibrio cholerae* encoded by phage CTXϕ (Waldor and Mekalanos, 1996), the structural (*tox*) gene of diphtheria toxin, from *Corynebacterium diphtheriae* (Buck *et al.*, 1985) and the Shiga toxins produced by phages 933W and

933V in *E. coli* O157:H7 (Perna *et al.*, 2001). Even λ encodes genes (*lom* and *bor*) that help *E. coli* strains adhere to human tissues and aid their survival in serum (Barondess and Beckwith, 1990; Barondess and Beckwith, 1995).

The events that lead to the exchange of genetic material tend to occur at short tracts of sequence homology. These types of exchanges are often referred to as illegitimate and are the result of phage-encoded recombinases that display promiscuous tendencies. Recombination between regions displaying only 30-50 base pairs (bp) of homology are not uncommon (Yu *et al.*, 2000). However, examination of the junctions left from previous recombination events mediated by lambdoid phages, revealed that as little as 5-10 bp of sequence identity may be sufficient to secure exchange (Juhala *et al.*, 2000). The remarkable diversity and widespread genetic mosaicism mentioned previously is a direct consequence of this recombinational promiscuity.

In order to understand the role phage play in the evolution of bacteria, we first need to unravel the molecular processes of phage recombination. Since its isolation in 1951 (Lederberg, 1951), λ has served as a pioneering model for numerous biological systems, for example: gene regulation, transcription termination, site-specific and general recombination, in addition to the developmental programs discussed earlier (Gottesman, 1999). As the best characterised of the lambdoid phages, it has been thoroughly characterised at both the genetic and biochemical levels (Hendrix *et al.*, 1999; Vander Byl and Kropinski, 2000; Juhala *et al.*, 2000; Barondess and Beckwith, 1990; Neely and Friedman, 1998; Plunkett *et al.*, 1999). Phage λ therefore provides an excellent model to elucidate the molecular mechanisms that promote bacteriophage genome rearrangements.

1.3 Genetic Recombination

Genetic recombination, also known as homologous recombination, is critically important for the maintenance of genomic integrity (Stahl, 1996; Kogoma, 1996; Cox *et al.*, 2000). Initial studies suggested that recombination primarily functions in the repair of chromosomal breaks and to generate genetic diversity within populations. However, it is now thought that the major role of homologous genetic recombination in bacteria, and in virtually all cellular organisms, is the non-mutagenic restoration of stalled or collapsed replication forks (Kreuzer, 2005; Cox *et al.*, 2000). Every mitotic division in mammalian cells is thought involve the recombinational repair of as many as ten replication forks (Haber, 1999). Similarly, bacteria grown under normal culture conditions are estimated to require the repair of a damaged replication fork in almost every cell, in every generation (Cox, 1998).

1.4 *E. coli* DNA replication

Wild-type *E. coli* cells grown in rich media at 37°C with suitable aeration will divide approximately every 20 minutes. The circular *E. coli* chromosome consists of 4.6 million base pairs (Mbp) and takes approximately 40 minutes to replicate. In order for an *E. coli* cell to divide faster than the chromosome can be replicated, the next round of replication must be initiated prior to cell division and segregation. The single *E. coli* chromosome is replicated bidirectionally, by two replication forks, from a single replication origin called *oriC* (Prescott and Kuempel, 1972). The semi-conservative nature of DNA replication requires that a replication fork must perform simultaneous DNA synthesis on both strands of parental DNA (Watson and Crick, 1953). DNA replication is catalysed in a highly ordered manner by a multisubunit

enzyme complex, the replisome (Figure 1.1). DnaA initiates replication by binding to multiple sites at *oriC* (Messer, 2002) and an association between DnaA and ATP leads to the localized melting of the adjacent duplex DNA (Erzberger *et al.*, 2002; Erzberger *et al.*, 2006; Simmons *et al.*, 2003). This complex is recognised by DnaC, which facilitates loading of DnaB, the replicative helicase, onto the exposed single-strand region (Carr and Kaguni, 2002). The helicase acts to disrupt the parental duplex, unwinding the individual template strands in preparation for DNA synthesis. Loading of DnaB recruits the clamp loader complex, which in turn loads the β–sliding clamp. This assembly forms the centre of the replication apparatus (Figure 1.1), coordinating simultaneous leading and lagging strand DNA synthesis (McHenry, 2003). The clamp loader complex consists of five proteins, δ, δ', $DnaX_3$ (consisting of two τ subunits and a truncated version of τ called γ) and two accessory factors χ and ψ (McHenry, 2003). The leading strand acts as a template for the DNA polymerase III (α) complex, which incorporates nucleotides continuously in the 5' to 3' direction. A second polymerase III (α) molecule, tethered to the other by the clamp loader complex, synthesises the lagging strand discontinuously as a series of short (1-3 kb) Okazaki fragments from RNA primers, which are synthesised by the DnaG primase. The polymerase and its associated β-sliding clamp are reloaded continually onto the lagging strand by the γ-complex clamp loader. Completion of lagging strand synthesis requires degradation of the RNA primers, filling of the gaps by DNA polymerase I and ligation of Okazaki fragments. Single-strand binding protein (SSB) binds to single-stranded DNA (ssDNA) around the replication fork, in particular on the lagging strand, and helps prevent reannealing of template sequences, eliminates hairpin formation at palindromic sequences and protects the

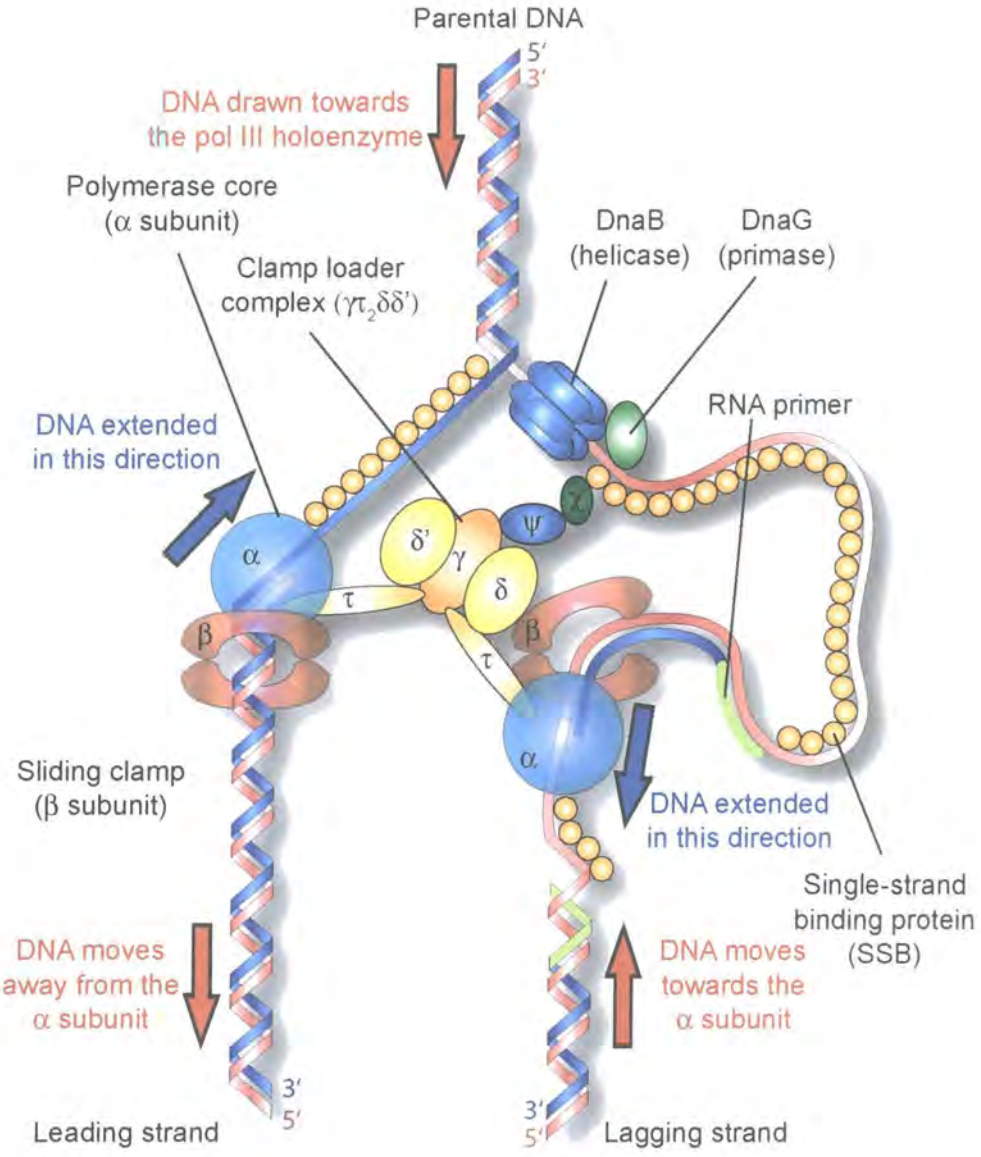

Figure 1.1. The *E. coli* replication fork. A schematic of the DNA polymerase III holoenzyme and associated factors. The replisome is a multisubunit enzyme responsible for catalysing DNA replication in *E. coli*. DnaB helicase unwinds parental DNA exposing the two individual strands which are coated with SSB. DNA extension on the leading strand is continuous in a 5' to 3' direction, the β clamp is loaded once by the clamp loader complex and the DNA continuously extended. On the lagging strand DNA synthesis is semi-discontinuous. RNA primers are generated by DnaG primase, these are then extended by Pol III (α subunit) to produce Okazaki fragments. DNA synthesis is repeatedly initiated on the lagging strand, and the β clamp is continually reloaded to extend each RNA primer. A single continuous strand is formed by degradation of the RNA primers, filling of the gaps by DNA polymerase I and ligase action upon adjoining fragments. Here on the lagging strand the β clamp is being loaded, note that DNA extension is also taking place. It is depicted in this fashion to show both activities, it is not meant to indicate that both events occur concurrently. Blue arrows indicate the direction of DNA extension; red arrows show the movement of DNA. Adapted from McGlynn and Lloyd, 2002.

exposed ssDNA. As we shall see later, SSB complexed with ssDNA can present a problem for recombination and repair proteins as it prevents access to the template (see section 1.6).

1.5 Repair of stalled replication forks in *E. coli*

DNA replication forks frequently encounter a damaged parental template and this can lead to fork stalling or even replisome disassembly (Cox *et al.*, 2000; McGlynn and Lloyd, 2002). For example, a replication fork will stall if it encounters a blockage in the DNA such as a thymine dimer. If it encounters a nick in the template DNA, one of the sister strands will be released producing a double strand break. Such situations may arise following exposure to DNA damaging agents, such as ultraviolet radiation or ionising radiation. However, this type of problem is not unique to cells subjected to exogenous sources of DNA damage, replication fork inactivation also occurs frequently during normal aerobic growth. This could be due, for example, to a lack of processive helicase activity, the presence of bound proteins or unusual DNA structures, natural replication pause sites in DNA (Rothstein *et al.*, 2000) or oxidative lesions produced by the action of endogenous hydroxyl radicals (Keyer *et al.*, 1995). It is estimated that at least 18% of cells suffer some form of replication fork damage during each replication cycle (Maisnier-Patin *et al.*, 2001). Consistent with this level of injury is the finding that *priA* mutants have very low viability (Marians, 1999). PriA is responsible for reloading DnaB helicase at sites other than the origin of replication. Hence, replication restart appears necessary in the majority of replicative cycles.

Repair of stalled replication forks is accomplished by a variety of methods including homologous recombination (Figure 1.2). Depending on the nature of the genetic insult received, different sets of presynaptic enzymes (e.g. RecBCD, RecFOR) aid assembly of the strand exchange recombinase RecA (see section 1.6). The 4-stranded Holliday junction, a central intermediate in the process of homologous recombination, can then be established (McGlynn and Lloyd, 2002; Cox *et al.*, 2000; Cox, 2001). Exactly which pathways are used and in which situations remains unclear, however the interconversion between a replication fork and a Holliday junction remains a fundamental principle of DNA repair (Sherratt, 2003). Because the Holliday junction lies within homologous duplexes it can be moved *via* branch migration by the action of RecG, RecQ or RuvAB and resolved by the action of the RuvC endonuclease. The actions of the RuvABC helicase-endonuclease complex will be discussed further in Chapter 8. However, the functions of RecBCD and RecFOR complexes in loading RecA onto ssDNA are relevant to phage λ recombination and are described in the following sections.

1.6 Genetic Recombination in *E. coli*

Genetic recombination in *E. coli* depends almost exclusively on RecA protein, which binds ssDNA and catalyses strand exchange between the bound strand and a homologous dsDNA partner (Kuzminov, 1999; Roca and Cox, 1997). Two RecA-dependent recombination systems operate in *E. coli*, the RecBCD and the RecFOR pathways, named after the critical enzyme participants (Clark and Sandler, 1994; Horii and Clark, 1973) (Figure 1.2). In both pathways, sufficient ssDNA needs to be exposed to allow RecA nucleoprotein filament formation and targeted assembly of

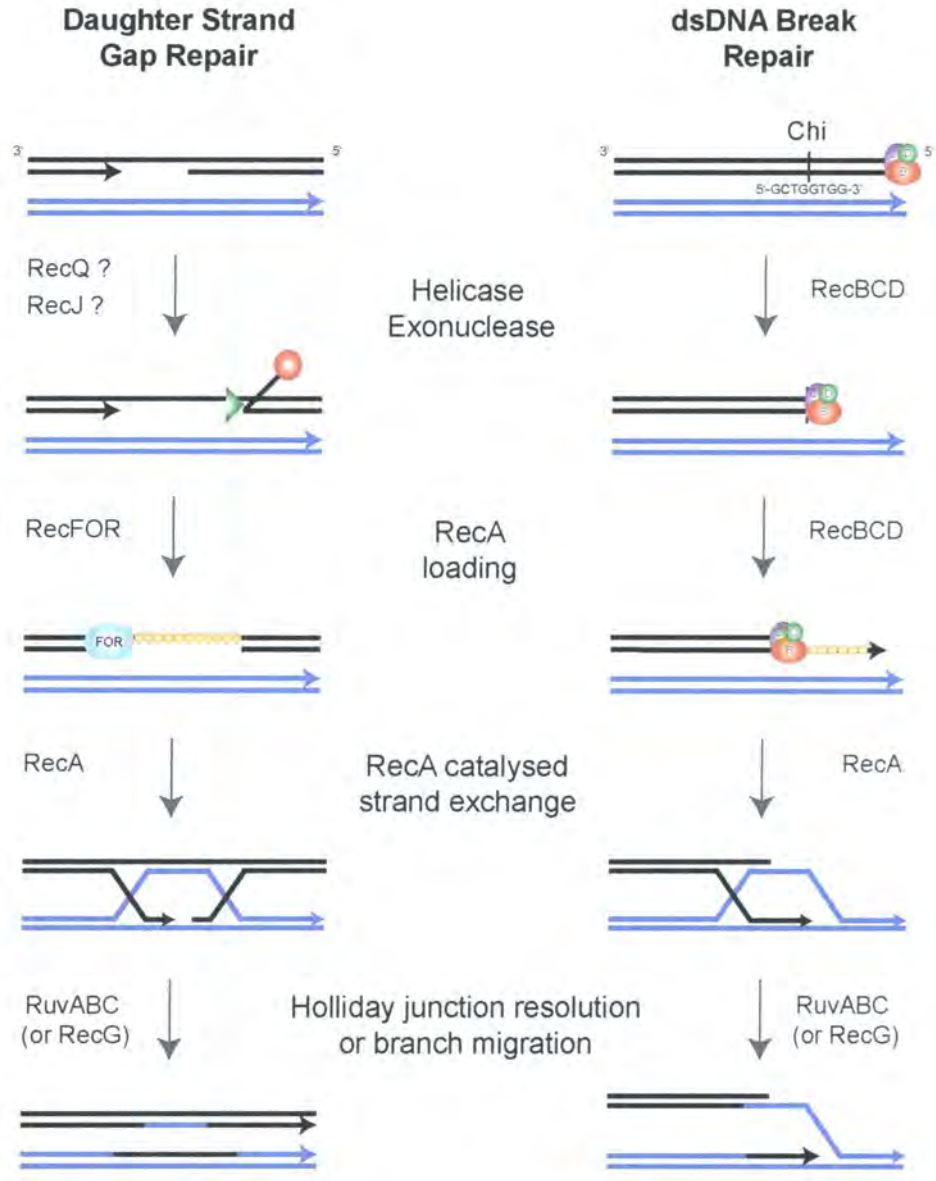

Figure 1.2. Models for the action of *E. coli* recombination proteins at ssDNA gaps and dsDNA ends. Enzymes with known biochemical activities are indicated. The presynaptic step, resulting in the formation of a RecA filament, entails unwinding of the duplex by helicases, followed by exonuclease degredation of DNA to give single strands to which SSB binds. In dsDNA break repair, RecBCD holoenzyme has both helicase and nuclease activities and facilitates loading of RecA onto SSB-coated ssDNA. In ssDNA gap repair, RecQ may act as the helicase, though no specific helicase is required for gap repair, and the 5'-3' RecJ nuclease expands the ssDNA region. RecFOR are proposed to act together at the dsDNA-ssDNA junction and mediate loading of RecA onto SSB-coated ssDNA. RecA mediates the search for a homologous duplex, and catalyses strand invasion and exchange, the synaptic step. The postsynaptic step involves the migration and resolution of Holliday junctions. RuvAB and RecG can perform the branch migration and resolution is achieved by RuvC (in concert with RuvAB). Note that RecQ and RecJ can potentially function at ends in the absence of RecBCD enzyme. In this scenario, RecFOR would be required to load RecA on SSB-coated ssDNA. Adapted from Rocha et al 2005.

RecA is required to avoid problems associated with the presence of SSB protein (Kowalczykowski *et al.*, 1994; Kuzminov, 1999; Anderson and Kowalczykowski, 1997).

RecBCD is primarily involved in the repair of chromosome breaks, initiating recombination at DNA ends by generating 3' ssDNA overhangs (Kowalczykowski, 2000; Kuzminov, 1999; Myers and Stahl, 1994) (Figure 1.2). RecBCD functions as a holoenzyme, incorporating all of the required functions, helicase activity, nuclease activity and the RecA loading function (Singleton *et al.*, 2004). The complex binds to dsDNA ends, unwinds DNA in the 3' to 5' direction and degrades both strands *via* the RecB helicase and nuclease function (Boehmer and Emmerson, 1991; Yu *et al.*, 1998) until it encounters a χ (chi) site (5'-GCTGGTGG-3') recognised by RecC (Handa *et al.*, 1997). Enzyme activity is modified at the χ site resulting in preferential degradation of the complementary DNA in a 5' to 3' direction to generate 3'-tailed ssDNA (involving RecD helicase function). RecA is loaded onto the resected ssDNA through a physical interaction with RecBCD (Kowalczykowski, 2000; Dillingham *et al.*, 2003; Taylor and Smith, 2003). The substrate preparation step is known as presynapsis and is closely followed by synapsis, catalysed by the RecA protein, during which homologous pairing and strand exchange with the intact sister duplex occur, resulting in the formation of a Holliday junction recombination intermediate (Lusetti and Cox, 2002). Holliday junctions can be eliminated by branch migration using the RecG helicase or by dual strand scission mediated by the RuvABC complex (Sharples et al., 1999b). The nicked duplex products of resolution are sealed by DNA ligase.

While RecBCD action is restricted to DNA ends, the RecFOR pathway appears to function primarily in the repair of daughter strand gaps, a process critical in the restoration of stalled replication forks (Cox, 2001; Kuzminov, 1999) (Figure 1.2). In certain circumstances, namely in a *recBC sbcBC* mutant background, the RecFOR pathway can mediate repair of dsDNA breaks, though at a much reduced frequency compared to that seen with RecBCD (Lloyd and Buckman, 1985; Kushner *et al.*, 1971). A detailed understanding of the RecFOR pathway remains unclear, however, it is thought that the 5' to 3' ssDNA exonuclease RecJ enlarges the ssDNA region, possibly with the help of other helicases (e.g. RecQ or UvrD), and the RecFOR proteins assist loading of RecA onto SSB-coated ssDNA (Figure 1.2). As in the RecBCD pathway, RecA catalyses strand invasion of the homologous duplex DNA, and recombination intermediates are processed by either RuvABC or RecG.

The RecF, RecO and RecR proteins display a wide range of *in vitro* activities. There is good evidence for pair-wise interactions between each of these proteins, though no direct evidence that they form a tricomponent complex (Webb et al., 1997; Umezu et al., 1993; Umezu and Kolodner, 1994; Hegde et al., 1996a; Morimatsu and Kowalczykowski, 2003). RecO binds DNA, associates with SSB and promotes annealing of homologous sequences, even when one of the strands is coated with SSB protein (Luisi-DeLuca and Kolodner, 1994; Luisi-DeLuca, 1995; Kantake *et al.*, 2002). RecF protein binds either ssDNA or dsDNA and has a weak ATPase activity important in its dissociation from dsDNA; several of these activities are enhanced by RecR binding (Griffin and Kolodner, 1990; Webb *et al.*, 1995; Webb *et al.*, 1999). RecR proteins from *Bacillus subtilis* and *Deinococcus radiodurans* have been shown to bind DNA and the RecR protein from the latter forms a ring-shaped tetramer with

a central hole that could potentially accommodate dsDNA (Lee *et al.*, 2004; Alonso, 1993). However, the structure of *E. coli* RecR has yet to be determined and the protein does not appear to bind DNA (Webb *et al.*, 1995; Umezu and Kolodner, 1994). Together, RecF and RecR limit the extension of RecA on duplex DNA after the initial filament assembly on ssDNA and RecOR help RecA regain access to ssDNA blocked by the presence of SSB protein (Webb *et al.*, 1997). Specific loading of RecA protein at the 5' end of a dsDNA-ssDNA intersection (i.e. a daughter strand gap) when SSB coats the ssDNA is mediated by the RecF, RecO and RecR proteins acting in concert (Morimatsu and Kowalczykowski, 2003). No interaction between RecF and RecO was detected in these experiments, suggesting that RecR is the key mediator of RecF and RecO association at dsDNA-ssDNA junctions. RecFOR also act to safeguard the nascent lagging strand by loading RecA filaments, thus maintaining the replication fork structure following UV-irradiation (Chow and Courcelle, 2004). This activity may be detrimental to survival in certain mutant backgrounds where DNA replication is impaired (Moore *et al.*, 2003; Petit, 2002). A model for RecFOR assembly at dsDNA-ssDNA junctions is illustrated in Figure 1.3.

1.7 λ Red recombination

The *E. coli recA* gene was discovered in 1965 (Clark and Margulies, 1965) and its mutation eliminates RecBCD and RecFOR mediated recombination pathways (Figure 1.2) (Clark, 1973; Horii and Clark, 1973). Shortly after its discovery it was found that λ recombination was unaffected by mutation of the *recA* gene (Brooks and Clark, 1967; Takano, 1966; van de Putte *et al.*, 1966). This discrepancy was explained by the discovery that phage λ encodes its own recombination system,

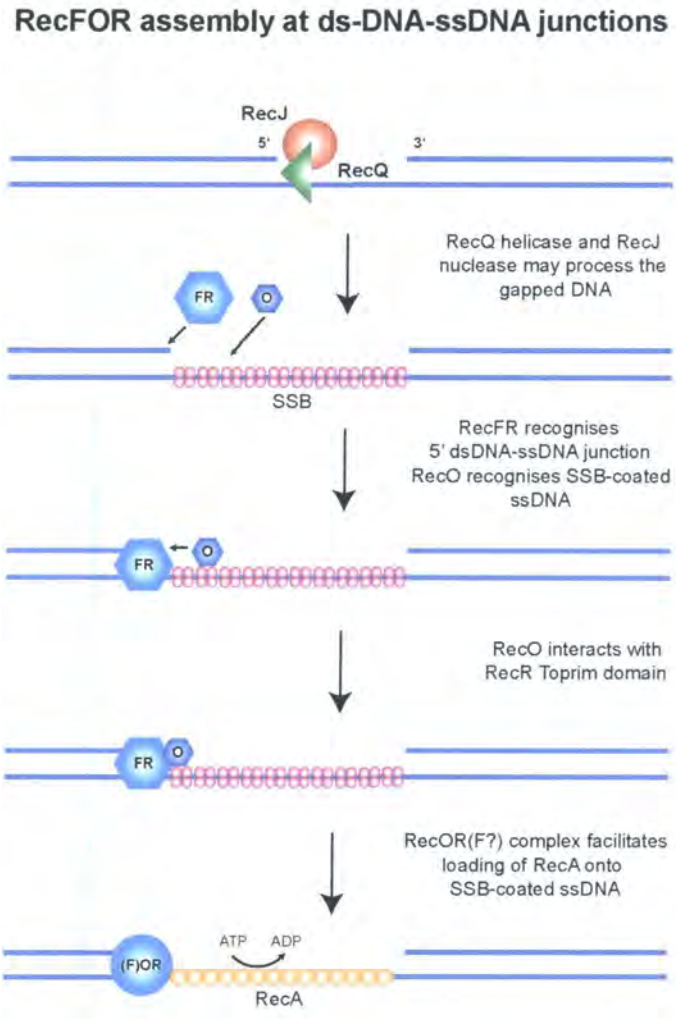

Figure 1.3. Model for assembly of RecFOR at dsDNA-ssDNA junctions
The process of RecFOR assembly at DNA gaps has not been fully characterised. The current model for their action is illustrated above. When ssDNA gaps are generated as a consequence of DNA damage (e.g by skipping an Okazaki fragment), the RecQ helicase and RecJ (5'-3') nuclease process the gap prior to recognition by the RecFOR proteins (Courcelle, 1999). The RecFR heterohexamer complex (Honda, 2005) recognises and binds to the 5'-ended dsDNA-ssDNA junction and RecO interacts with the SSB-coated ssDNA region (Morimatsu, 2003; Luisi-DeLuca, 1994; Umezu, 1994). RecO binds to the Toprim domain of RecR in the RecFR complex and may compete with RecF for binding to produce a RecOR complex (Honda, 2005). The RecOR complex, which may retain RecF, facilitates loading of RecA onto SSB-coated ssDNA. DNA damage is subsequently repaired by homologous recombination. RecA catalyses strand invasion and exchange with a homologous duplex resulting in the formation of recombination intermediates. Resolution of these structures is mediated by RuvABC or RecG. Adapted from Honda, 2005.

consisting of three adjacent genes (*exo*, *bet* and *gam*), collectively termed the Red system. The *red* genes lie within the *pL* operon and are expressed early in lambda development after infection of a new host cell or following prophage induction. Mutation of λ *exo* (*redα*) and *bet* (*redβ*) genes eliminates all homologous recombination in *recA* mutant strains (Echolas and Gingery, 1968; Franklin, 1967). The *gam* gene product serves to inhibit two host nucleases, RecBCD and SbcCD, both of which are involved in recombinational repair of dsDNA breaks (Karu *et al.*, 1975; Murphy, 1991; Chalker *et al.*, 1988; Kulkarni and Stahl, 1989). Gam ensures the preservation of phage genomic DNA during rolling circle replication (see section 1.8).

λ Exo is a 24 kDa protein with 5'-3' exonuclease activity (Carter and Radding, 1971; Little, 1967). It requires a dsDNA end, nicks or gaps are not targeted, and degradation of the 5'-containing strand produces long 3' ssDNA overhangs (Cassuto *et al.*, 1971; Cassuto and Radding, 1971; Carter and Radding, 1971; Hill *et al.*, 1997; Little, 1967). Mononucleotides are released in a highly processive manner at a rate of 1000 bases per second (Matsuura *et al.*, 2001). The biologically active form of Exo is a trimer arranged as a ring so dsDNA can be accommodated into the tapering cavity while the ssDNA product is extruded through the central channel (Kovall and Matthews, 1997) (Figure 1.4). Since λ Gam effectively blocks all *E. coli* RecBCD exonuclease activities at DNA ends, Exo acts as an efficient replacement for the production of ssDNA suitable for recombination.

The *bet* (*redβ*) gene encodes the 29 kDa β protein, which binds stably to ssDNA (Radding, 1971) and protects it from attack by single-strand specific nucleases (Karakousis *et al.*, 1998; Muniyappa and Radding, 1986). β does not bind

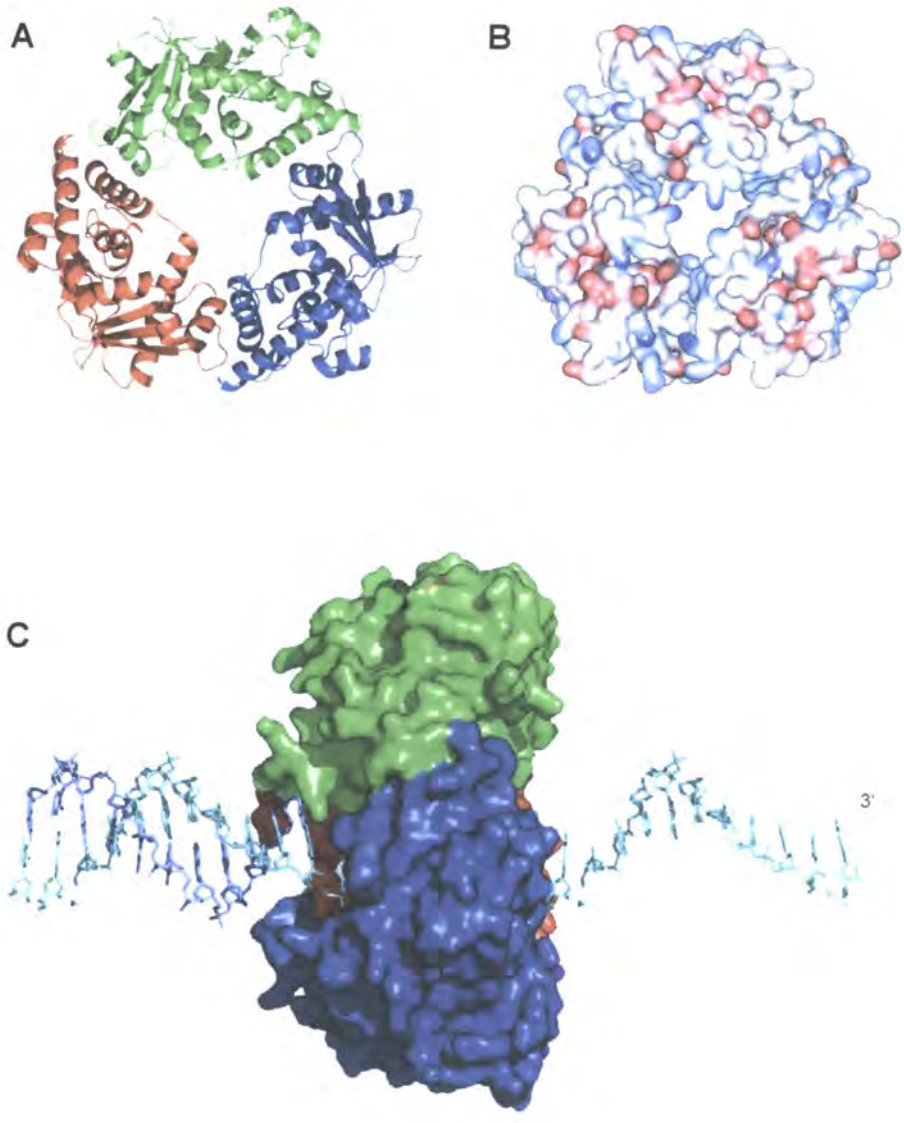

Figure 1.4. Crystal structure of the λ Exo protein.
A. Ribbon diagram of the trimeric λ exonuclease structure. **B.** Electrostatic surface potential of the λ exonuclease, blue: positively-charged, red: negatively-charged. **C.** Model for the interaction of λ Exo with DNA. The trimer has a toroidal structure with a tapered central cavity that has the potential to accommodate dsDNA at the wider end but only ssDNA at the other. The enzyme recognises dsDNA ends, binds and processively degrades DNA in the 5' to 3' direction. The 3'-tailed ssDNA product is extruded through the central cavity. The protein requires a free dsDNA end and Mg^{2+}, and degrades 1000 bases per second. Adapted from Kovall and Mathews, 1997. Figure created using MacPyMOL.

dsDNA directly but it does promote annealing of complementary ssDNA strands where it remains tightly associated with the duplex product of annealing (Karakousis *et al.*, 1998; Li *et al.*, 1998; Muniyappa and Radding, 1986). Under certain circumstances, β protein can promote strand exchange by displacing a strand of duplex DNA. However, it cannot carry out the duplex strand invasion reactions typical of those performed by RecA (Li *et al.*, 1998). β is also thought to modulate Exo nucleolytic and recombination promoting activities and the two proteins appear to form a complex (Cassuto *et al.*, 1971; Cassuto and Radding, 1971; Kmiec and Holloman, 1981; Muniyappa and Radding, 1986; Passy *et al.*, 1999). The β protein belongs to a family of annealing proteins which includes RecT from the *E. coli* cryptic Rac prophage (Iyer *et al.*, 2002). RecT is a functional analog of β and the two proteins share 15% sequence identity. The Rac prophage also specifies a functional equivalent of λ Exo called RecE. In *E. coli* strains carrying the Rac prophage, *sbcA* mutations upregulate the expression of RecET and allow efficient recombination in the absence of RecBCD (Kolodner *et al.*, 1994; Aravind *et al.*, 2000). There is currently no representative crystal structure available for this family of proteins, although the structure of β protein has been studied by electron microscopy (Passy *et al.*, 1999). The protein forms multisubunit rings in solution or in the presence of ssDNA and forms left-handed helical filaments when two complementary strands of ssDNA are annealed (Passy *et al.*, 1999). Merging of β protein rings probably gives rise to the helical filaments observed in the presence of complementary oligonucleotides. Recent deletional analysis suggests that each β subunit contains a stable N-terminal core, a flexible central domain involved in binding DNA and a largely disordered C-terminal domain (Wu *et al.*, 2006). The human Rad52 protein,

which has similar ssDNA renaturation properties to β and RecT (Passy *et al.*, 1999; Singleton *et al.*, 2002), also forms ring-like structures (Figure 1.5). Its single-strand binding site forms a positively charged groove around the central hub, leaving nucleotide bases accessible for annealing to complementary ssDNA (Singleton *et al.*, 2002; Iyer *et al.*, 2002). It seems likely that Rad52, RecT and β share a structurally similar mode of ssDNA binding and annealing (Passy *et al.*, 1999).

1.8 Recombination pathways in λ

Two recombination pathways predominate in phage λ, strand-annealing and strand-invasion; both involve the restoration of a genomic dsDNA break by exchange with a second λ genome (Stahl *et al.*, 1997) (Figure 1.6) and both require the presence of Exo and β (Muyrers *et al.*, 2000). *In vivo* experiments have revealed that these enzymes can promote exchange at relatively short regions of homology (≥50 bp) (Murphy, 1998). When linear λ DNA is injected into the bacterial host it circularises, and assuming it follows the lytic cycle, begins to replicate in the θ mode (Figure 1.7). Subsequent displacement of the 5' end of the leading strand allows replication in the σ mode, referred to as rolling circle replication (Taylor and Wegrzyn, 1995; Baranska *et al.*, 2001; Taylor, 1998). The Gam protein specifically inhibits RecBCD and SbcCD host nucleases to preserve the ends of the phage genome during rolling circle replication.

The strand-invasion pathway depends on RecA for strand exchange utilising the 3' ssDNA tails generated by Exo at a dsDNA end (Figures 1.6 and 1.8). RecA, and possibly β, initiate strand pairing between the ssDNA and the homologous duplex resulting in the formation of a recombinant joint (Passy *et al.*, 1999; Stahl *et*

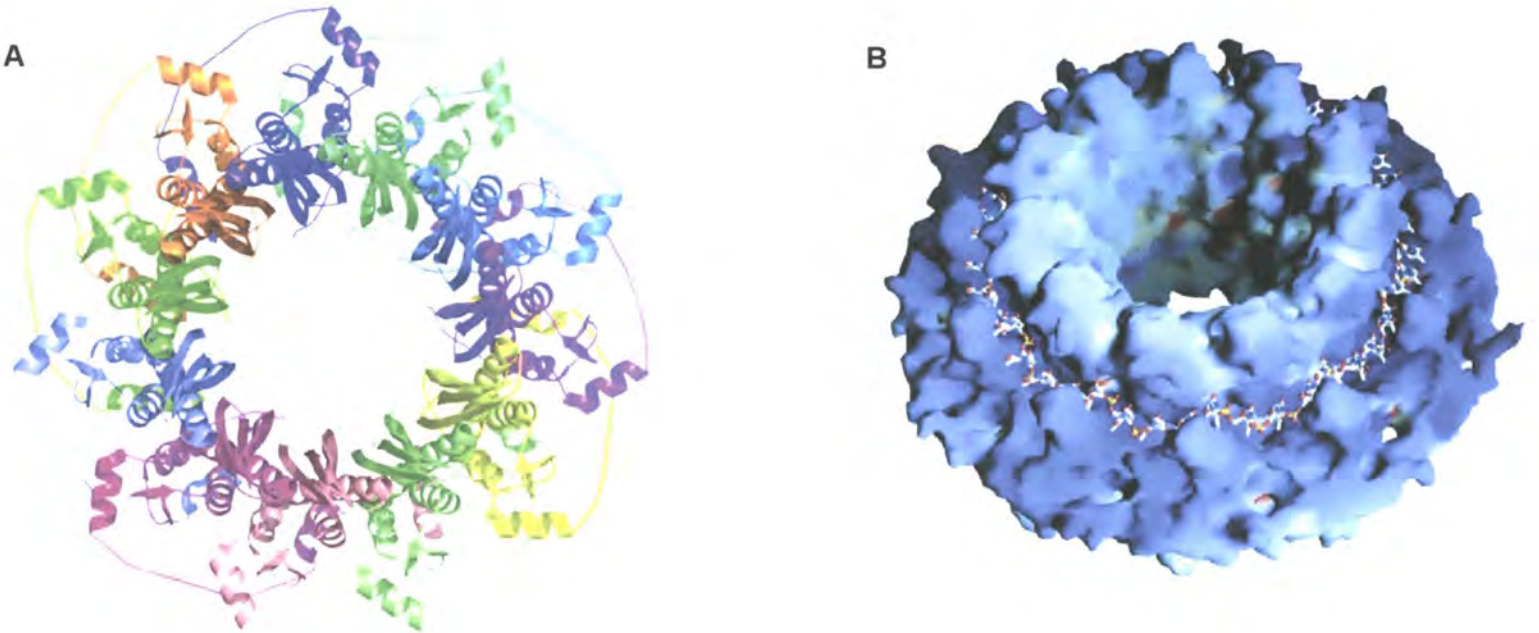

Figure 1.5. Ring-like structure of the Rad52 (residues 1-209) undecamer.
A. The RAD52 undecamer viewed along the 11-fold symmetry axis. **B.** Molecular surface representation of the protein. Surface charge was determined using *GRASP;* Blue: positively-charged, Red: negatively-charged. The positively charged surface groove encircling the ring forms a potential ssDNA binding site. DNA is modelled onto the image for reference. A similar ring-like structure may be adopted by β protein. Adapted from Singleton et al, 2002.

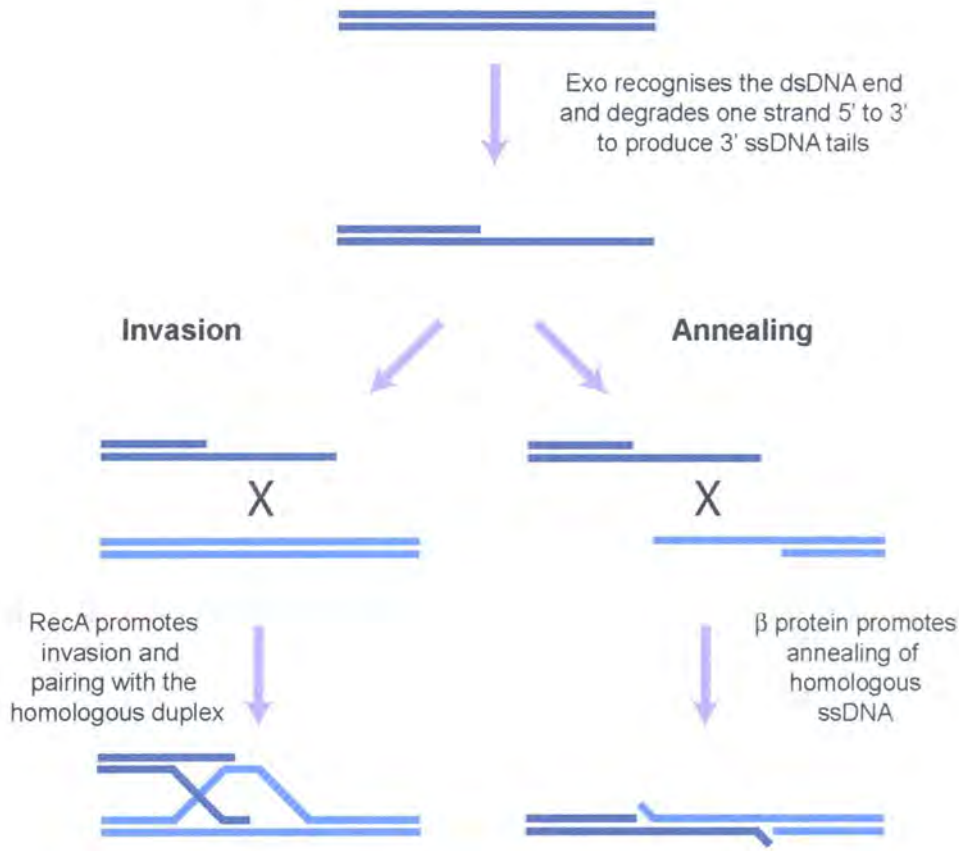

Figure 1.6. Model for recombination pathways in phage λ.
The two pathways of recombination which predominate in λ. In both pathways, Exo recognises and degrades dsDNA ends in the 5' to 3' direction, resulting in the formation of SSB-coated 3'-ssDNA ends. The invasion pathway is RecA-dependent and is probably the major pathway in wild-type *E. coli*. RecA is loaded onto SSB-coated DNA by RecFOR and catalyses strand exchange. The annealing pathway functions independently of RecA. In this pathway, the β protein coats the 3'-overhanging ssDNA and promotes annealing with a complementary ssDNA.

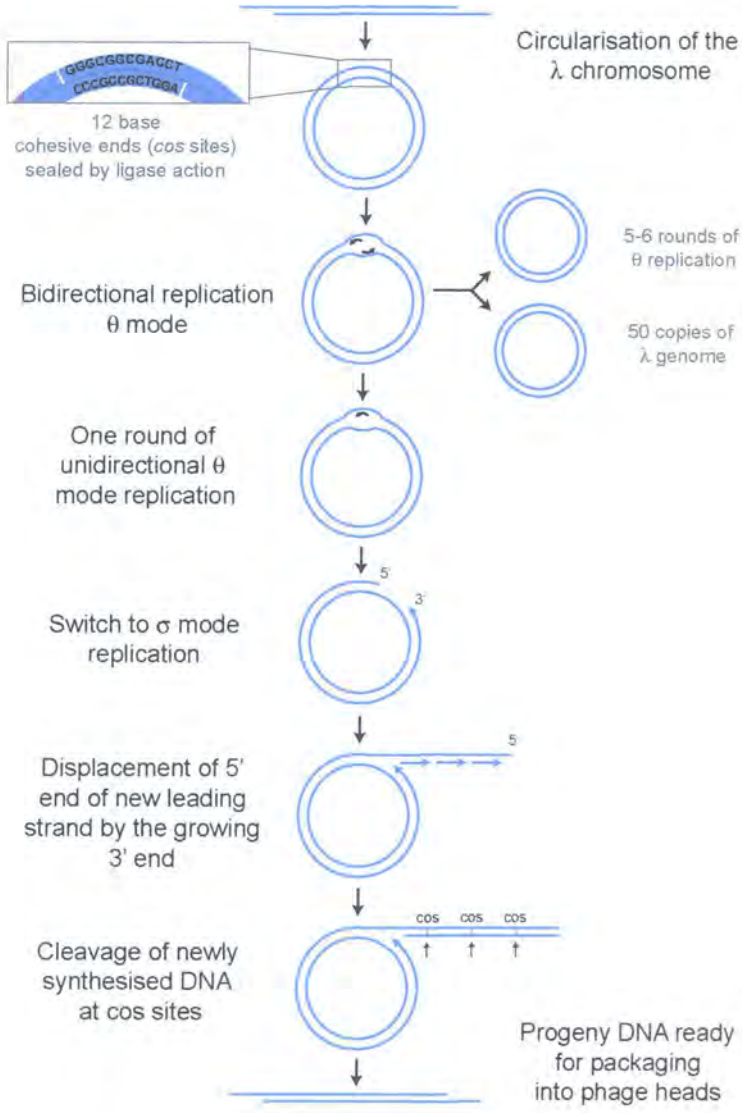

Figure 1.7. Model for phage λ replication.
Upon injection of λ linear dsDNA into the host cell it circularises by pairing of the cohesive ends followed by sealing of the nicks by DNA ligase. λ O and P replication proteins are synthesised. O binds to the origin of replication *oriλ* and directs viral and host proteins to this site in the process of replication complex assembly. *E. coli* DnaB helicase is involved, as is DnaA which is thought to mediate stimulation of transcriptional activation of *oriλ*. Bidirectional θ replication of λ DNA is initiated and continues for 5-6 rounds, producing around 50 copies of the λ genome. The switch from θ to σ (rolling-circle) replication is thought to be influenced by host DnaA. Due to the number of λ genomes present in the cell, cellular DnaA protein is titrated out and activation of *oriλ* is less efficient. One round of unidirectional θ replication is initiated followed by displacement of the 5' end of newly synthesised leading strand by its growing 3' end. σ mode replication leads to the production of long concatemers of λ DNA, up to10 genome equivalents in length, which are cut at cos sites and serve as substrates for the phage packaging machinery. Note at the switch from θ to σ replication, only the fate of the parental strand directing the synthesis of the leading strand is shown. Reviewed in Taylor and Wegrzyn 1995,1998.

al., 1997). The resulting displacement loop (D-loop) contains a 3'-end that can prime DNA replication, and extensive DNA synthesis is a potential escape mechanism at this stage. Alternatively, strand cleavage at appropriate points in the D-loop may generate a splice-type recombinant product (Stahl *et al.*, 1985; Hill *et al.*, 1997), whereas further strand exchange into the adjacent duplex DNA of the invading molecule creates a Holliday junction (Lloyd, 1996). Resolution of the Holiday junction and subsequent DNA replication will repair the double-strand break and restore intact chromosomes (Takahashi and Kobayashi, 1990; Szostak *et al.*, 1983). The λ Rap structure-specific endonuclease could potentially eliminate any of the branched intermediates formed during these exchanges, while the host RuvC endonuclease could resolve any Holliday junctions (Sharples *et al.*, 2004; Sharples, 2001). Properties of the Rap endonuclease will be discussed more fully in Chapter 8.

The strand-annealing pathway functions independently of RecA and involves, as its name suggests, pairing of homologous ssDNA partner sequences (Figure 1.6). This reaction requires two dsDNA breaks to be located at different sites in separate λ genomes and depends upon the ssDNA annealing properties of β protein (Stahl *et al.*, 1997; Poteete, 2001). As in the strand-invasion reaction, Exo resects the overlapping double-strand ends to yield complementary single strands. β protein then anneals the two strands and continuing Exo degradation trims the joint so that excess 3' single strand branches are assimilated into duplex DNA (Figure 1.6). Sealing of the resulting nicked duplex by DNA ligase completes the process and the fused genomes can be packaged into phage heads (Cassuto *et al.*, 1971; Cassuto and Radding, 1971; Kmiec and Holloman, 1981; Muniyappa and Radding, 1986).

1.9 Role of λ Orf in phage recombination

Studies of λ recombination in *E. coli* strains lacking the RecFOR pathway led to the discovery of a λ equivalent of the *E. coli recF*, *recO*, and *recR* genes (Sawitzke and Stahl, 1992). RecFOR function is necessary for bacterial conjugational recombination in *recBC sbcBC* strains. However, they are superfluous in recombination between λ chromosomes in this background, even when λ lacks its own Red recombination system and is dependent upon host recombinases (Sawitzke and Stahl, 1992; Sawitzke and Stahl, 1994). The functional analog responsible for replacing RecFOR activity was found to be a single reading frame (*ninB*) and subsequently designated *orf* for suppression of RecO, RecR and Rec F phenotypes. Orf maps to the 3612 bp *ninR* region of the λ genome within a chain of overlapping reading frames (Hollifield *et al.*, 1987; Kroger and Hobom, 1982). This accessory gene segment is conserved in lambdoid phage genomes (see Chapter 3), despite the fact it can be deleted in its entirety without loss of phage viability. The functional relevance of most of the putative *ninR* genes remains unknown, however two, *orf* (*ninB*) and *rap* (*ninG*), have been shown to participate in genetic recombination.

Orf substitutes for the RecFOR complex in λ crosses but not during host conjugational exchange (Sawitzke and Stahl, 1992; Sawitzke and Stahl, 1994; Tarkowski *et al.*, 2002). It stimulates the formation of recombinational crossovers close to an initiating double-strand breaks in both the RecF and Red pathways (Tarkowski *et al.*, 2002; Sawitzke and Stahl, 1997). Orf can also partially substitute for *recFOR* mutants in *E. coli* recombination when Exo and β are present (i.e. Red-mediated) (Poteete, 2004). The ability of Orf to substitute for RecFOR under certain conditions suggests that it may perform the same mechanistic function as the *E. coli*

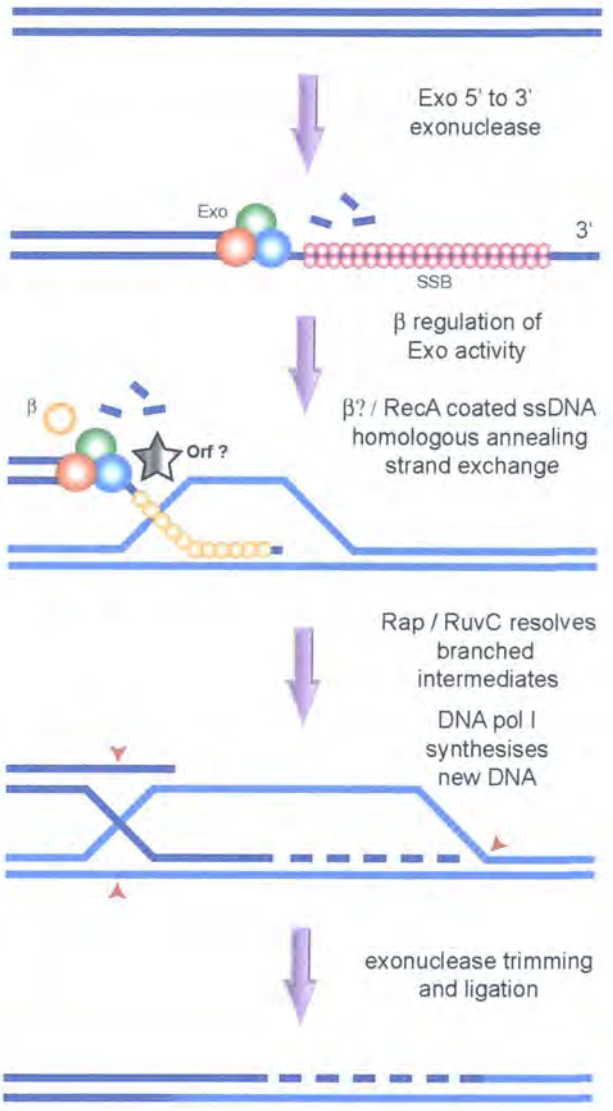

Figure 1.8. Model for the Red-mediated strand invasion pathway of recombination.
Lambda exonuclease requires a dsDNA to begin digestion of a single strand in the 5'-3' direction, producing 3'-tailed ssDNA. The β protein is thought to interact with Exo and modulate its activities. β can also bind ssDNA and may coat the 3'-ends to promote annealing between homologous strands. Although β can promote strand exchange, it cannot perform the duplex invasion reactions carried out by RecA. It is not known if β participates in the invasion pathway of λ recombination, it may only be required to regulate Exo activity when RecA is available. λ encodes is own structure-specific endonuclease, Rap, which may resolve the branched intermediates generated during λ recombination. The function of λ Orf is unknown. Based upon its ability to supply a function analogous to the *E. coli* RecFOR, it is probably involved in mediating the loading of RecA onto SSB- or β-coated ssDNA.

complex (Hegde et al., 1996a; Morimatsu and Kowalczykowski, 2003; Wang, 1993; Bork et al., 2001; Umezu et al., 1993; Umezu and Kolodner, 1994). Thus, Orf is likely to contribute to the initial stages of daughter strand gap repair, possibly by displacing SSB to allow loading of RecA or β proteins (Figure 1.8).

The objective of this study was to elucidate how the 17 kDa λ Orf protein could substitute for a complex of three (RecFOR) much larger host proteins. The Orf protein will be purified and biochemical analyses conducted to investigate potential interactions with DNA and SSB. The ultimate aim was to investigate its role in ongoing strand exchange reactions mediated by Red and RecA. The results should offer fresh insights into the molecular mechanisms of recombination in phage λ at legitimate and illegitimate sites and how such exchanges influence the emergence of new pathogenic bacteria.

Chapter 2

Materials and Methods

2.1 Computer Software

The text for presentation of this thesis was written in Microsoft Word. Figures were prepared using Adobe Photoshop and Adobe Illustrator. Graphs and numerical analyses of data were produced in Excel and charts converted for figures in pro Fit. Percentage DNA binding/cleavage was calculated by quantifying ^{32}P-labelled substrate visualised by Phosphor Imaging using Image Gauge. Images of protein structures were generated using PyMol (DeLano, W.L., The PyMOL Molecular Graphics System (2002), http://www.pymol.org). Protein sequences were obtained from the National Center for Biotechnology Information (NCBI, www.ncbi.nlm.nih.gov) and potential homologs of phage λ Orf were identified by the PSI-BLAST algorithm (Altschul *et al.*, 1997) using default parameters. Sequence alignments were constructed using ClustalW (Chenna *et al.*, 2003). The following residues were considered to be functionally related: (V, I, L, M), (F, Y, W), (K, R), (D, E), (S, T), (G, A) and (N, Q). Secondary structure predictions and alignments were generated using the Phyre protein fold recognition server (Kelley *et al.*, 2000). Protein structures were obtained from the Protein Data Bank (www.rcsb.org/pdb).

2.2 Materials

2.2.1 Chemicals and reagents

Analytical grade chemicals and reagents used in this study were obtained from Sigma-Aldrich, BDH Laboratory Supplies, Fischer Science, New England Biolabs, Amersham-Pharmacia, Difco and Invitrogen.

2.2.2 Growth media and antibiotics

Luria-Burrows (LB) broth contained, 10 g Difco tryptone, 5 g Difco yeast extract, 0.5 g NaCl, and 2 ml 1M NaOH (pH 7) made up to 1L SDW and autoclaved at 121°C for 15 min. LB media was solidified with 15 g/l Difco agar. For growth of BL21-SI strains, LB media was made without NaCl (LBON). Strains carrying a pMALc2 expression vector were cultured in LB broth or agar containing 2 g/L glucose.

Antibiotic stocks were made in sterile distilled water (SDW), unless specified otherwise, and stored at 4°C. Antibiotics were added to media at the following concentrations: Ampicillin (Ap) at 150 μg/ml, chloramphenicol (Cm) at 50 μg/ml, and kanamycin (Km) at 40 μg/ml.

2.2.3 Bacterial strains and plasmids

Escherichia coli strains used in this work are listed in Table 2.1 and plasmids listed in Table 2.2. The plasmids specifically constructed for this work (designated pPR) were generated using standard molecular biology techniques. Plasmids pPR100, pPR101, pPR111, and pPR112 were generated by subcloning using the restriction enzymes indicated in parenthesis (Table 2.2). pPR109, pPR110, pPR113 and pPR114

were constructed by PCR amplification of the relevant sequences using oligonucleotides matching the 5' (containing NdeI or EcoRI restriction sites) or 3' (containing BamHI or HindIII restriction sites) ends of the required gene. Restriction of the DNA product with NdeI/BamHI, EcoRI/HindIII or EcoRI/BamHI was followed by ligation into the appropriately cleaved vector.

Table 2.1 Genotype and sources of *E. coli* strains

Strain	Genotype	Reference
AB1157	*thr-1, ara-14, leuB6, Δ(gpt-proA)62, lacY1, tsx-33, qsr'-, glnV44*(AS), *galK2*(Oc), λ⁻, *rac*⁻, *hisG4*(Oc), *rfbD1, mgl-51, rpoS396*(Am), *rpsL31*(strR), *kdgK51, xylA5, mtl-1, argE3*(Oc), *thi-1*	(Bachmann, 1972)
DH5α	F'/*endA1, hsdR17*(r_k-m_k), *glnV44, thi-1, recA1, gyrA* (Nal^r), *relA1, Δ(lacIZYA-argF)U169, deoR, (φ80dlacΔ(lacZ)M15)*	(Hanahan, 1983)
BL21(DE3)	F⁻ *ompT hsd*S_B($r_B^-m_B^-$) *gal dcm* (λcl*ts*857 *lac*UV5-T7 gene *1 ind1 Sam7 nin5*)	(Studier and Moffatt, 1986)
BL21-SI	*F⁻ ompT lon hsdSB*($r_B^-m_B^-$) *gal dcm endA pro*U-T7 RNAP:: *malQ-lacZ* Tet^s	(Hanahan, 1983) and US Patent
BL21-Codonplus	*F⁻ ompT hsdS*($r_B^-m_B^-$) *dcm*⁺ Tet^R *gal* λ(DE3) *endA Hte [argU ileY leuW* Cam^R*]*	(Weiner, 1994)
SR2210	F⁻ as AB1157 but *ruvA200*	(Sargentini and Smith, 1989)
N1057	F⁻ as AB1157 but *ruvB4*	(Sharples *et al.*, 1990)
N2057	F⁻ as AB1157 but *ruvAB60*::Tn*10*	(Shurvinton *et al.*, 1984)
GS1922	F⁻ as AB1157 but Δ*ruvC*::cat	(Lloyd, 1991)
GS2053	F⁻ as AB1157 but Δ*ruvC64*::kan	(Mandal *et al.*, 1993)

Strain	Genotype	Reference
CS85	F⁻ as AB1157 but *ruvC53 eda*51::Tn10	(Shurvinton *et al.*, 1984)
N4454	F⁻ as AB1157 but Δ*ruvABC*::cat	(Seigneur *et al.*, 1998)

Table 2.2 Plasmids

Plasmid	Description	Source/Reference
pT7-7	Expression vector with T7 ϕ10 promoter, Ap^R	(Tabor and Richardson, 1985)
pET-14b	Expression vector for tagging proteins with N-terminal His_6, Ap^R	Novagen
pET-22b	Expression vector for tagging proteins with N-terminal His_6, Ap^R	Novagen
pET-24a	Expression vector for tagging proteins with N-terminal His_6, Km^R	Novagen
pMALc2	Expression vector for tagging proteins with maltose binding protein (MBP), Ap^R	NEB
pUC18	Cloning vector, Ap^R	MBI Fermentas
pHP13	Shuttle plasmid with replication functions from pTA106 (pUC derived), Cm^R	DSMZ
pACYC184	Control vector for pGS739, Cm^R	(Sharples *et al.*, 1990)
pHSG415	Control vector for pPVA101, Ap^R	(Sharples and Lloyd, 1991)
pGS903	λ *orf* in pT7-7 (NdeI-BamHI), Ap^R	GJ Sharples
pGS904	*E. coli orf151* in pT7-7 (NdeI-BamHI), Ap^R	GJ Sharples
pPR100	λ *orf* in pET-14b (NdeI-BamHI), Ap^R	This work
pPR101	*E. coli orf151* in pET-14b (NdeI-BamHI), Ap^R	This work
pPR109	λ *orfΔC6* in pT7-7 (NdeI-BamHI), Ap^R	This work
pPR110	λ *orfΔC19* in pT7-7 (NdeI-BamHI), Ap^R	This work
pPR111	λ *orfΔC6* in pET-14b (NdeI-BamHI), Ap^R	This work
pPR112	λ *orfΔC19* in pET-14b (NdeI-BamHI), Ap^R	This work
pPR113	*S. aureus* phage ϕETA *orf20* in pMALC2 (EcoRI-HindIII), Ap^R	This work
pPR114	λ *orf* in pMALc2 (EcoRI-BamHI), Ap^R	This work
pLB101	*E. coli orf151* in pMALc2 (NdeI-EcoRI), Ap^R	Laura Bowers

Plasmid	Description	Source/Reference
pCC146	*E. coli ssb* in pET-22b, Ap^R	(Cadman and McGlynn, 2004)
pCC180	*E. coli ssbΔC10* in pET-22b, Ap^R	(Cadman and McGlynn, 2004)
pFC149	λ *bet* in pET-14b (NdeI-BamHI), Ap^R	Dr Fiona Curtis
pFC150	λ *exo* in pET-14b (NdeI-BamHI), Ap^R	Dr Fiona Curtis
pGTI4	*E. coli ruvA* in pUC18, Ap^R	(Sharples *et al.*, 1990)
pGTI19	*E. coli ruvB* in pUC18, Ap^R	(Sharples *et al.*, 1990)
pGS711	*E. coli ruvA* and *ruvB* in pUC18, Ap^R	(Sharples *et al.*, 1990)
pGS739	*E. coli ruvC* in pACYC184, Cm^R	(Sanchez *et al.*, 2005)
pGS775	*E. coli ruvC* in pT7-7, Ap^R	(Sharples *et al.*, 1990)
pCB593	*B. subtilis ruvA* in pUC18, Ap^R	(Sanchez *et al.*, 2005)
pCB594	*B. subtilis ruvB* in pUC18, Ap^R	(Sanchez *et al.*, 2005)
pCB559	*B. subtilis ruvAB* in pUC18, Ap^R	(Sanchez *et al.*, 2005)
pCB564	*B. subtilis recU* in pHP13, Cm^R	(Sanchez *et al.*, 2005)
pCB620	*B. subtilis recU* (E36A) in pUC18, Ap^R	JC Alonso
pCB624	*B. subtilis recU* (H119A) in pUC18, Ap^R	JC Alonso
pFC105	Phage bIL67 *ruvC* in pET-24a (NdeI-BamHI), Km^R	FA Curtis
pFC182	Phage bIL67 *ruvC* D8N mutant in pET-24a (NdeI-BamHI), Km^R	FA Curtis
pFC120	*H. pylori ruvC* in pT7-7, (NdeI-BamHI), Ap^R	FA Curtis
pPVA101	*E. coli ruvABC* in pHSG415, Ap^R	(Sharples and Lloyd, 1991)
pMP101	*H. pylori ruvC* in pMALc2, Ap^R	GJ Sharples
pLysS	Encodes T7 lysozyme active against T7 RNA polymerase, Cm^R	(Studier and Moffatt, 1986)

2.2.4 Oligonucleotides

DNA oligonucleotides (HPSF-purified) were purchased from MWG-Biotech AG. Those used for PCR (Table 2.3) were synthesised at the 0.05 μmol scale and those for preparation of DNA substrates (Table 2.4) at 0.2 μmol. The DNA was supplied as a precipitate and resuspended in an appropriate volume of TE prior to use.

Table 2.3 Oligonucleotides used for PCR amplification

Name	Nucleotide Sequence (5'-3')	Restriction Site
NinB1	GAGAGGGAACATATGAAAAAACTAA	NdeI
NinB2	CCTGCCACCGGATCCACTAACGACA	BamHI
MBP-NinB1	AGAGGGAATTCATGAAAAAACTAAC	EcoRI
NinB-C6	CATGGATCCCTGTCTCCTCATCTCG	BamHI
NinB-C19	GTCTGGATCCGTCTCACCACTTAAC	BamHI
Orf151-1	CGGAGGGGACATATGAACCTCTCAC	NdeI
Orf151-2	GACGATTTGGATCCCTGTAGATGTG	BamHI
Orf151-3	CGGAGGGAATTCATGAACCTCTCAC	EcoRI
ETA20-1	ACGTGGTTCATATGCAATACATTAC	NdeI
ETA20-2	CTTCCGCCAAGCTTACGATTAGGAG	HindIII
ETA20-C82	TTACAAGCTTGCGCTAGATTGTAGC	HindIII
MBP-ETA20-1	ACGTGGAATTCATGCAATACATTAC	EcoRI
Exo-1	TGGCCATATGACACCGGACATTATC	NdeI
Exo-2	TGAGGATCCGTCATCGCCATTGCTC	BamHI
Bet-1	TAAAACATATGAGTACTGCACTCGC	NdeI
Bet-2	TGCAGGATCCTGTCCGGTGTCATGC	BamHI

Name	Sequence 5' - 3'	Length (nt)	Reference
NH7	GGCGACGTGATCACCAGATGATTGCTAGGCATGCTTTCCGCAAGAGAAGC	50	(Whitby,1996)
NH8	GGCTTCTCTTGCGGAAAGCATGCCTAGCAATCCTGTCAGCTGCATGGAAC	50	(Whitby,1996)
NH11	GCTTCTCTTGCGGAAAGCATGCCTAGCAATCATCTGGTGATCACGTCGCC	50	(Whitby,1996)
FAC5	GATTACATTGCTAGGACATCTTTGC	25	This work
FAC6	CCACGAACGTCATAGACGATTACATTGCTAGGACATCTTTGCCCACGTTGACCCAAGTCG	60	This work
FAC7	CGACTTGGGTCAACGTGGGCAAAGATGTCCTAGCAATGTAATCGTCTATGACGTTCGTGG	60	This work
FAC8	GTCTATGACGTTCGTGG	17	This work
FAC9	CGACTTGGGTCAACGTG	17	This work
ϕX174-25	GTAAGAGCTTCTCGAGCTGCGCAAG	25	This work
ϕX174-50	GAAAGGTCGCAAAGTAAGAGCTTCTCGAGCTGCGCAAGGATAGGTCGAAT	50	This work
GF-51	CTACCGGTTGGTCACGGGTGACCATTGCTGAAAAACTCGGCGGCAAACAGC	51	J. Yates

Table 2.4 Oligonucleotides used for DNA binding, cleavage and annealing assays in this work.

2.3 Molecular Biological and Biochemical Techniques

2.3.1 Growth and maintenance of bacterial strains

Primary "overnight" cultures for routine use were prepared by suspending a single colony in 5 ml of LB or LBON broth. Appropriate antibiotics were included to ensure the maintenance of plasmids. Cultures were typically grown at 37°C overnight with gentle aeration (150 rpm) unless stated otherwise. Strains prepared in this way could be stored for 2-4 weeks at 4°C. For longer-term storage, 2.5 ml of a fresh overnight culture was mixed with 1.5 ml of 80% glycerol and stored at -20°C. Experimental cultures, for UV survival studies were typically prepared by diluting overnight cultures 20-fold into 8 ml of fresh media and incubated at 37°C with vigorous shaking. Bacterial cell density was measured at A_{650nm} in a Spectronic 20+ spectrophotometer. For protein overexpression, 500 ml or 1L of fresh media were inoculated with 5-15 ml of an overnight culture and grown in 1L or 2L baffled flasks. Cultures were incubated at 37°C with shaking (typically at 150 rpm). Strains grown on solid media were streaked on LB or LBON agar plates using sterile disposable or Tungsten loops to obtain single colonies. Plates were incubated for 8-16 hours at 37°C and stored at 4°C for up to 3 weeks.

2.3.2 Harvesting bacterial cells from liquid culture

Cells were harvested from liquid culture by centrifugation. Samples ≤1.5 ml were centrifuged in an Eppendorf bench top microcentrifuge at either low (4000-6000 rpm) or high (13000 rpm) speed settings. Larger volumes were harvested at 4°C in a Beckman Coulter Avanti J-E centrifuge using Beckman JA-20 or JA-10 rotors.

2.3.3 Competent cells

Chemically competent *E. coli* cells were obtained by the standard calcium chloride protocol (Sambrook, 2001). Cells, in 40 µl aliquots, were snap frozen in ethanol and dry ice and stored at -80°C. Electrocompetent cells were prepared by diluting 2 ml of overnight culture into 100 ml fresh LB broth and growing to A_{650nm} 0.5-0.7. Cultures were chilled on ice for 20 min before centrifugation at 4000 rpm for 15 min at 4°C. The supernatant was discarded and the cell pellet resuspended in 50 ml of sterile ice-cold 10% glycerol. This step was repeated, resuspending the pellet in decreasing volumes of 10% glycerol to a final volume of 1 ml. Cells were stored after snap freezing in 40 µl aliquots at -80°C.

2.3.4 Bacterial transformations

Bacterial cell transformations were performed using the heat-shock or electroporation methods. Plasmid DNA (1-2 µl of stock samples) was added to 40 µl of competent cells in both protocols. Mixtures for electroporation were incubated on ice for 1 min before transfer to a 1.5 ml cuvette. Cells were pulsed (2.5 Kv for 5.9 ms) in a BioRad MicroPulser™ and 1 ml of SOC medium (2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose) was added immediately. The sample was mixed, transferred to a sterile 1.5 ml Eppendorf tube and incubated at 37°C for 1 hour without shaking to allow expression of the plasmid-encoded antibiotic resistance. Cells were pelleted in an Eppendorf bench top microcentrifuge at 6000 rpm and resuspended in 50-100 µl of fresh LB. The cell suspension was spread on appropriate selective media using sterile plastic spreaders. For transformation by heat-shock, mixtures were incubated on ice

for 30 min, transferred to 42°C for 1 min in a heat block or water bath and returned to ice for 1 min. Cells were rescued in 1 ml SOC or LB and incubated at 37°C for 30-60 min, then prepared for spread plating as described above.

2.3.5 Measurements of *E. coli* UV light sensitivity

Bacteria were grown from fresh overnight cultures in 8 ml of LB broth and incubated at 37°C (unless stated otherwise) with good aeration and appropriate antibiotic selection. Cultures were chilled on ice and 10-fold serial dilutions down to 10^{-5} were made in chilled broth in sterile1.5 ml Eppendorf tubes. A series of dried LB agar plates, containing appropriate antibiotics, were spotted with 10 µl of each dilution and spots allowed to dry before exposure to UV light from a germicidal lamp. Plates were irradiated at 1 J/m^2 for 5, 10, 20, 30, 45 and 60 seconds and an unirradiated plate served as a control. Plates were incubated for 16-24 hours and surviving colonies within appropriate dilution spots (typically 20-100 colonies) were counted. The fraction surviving at each dose was calculated (relative to the unirradiated control) from duplicate experiments and the mean plotted as the log survival against UV dose.

2.3.6 PCR amplification of DNA

DNA for cloning purposes was amplified by polymerase chain reaction (PCR) according to the manufacturers instructions using buffers provided in the PCR kit (Invitrogen™). Typically, 50 µl reactions contained, ~100 pmoles of each oligonucleotide, the required template DNA, 2.5 units Platinum *Pfx* DNA polymerase (Invitrogen™), 1 mM $MgSO_4$, standard reaction and enhancer buffers

(1x final concentration), 2 mM of each of the 4 dNTPs and made up to 50 µl with SDW. Mineral oil (50 µl) was pipetted onto the mixture to prevent evaporation. All amplifications were performed in a Techgene Thermocycler. Programs varied, but typically a three-step cycle (repeated 25-35 times) was employed involving denaturation at 94°C for 40 sec, annealing at 45-55°C for 40 sec and extension at 68°C for 1-3 (1 min for every kilobase amplified). Annealing temperatures varied according to the calculated T_m of the primers, typically the reaction temperature was set at 5°C below the T_m of the primers. If problems were encountered in amplification, annealing temperatures were reduced, the number of cycles increased, primer and $MgSO_4$ concentrations increased or the template concentration varied.

2.3.7 Purification of DNA, restriction digestion and ligation

Plasmid DNA preparations from 1-100 ml cultures and purification of PCR and digestion products from agarose gel slices were performed using kits and protocols supplied by Qiagen. Recovered DNA was stored at -20°C in EB buffer (10 mM Tris-HCl pH 8.5). Restriction enzymes were purchased from Invitrogen and stored at -20°C. Digestion products were separated on TBE-agarose gels (see below) containing 0.4 µl/ml ethidium bromide (EtBr). Digestion products were excised in gel fragments and purified using Qiagen kits. DNA ligations were incubated for at least 8 hours at 16°C in the presence of 1-2 units of T4 DNA ligase (Invitrogen). Ligation mixtures containing a 5:1, 10:1 or 25:1 ratio of insert:vector were introduced to DH5α competent cells and plated on appropriate selective media.

2.3.8 Electrophoresis

i. Agarose gel electrophoresis

1-1.5% w/v agarose gel was prepared by melting powdered agarose in Tris-borate EDTA buffer (TBE: 90 mM Tris-borate pH 7.5, 2 mM EDTA). The solution was allowed to cool to ~55°C before addition of 0.2 μg/ml ethidium bromide to allow visualisation of DNA under UV light. Gels (40 ml) were poured into the casting trays (BioRad Minigels), combs inserted and the mixture allowed to solidify at room temperature. DNA samples (20 μl) were mixed with 5 μl loading dye (0.25% bromophenol blue, 0.25% xylene cyanol, 15% Ficoll type 400) and electrophoresised at 50-100 V in TBE buffer. A λ BstEII or 1 kb ladder (New England Biolabs) was used as a marker. DNA was visualised under long-wavelength UV illumination using a BioRad gel documentation system. Gels requiring a higher degree of sensitivity were stained with SYBR green (Cambridge Bioscience). SYBR green stained gels were visualised by fluorescence (FujiFilm Fluorescent Image Analyzer FLA-3000, excitation wavelength 473 nm, visualisation wavelength 532 nm).

ii. SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

12.5-15% polyacrylamide gels were made with 29:1 acrylamide-bisacrylamide (Sigma) and electrophoresed using standard apparatus and protocols (BioRad Mini-PROTEAN III) in Tris-glycine running buffer (25 mM Tris-HCl, 250 mM glycine, 0.01% SDS). Protein samples were mixed with 0.5 volumes of SDS loading dye (100 mM Tris-HCl pH 6.8, 200 mM dithiothreitol, 4% SDS, 0.2% bromophenol blue, 20% glycerol). Samples were incubated at 100°C for 5 min immediately before loading. Gels were run at 190 V and stained with Coomassie blue (200 ml methanol,

200 ml H_2O, 80 ml acetic acid, 0.48 g Coomassie Blue R-250) at room temperature for 15-20 min. Gels were destained in 20% methanol, 10% acetic acid for 1-2 hours and stored in distilled water.

iii. Native PAGE

^{32}P-abelled oligonucleotides and annealed substrates were purified on 10% polyacrylamide in TBE buffer using the BioRad Protean II system. Samples mixed with loading dye were applied to the gel and electrophoresed at 60 V overnight with buffer recirculation. The same protocol was used in DNA cleavage assays except gels were electrophoresed at 190 V for 2 h.

DNA gel retardation assays employed low-ionic strength 4% polyacrylamide gels. Gels were electrophoresed in LIS buffer (6.7 mM Tris-HCl, 3.3 mM sodium acetate, 2 mM EDTA) at 160 V for 2 h. Loading dye was omitted from samples applied to the gel, although a sample of dye was included in the first lane as a marker.

2.3.9 Protein overexpression

Target protein overexpression was from pT7-7, pET or pMALc2 vectors in BL21 (DE3) pLysS, BL21 (DE3) Codonplus or BL21-SI backgrounds (see Chapters 4 and 5). *E. coli* strains were transformed with the relevant expression vector and transformants grown in overnight cultures with selection. These were used to inoculate fresh medium (LB or LBON) containing the appropriate antibiotics. At A_{650nm} ~0.5, overexpression was induced by addition of IPTG or NaCl as described

(Chapters 4 and 5), followed by a further 3 hours incubation with vigorous aeration. Cells were harvested by centrifugation at 6000 rpm at 4°C for 10 min.

2.3.10 Protein purification

Full details of purification protocols used in this study are described in the relevant chapters.

i. Purification buffers

All buffers were derived from those described below; salt, glycerol, imidazole or maltose concentrations were adjusted as detailed in the relevant sections. PMSF (phenylmethylsulphonyl fluoride; Sigma-Aldrich) a serine protease inhibitor, was prepared as per the manufacturers instructions and used as detailed in Chapter 4 (section 4.2.3.1) to inhibt proteolysis of wild-type Orf.

Buffer A: 20 mM Tris-HCl pH 8.0, 1 mM EDTA, 0.5 mM dithiothreitol (DTT), 10% glycerol.

Lysis buffer (Equilibration buffer): 50 mM sodium phosphate (NaH_2PO_4) pH 8.0, 300 mM NaCl, 10 mM imidazole.

Elute buffer: 50 mM sodium phosphate pH 8.0, 300 mM NaCl, 250 mM imidazole.

Column buffer: 20 mM Tris-HCl pH8.0, 200 mM NaCl, 1 mM EDTA.

ii. Sonication

Cells were disrupted by sonication using an MSE Soniprep 150 at an ampitude of 7-7.5 in 20 second bursts. Samples were kept in a water-ice mixture throughout sonication to help prevent localised heating.

iii. Buffer exchange

Buffer exchange was performed by dialysis with the required buffer for ≥ 3 hours at 4°C with mixing on a magnetic stirrer. Dialysis tubing (17.5 mm diameter; Medicell International Ltd) was prepared as per the manufacturers instructions and stored in SDW at 4°C.

iv. Ion exchange and affinity chromatography

Ion exchange and affinity chromatography was performed using BioRad Econocolumns. Matrices were prepared and poured following the manufacturers instructions using the buffers listed in section 2.3.10(i). All columns were run at 4°C, column flow-through, wash and elute fractions were collected and kept at 4°C. In some instances proteins were incubated with matrices prior to pouring of the column as described in the relevant sections of Chapters 4 and 5.

Q-sepharose fast flow medium (Amersham-Pharmacia), was prepared by decanting the required volume of slurry into the column, and equilibrating with 10 column bed volumes of Buffer A containing appropriate concentrations of salt (see Chapters 4 and 5). The supernatant following cell lysis was applied to the column in the same buffer and proteins allowed to bind. Bound proteins were washed with 10 column volumes of the same buffer to remove unbound proteins. Proteins were eluted from the column on a stepwise salt gradient.

An appropriate amount of ssDNA cellulose powder (Sigma-Aldrich), was prepared by adding Buffer A and allowing absorption for 20 minutes. Removal of fines was achieved by mixing of the solution, followed by a settling period, and then removal of half of the buffer. This was repeated, with the addition of fresh buffer

each time, until the buffer no longer appeared cloudy. This ensured removal of fines from the slurry, which can inhibit buffer flow through the column. The required amount of matrix was then poured in to a column and equilibrated with 10 bed volumes of the appropriate buffer (Buffer A + KCl), before addition of the sample. Bound proteins were washed with 10 column volumes of the same buffer, and then eluted on a step-wise salt gradient.

Heparin agarose slurry (Sigma-Aldrich) was poured in to a column and equilibrated with 10 column volumes of appropriate buffer (Buffer A + KCl). Protein solutions, in the same buffer, were loaded on to the column and then washed with 10 column volumes of the equilibration buffer. Bound proteins were eluted on a salt gradient.

His-Select™ Nickel Affinity Gel (Sigma-Aldrich) was poured in to a column and washed with 1 to 2 column volumes of deionised water to remove the storage ethanol. The matrix was equilibrated with 5 volumes of Lysis buffer (equilibration buffer), before addition of the recombinant protein soluion. The column was washed with 10 column volumes of the same buffer and proteins eluted with 5 column volumes of Elute buffer.

Amylose resin was prepared by equilibrating with 10 column volumes of Column buffer. Protein solutions were added to the matrix and washed with 12 column volumes of the same buffer. Bound proteins were eluted with Column buffer + 10 mM maltose.

v. Gel filtration

Size exclusion chromatography of His-Orf was performed using a BioLogic DuoFlow system (BioRad). 0.5 mg/ml of protein was injected onto a S-200 HR sephacryl (1.6 x 60 cm column bed) column in Buffer A containing 250 mM KCl. Molecular weight standards (BioRad) were run under the same conditions and contained thyroglobulin (670 kDa), gamma globulin (158 kDa), ovalbumin (44 kDa), myoglobulin (17 kDa) and vitamin B12 (1.35 kDa). All other gel filtration experiments were performed using an AKTA FPLC system with a 24 ml Superose 6HR 10/30 column (GE Healthcare). Proteins (1 mg/ml unless otherwise stated) were mixed in a 200 µl volume with 250 mM KCl Buffer A and incubated on ice for 15 min prior to loading 100 µl onto the column at a flow rate of 0.3 ml/min. A 51-mer oligonucleotide at 0.1 mg/ml was added last to some protein mixtures. The same molecular weight standards were used and all column runs were performed at 4°C. Protein and DNA elution was detected by UV absorption at 260 nm and 280 nm. Peak fractions were visualised by SDS-PAGE for proteins and on 3% (w/v) agarose gels with SYBR green staining for DNA.

vi. Protein concentration measurements

Protein concentration was estimated using the BioRad Protein Assay kit, which is a modified version of the Bradford Assay. Bovine serum albumin (BSA) was used as a standard. Samples were placed in a 1 ml cuvette and absorbance measured at 595 nm in a Cecil CE3041 spectrophotometer.

2.3.11 5'-end labelling of DNA substrates and DNA substrate assembly

i. ^{32}P-labelling

ssDNA oligonucleotides NH7 and FAC6 were 5'-end labelled using T4 polynucleotide kinase (Invitrogen) and γ-^{32}P ATP (Amersham Pharmacia). Typical reactions contained 3 µl of the desired substrate (100 ng/µl), 4 µl 5x Forward reaction buffer (350 mM Tris-HCl pH 7.6, 50 mM $MgCl_2$, 500 mM KCl, 5 mM 2-mercaptoethanol), 1 µl T4 polynucleotide kinase, 2 µl γ-^{32}P ATP and 10 µl SDW. Reactions were incubated at 37°C for 1 h followed by 15 min at 65°C to inactivate the kinase. 30 µl SDW was added to the mix and labelled DNA separated from unincorporated ^{32}P ATP using Micro Bio-spin columns (BioRad). The sample was made up to 50 µl with SDW and 90% recovery of the oligonucleotide was assumed at this stage.

ii. Annealing

Labelled oligonucleotides were annealed with their complementary partners (3-fold excess of unlabelled substrate) in SCC buffer (150 mM NaCl, 15 mM sodium citrate pH 7.0). Samples were incubated at 92°C for 2 min and allowed to cool gradually in a heat block (Techne Dri-Block® DB-2D) for ~2 h. Loading dye (5 µl) was added and products separated on native polyacrylamide gels (10% polyacrylamide in TBE buffer). Gels were wrapped in Saranwrap (cling-film), corners marked with luminescent marker and exposed to X-ray film for 1-2 min. Bands corresponding to the correct products were excised using an X-ray film template and eluted by soaking in TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA) overnight at 4°C. Concentration of the DNA substrate was estimated by relating the radioactivity

recovered (measured in a Scintillation counter) to that of the labelled oligo, the concentration of which was known.

2.3.12 Western blot analysis

Proteins were separated on 15% SDS-PAGE as described (2.3.8ii) using pre-stained molecular weight markers (BioRad). Gels were soaked in methanol-transfer buffer (20% methanol, 10 mM cyclohexylaminopropanesulfonic acid (CAPS)) before assembly in the transblot apparatus (BioRad). Proteins were transferred onto a PVDF membrane (BioRad) at 50-100 mA overnight and incubated for 90 min in blocking buffer (50 mg/ml milk powder in Tris-Tween buffer (10 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1% Tween)). Blots were incubated with a 1:1000 dilution of monoclonal anti-His antibody in fresh blocking buffer for 90 min. The membranes were rinsed in Tris-Tween buffer and incubated with a 1:10 000 dilution of Goat Anti-Rabbit HRP conjugate (BioRad) in fresh blocking buffer for 90 min. Membranes were washed in Tris-Tween buffer for 1 h, changing the buffer every 10 min. Blots were visualised using enhanced chemiluminescence (ECL) reagents (Amersham Pharmacia) for 1 min followed by exposure to X-ray film for 5 min.

2.4 Biochemical Assays

2.4.1 Gluteraldehyde cross-linking

Protein oligomerisation was investigated by cross-linking with gluteraldehyde (Sigma). Protein samples were dialysed against 1L 100 mM $NaHCO_3$ for ≥ 3 hours at 4°C. Protein samples (0.1-0.5 mg in 38 µl of this buffer) were mixed with 2 µl of 0.00025%, 0.0025%, 0.025% and 0.25% gluteraldehyde and incubated at room

temperature for 2 min. 10 µl of SDS-loading dye was added and reactions boiled for 2 min before separation on 12.5% SDS-PAGE. Protein species were visualised by staining with Coomassie blue or by Western blotting (section 2.3.12).

2.4.2 Gel retardation assays

Band shift assays to determine protein binding to DNA were performed by mixing ^{32}P-labelled DNA substrates with increasing concentrations of protein in binding buffer (50 mM Tris-HCl pH 8.0, 5 mM EDTA, 1 mM DTT, 5% v/v glycerol, 100 µg/ml BSA). Protein was added last in 20 µl reactions and binding mixtures incubated on ice for 15 min. Samples were separated on 4% polyacrylamide gels in LIS buffer for 90 min at 190 V. Gels were transferred to 3 mm Whatman filter paper and dried under vacuum in a gel dryer (BioRad). Dried gels were exposed to X-ray film overnight at -80°C in autoradiography cassettes. Alternatively, gels were analysed by phosphorimaging on a GE Healthcare Typhoon 9200 gel imager.

2.4.3 DNA annealing assay

Polyacrylamide gels (10%) in TBE buffer were used to monitor His-Orf stimulated annealing of a ^{32}P-labelled oligonucleotide with unlabelled DNA. The labelled ϕX174 oligonucleotides (25 nt or 50 nt at 0.1 ng/µl) were mixed with protein in buffer (50 mM Tris-HCl pH8.0, 1 mM DTT, 100 µg/ml BSA) before addition of ϕX174 virion DNA (20 ng/µl) to the 20 µl reaction. Samples were incubated at 37°C for 10 min and 5µl of stop buffer (20 mM Tris-HCl pH8.0, 0.5% SDS, 20 mM EDTA, 2 mg/ml proteinase K) added and incubation continued for a further 10 min. Samples were separated at 190 V for 1 h and gels dried and analysed as described

(2.4.2). Appropriate controls were separated alongside the reactions containing protein for reference. These included: i) the labelled oligonucleotide alone, ii) a pre-annealed duplex, and iii) the labelled oligonucleotide and homologous DNA without prior annealing.

2.4.4 DNA cleavage assay

Cleavage assays (20 µl) were performed by mixing ^{32}P-labelled DNA with increasing concentrations of protein in cleavage buffer (50 mM Tris-HCl pH 8.0, 1 mM DTT, 100 µg/ml BSA) containing 1 mM $MgCl_2$. Reactions were incubated at 37°C for a 30-60 min prior to addition of cleavage stop buffer (20 mM Tris-HCl pH 8.0, 0.5% v/v SDS, 20 mM EDTA, 2 mg/ml proteinase K) and a further 10 min incubation at 37°C. Controls without protein were subjected to the same reaction conditions and treatment. Loading dye was added to samples and DNA separated on 10% polyacrylamide gels in TBE buffer at 190 V for 2 h. Gels were dried and analysed as described above.

2.4.5 Affinity chromatography

To analyse the interaction of MBP-Orf or MBP-Orf151 with SSB or SSBΔC10, purified MBP-Orf (0.5 mg/ml) or MBP-Orf151 (0.5 mg/ml) was coupled with 200 µl amylose resin (typically binds 3 mg/ml bed volume, prepared as described in section 2.3.10iv) by incubation at 4°C for 15 min with rotation. The protein-bound resin was washed, packed in a glass Econocolumn (BioRad; 5 x 50 mm), and equilibrated with 10 column volumes of column buffer (20 mM Tris-HCl pH 7.4, 200 mM KCl, 1 mM EDTA) prior to use. SSB or SSBΔC10 (0.5 mg/ml in column buffer) was applied and

the column washed to elute unbound protein. Bound proteins were eluted with column buffer containing 10 mM maltose. Fractions were analysed by SDS-PAGE and proteins visualised by Coomassie blue staining. SSB and SSBΔC10 were also applied to amylose resin under the same conditions in the absence of MBP-Orf or MBP-Orf151 as a control.

2.4.6 Far Western analysis

Far Western blotting was used to detect MBP-Orf interactions with other proteins bound to a PVDF membrane. Prey proteins, SSB, SSBΔC10, His-Exo and His-Bet were separated on 15% SDS-PAGE as described (2.3.8ii) using pre-stained molecular weight markers (BioRad). The proteins were transferred to PVDF membrane as described (2.3.12) and blocked in blocking buffer for 10 min. The membrane was incubated with bait protein (5 μg of MBP-Orf) in blocking buffer for 4 hours at 4°C with gentle agitation. After washing with fresh Tris-Tween buffer (2.3.12) the membrane was incubated with a 1:1000 dilution of monoclonal Anti-MBP antibodies (Sigma) for 30 min. The membrane was washed again then incubated with a 1:10000 dilution of rabbit anti-mouse-IgG peroxidase conjugate (Sigma) for 30 min with agitation. Final wash and detection steps were as described in the Western blotting section (2.3.12). Control membranes only exposed to the secondary labelled antibody were also treated with ECL reagents and exposed to X-ray film. A control SDS-PAGE gel was stained with Coomassie blue to check for presence of the prey proteins. In addition, the blotted polyacrylamide gel was stained to check that the prey proteins had been successful transferred to the PDVF membrane.

2.5 Collaborations

The crystal structure of Orf was determined by Dr Karen Maxwell and Prof Aled Edwards at the Ontario Cancer Institute and Department of Medical Biophysics, Toronto University, Toronto. Complementation studies using *Bacillus subtilis* recombination genes were performed in collaboration with Dr Juan Alonso at the Department of Microbial Biotechnology, Centro Nacional de Biotecnologia, Madrid.

Chapter 3

New phage Orf family sequences

3.1 Introduction

λ *ninB* lies within a chain of interlinked open reading frames in the *nin*R cluster located between *P* and *Q* genes (Kroger and Hobom, 1982). This region shows remarkable organisation with the end of the preceding gene overlapping the start codon of the next (Kroger and Hobom, 1982). The *ninB* gene encodes a protein of 146 amino acids with a predicted molecular weight of 16.6 kDa (Daniels, 1983; Kroger and Hobom, 1982). Expression of the *ninB* product was detected in these early studies but no biochemical analysis was undertaken. A decade later, Frank Stahl's group discovered that phage λ carried a functional analog of the *E. coli recO*, *recR*, and *recF* recombination genes located within the *nin*R region (Sawitzke and Stahl, 1992). They performed crosses using λ containing a series of deletions within this region to determine which open reading frame was responsible. The analog corresponded to *ninB* and was renamed *orf* for *recO*-, *recR*-, *recF*-like function (Sawitzke and Stahl, 1992). Plasmids carrying *orf* were found to be capable of replacing the functions of *E. coli* RecO, RecR and RecF proteins during λ phage recombination but not in *E. coli* (conjugative) recombination (Sawitzke and Stahl, 1994). This latter observation has been challenged by recent evidence suggesting that Orf can substitute for *recFOR* mutants in *E. coli* recombination and DNA repair when Exo and β are present (Poteete, 2004); this feature will be discussed more fully later in this work. *E. coli recO*, *recR*, and *recF* encode proteins of 243 (Morrison *et*

al., 1989), 201 (Mahdi and Lloyd, 1989; Yeung *et al.*, 1990) and 358 amino acids (Blanar *et al.*, 1984), respectively. We were eager, therefore, to explain how a single phage enzyme of only 146 amino acids could replace the function of a complex of three much larger host proteins.

Prior to the commencement of this project, little was known concerning Orf. Few homologs had been identified, the structure was unknown, no biochemical analysis had been undertaken and, although it had been shown to participate in recombination, initial studies had failed to define a specific role for Orf in λ recombination. As a first step in the characterisation of Orf, the primary sequence was investigated by employing modern sequence analysis tools to probe protein databases. It was hoped that identification of Orf homologs would help in the elucidation of its biochemical activities, identify conserved functional elements and reveal its distribution among phage populations. A flowchart outlining a systematic approach to analysing protein sequences is shown in Figure 3.1. The three-dimensional structure of a polypeptide essentially determines its activity and function in biology. Hence, knowledge of the architecture of a novel protein is extremely valuable, helping us understand possible substrate and ligand binding mechanisms and allowing us to devise rational mutagenesis approaches and appropriate biochemical assays. Methodologies used for analysing the protein sequence of Orf, as outlined in the flowchart, are described in this chapter. The crystal structure of Orf (Maxwell *et al.*, 2005) was obtained relatively early in these studies, so that the latter stages of tertiary structure prediction and comparative homology modelling proved unnecessary. The experimental data arm of the flowchart (Figure 3.1) is addressed in

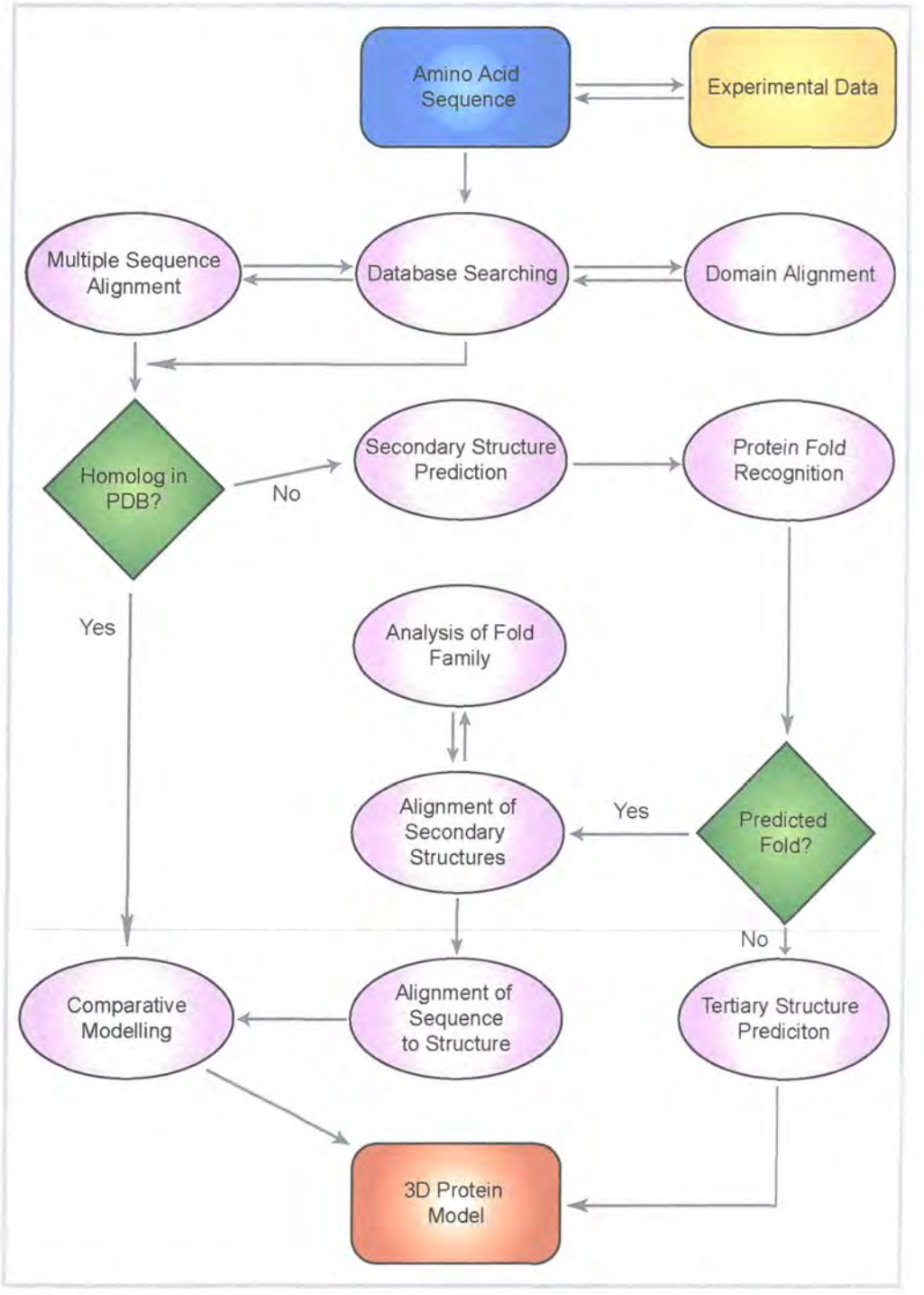

Figure 3.1. Protein sequence analysis flowchart.
This flowchart outlines a systematic approach to protein sequence analysis and a generalised *approach* to protein structure prediciton. The procedures shown below the grey line indicate those not performed during this study, but are included to show the complete set of analysis techniques required to predict protein structure. The blue box denotes the starting point of the study, the orange box incorporates experimental procedures that could be performed to further elucidate protein function, purple ellipses indicate techniques performed (or that can be performed), green diamonds indicate results returned and the red box refers to the intended end point for this procedure.

later chapters of this thesis, which describe the cloning, purification and biochemical analysis of Orf.

Using this systematic approach for analysing the Orf protein sequence we uncovered both closely related and distantly related Orf homologs residing in a diverse number of phages and discovered that the *orf* gene is located in a conserved genomic arrangement within these phages.

3.2 Orf belongs to a family of proteins conserved in diverse phages

The initial step in analysing a protein sequence involves the employment of database searches to construct pair-wise alignments of homologous sequences. Although there are many methods for searching sequences, the most well known of these is the BLAST (basic local alignment search tool) suite of programs available from NCBI (Altschul *et al.*, 1990). Based upon the Smith-Waterman algorithm (Smith and Waterman, 1981), these programs align sequences on the basis of local similarity between them. With the introduction of the Position-Specific Iterated BLAST (PSI-BLAST) and gapped-BLAST programs (Altschul *et al.*, 1997) it is now possible to detect very remote homologs with relative ease. These new programs are more sensitive to weak, but frequently biologically relevant, sequence similarities. By compiling multiple, as opposed to pair-wise alignments, distant relationships can be revealed by PSI-BLAST that would otherwise be indistinguishable from chance sequence similarity.

The Orf amino acid sequence was entered into the PSI-BLAST program using the default parameters to search for homologous sequences. The program works by taking the amino acid sequence and comparing it to protein databases using the

gapped-BLAST program. A multiple alignment and profile are constructed from the significant local alignments retrieved, using the input sequence as a template (unseen by the user). The constructed profile is then used to search the protein database again for further homologs. The program can repeat the process (iterate) a number of times or until convergence is achieved, allowing detection of distantly related sequence similarities. Using this approach we performed the first three steps (database searching, domain alignment and multiple sequence alignment) delineated in Figure 3.1.

The PSI-BLAST search uncovered a number of proteins (or putative proteins) from a variety of phage and prophage sources that display significant sequence similarity to λ Orf. The predicted Orf homologs recovered by these searches are listed in Table 3.1. It should be noted that the words homolog and homology are used with the understanding that significant sequence identity is indicative of homology (i.e. similarity attributed to common ancestry), it is not absolute proof of such. In cases of limited sequence identity, putative homology is supported by the position of *orf* relative to adjacent gene sequences (see section 3.3). Alignment of selected Orf-like protein sequences (Figure 3.2) was performed using the ClustalW alignment software (Chenna *et al.*, 2003). Secondary structure elements taken from the Orf crystal structure (Maxwell *et al.*, 2005) are displayed above the amino acid sequence.

Most *orf* homologs are encoded by prophages and phages infecting Gram-negative bacteria belonging to the β (e.g. *Neisseria* and *Bordetella* species) and γ (e.g. *E. coli* and *Salmonella* species) subdivisions of proteobacteria. Such a distribution is not too surprising given that *orf* resides within the enterobacterial phage λ genome. Lambdoid (lambda-like) phages are prevalent within pathogenic

ORF Designation	Source	Host Organism	Length	% Identity to NinB	Accession N°	Source/Reference
hypothetical protein	*E. coli* UTI89 (UPEC)	UPEC	146	99	YP_541667	(Chen et al., 2006)
gp63	phage HK97	*E. coli*	146	98	NP_037743	(Juhala et al., 2000)
gp64	*S. typhimurium* phage ES18	*Salmonella typhimurium*	145	97	YP_224202	(Casjens et al., 2005)
NinB	*S. typhimurium* phage ST104	*S. typhimurium* DT104	145	96	YP_006386	(Tanaka et al., 2004)
NinB	phage P22	*S. typhimurium*	145	95	NP_059612	direct submission, NCBI Genome project
NinB	phage HK022	*E. coli*	146	95	NP_597901	(Juhala et al., 2000)
gene *50*	phage Sf6	*Shigella flexneri*	146	94	NP_958225	(Casjens et al., 2004)
Y2223	*Yersinia pestis* KIM	*Y. pestis* KIM	148	54	NP_669532	(Deng, et al., 2002)
NinB	*Aeromonas hydrophila* cryptic prophage	*A. hydrophila* PPD134/91	146	53	AAS46709	(Yu et al., 2005)
NinB	phage ST64T	*S. typhimurium*	101	47	NP_720308	(Mmolawa et al., 2003)
NinB	*Pseudomonas putida* F1 prophage	*P. putida* F1	132	46	ZP_00897726	direct submission, NCBI Genome project
NinB	phage VT2-Sa	*E. coli* 0157:H7	181	44	NP_050530	(Miyamoto et al., 1999)
hypothetical protein	phage Nil2	*E. coli* 0157:H7 strain Nil653	181	42	CAC95094	direct submission, NCBI Genome project
p27	phage HK620	*E. coli*	136	37	NP_112060	(Clark et al., 2001)
hypothetical protein	*Haemophilus influenzae* R2866	*H. influenzae* R2866	139	34	ZP_00156617	direct submission, NCBI Genome project
hypothetical protein BB2204	*Bordetella bronchiseptica* RB50	*Bordetella bronchiseptica* RB50	135	34	NP_888748	(Parkhill et al., 2003)
NinB	Bacteriophage Aaphi23	*Actinobacillus actinomycetemcomitans*	165	32	NP_852743	(Resch et al., 2004)
hypothetical protein	*Neisseria gonorrhoeae* FA 1090	*Neisseria gonorrhoeae* FA 1090	127	27	YP_208182	direct submission, NCBI Genome project
ybcN (Orf151)	DLP12 prophage	*E. coli* K12	151	19.9	NP_415079	(Blattner et al., 1997)
Orf20	phage phi ETA	*Staphylococcus aureus*	223	13.8	NP_510914	(Yamaguchi et al., 2000)

Table 3.1. Putative homologs of λ Orf.
Homologs were identified using PSI-BLAST database searches and are listed in order of similarity to λ Orf. The name of each of the putative homologs shown refers either to how they are listed in GenBank or their putative homology. Phages or prophages encoding each gene product are shown in column 2 and their host bacterial species listed in column 3. Amino acid length is indicated in column 4. Accession numbers refer to GenBank data entries.

bacteria, including those from the β and γ proteobacteria families. A number of lambdoid genomes have been sequenced and found to share many of the same genes as λ (Hendrix *et al.*, 1999; Vander Byl and Kropinski, 2000; Juhala *et al.*, 2000; Plunkett *et al.*, 1999; Neely and Friedman, 1998; Miyamoto *et al.*, 1999). Bacteriophages are highly diverse in nature due to frequent recombination-driven genetic rearrangements (Hendrix, 2002). It is therefore quite likely that functionally important genes will be retained or transferred between phages of distant common ancestry. This theory would explain the widespread occurrence of Orf-like proteins and suggests that it fulfils an important role in phage DNA metabolism.

Many of the Orf homologs found using the PSI-BLAST program show significant similarity to λ Orf (Table 3.1). It is common practice, when analysing protein sequences using BLAST programs, to assume that hits with a sequence identity of ≥25% and E-values of $\geq 10^{-3}$ are significant alignments and not due to chance occurrence. However, these cut-offs cannot be followed blindly and those sequences with a low percentage identity recovered in PSI-BLAST searches should be considered further before being dismissed as unrelated.

In addition to the numerous closely related homologs retrieved, a significant number of distantly related relatives were found (Table 3.1). Of these, a putative phage protein (Orf151) encoded by the *E. coli* K12 cryptic prophage DLP12 and the Orf20 protein from *Staphylococcus aureus* phage ϕETA, designated ETA20, were of interest because of their relative dissimilarity (<25% identity) to λ Orf. To assess their relatedness in more detail we constructed pair-wise alignments of the two subject entries with the λ Orf sequence (Figure 3.2B and C). Alignment of DLP12 Orf151 and λ Orf revealed a number of short conserved segments across the length

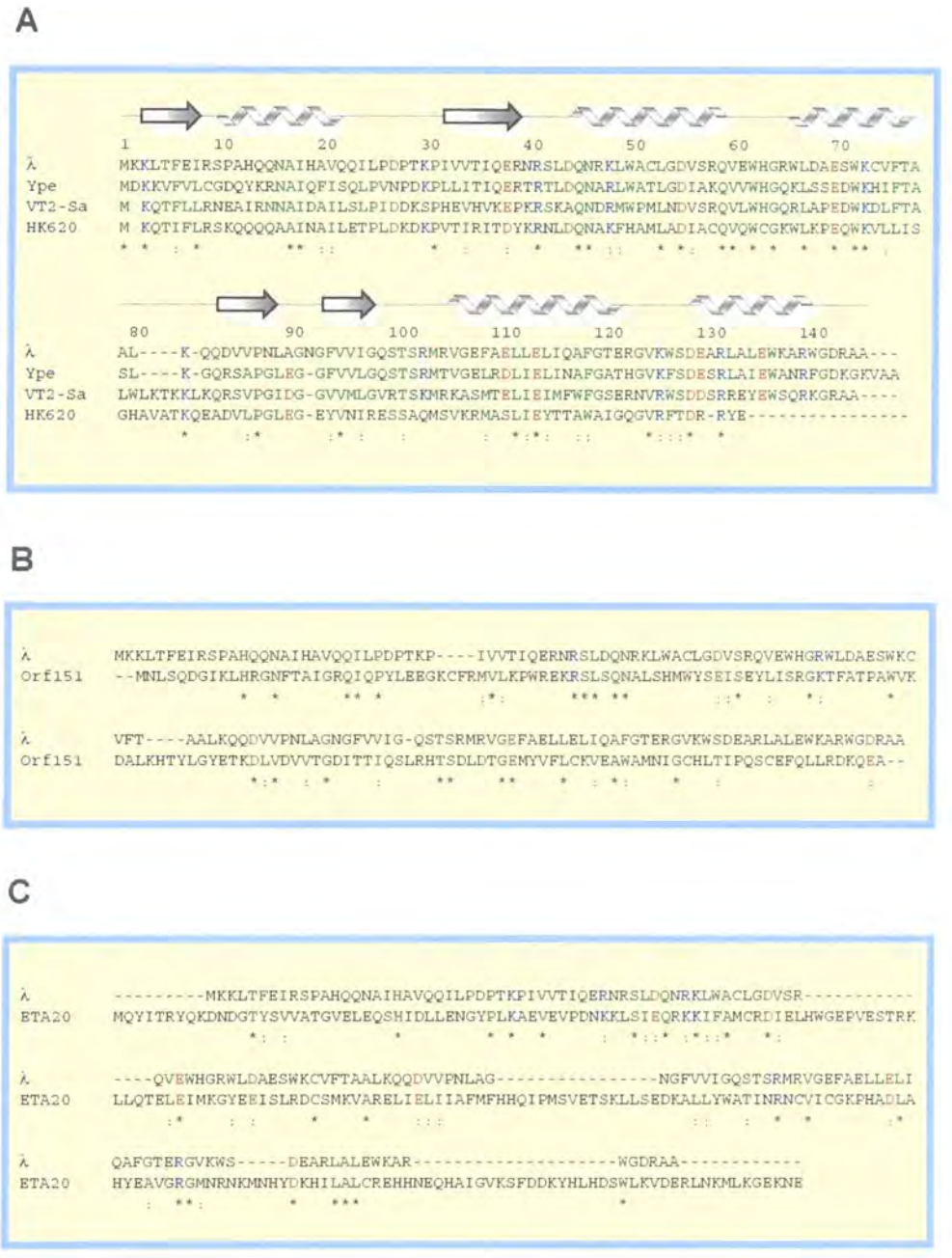

Figure 3.2. Sequence alignment of selected Orf homologs.
Conserved residues are highlighted: acidic (red), basic (blue), and others (green). Asterisks denote the presence of the same amino acid in all instances, while colons denote conserved functionally related residues. **(A)** Amino acid sequence alignment of Orf homologs with >30% sequence identity to Orf. The representative homologs come from a *Yersinia pestis* prophage (Ype), *E. coli* 0157:H7 phage VT2-Sa, and *E. coli* H phage HK620. Secondary structure elements are indicated above the residue numbers given for the λ Orf sequence. **(B)** Pair-wise sequence alignment of Orf and a homolog, Orf151, from the *E. coli* cryptic prophage DLP12. **(C)** Pair-wise sequence alignment of a *Staphylococcus aureus* phage ϕETA homolog with the λ Orf sequence.

of the proteins with only two internally positioned gaps. The relative absence of gaps within the alignment supports the idea that the proteins are ancestrally related. In contrast when ETA20 and Orf were aligned, multiple large gaps were evident and only a small number of scattered conserved residues were noted, although some of these coincide with those conserved in the Orf151-Orf alignment. The gaps within the sequence are to be expected as ETA20 is 77 residues longer than Orf. The C-terminal region of ETA20 incorporates a zinc-finger motif belonging to the HNH family nuclease domain (Mehta *et al.*, 2004). The presence of this motif suggests that ETA20 might function as a nuclease, however, experimental evidence is needed to confirm this possibility.

The use of alternative sequence alignment programmes, such as EMBOSS (European Bioinformatics Institute (EBI) alignment programs), yielded different pair-wise alignments for Orf, Orf151 and ETA20. This is not too surprising as the sequences share limited identity; however, it does cast some doubts over the validity of the proposed relationship. To help address this ambiguity we applied secondary structure prediction methodologies on each of the three protein sequences (Figure 3.3 and Figure 3.4).

A significant number of protein structures have been solved over the last 30 years. Analysis of these has revealed multiple examples where proteins lacking obvious sequence homology share very similar domain folds. Sequence alignments alone are therefore unreliable tools and this has encouraged the use of secondary structure and protein fold comparisons to verify homologous relationships (Sandhya *et al.*, 2003; Geourjon *et al.*, 2001; Errami *et al.*, 2003). Experimentally determined and predicted secondary structures of proteins have been used to identify related

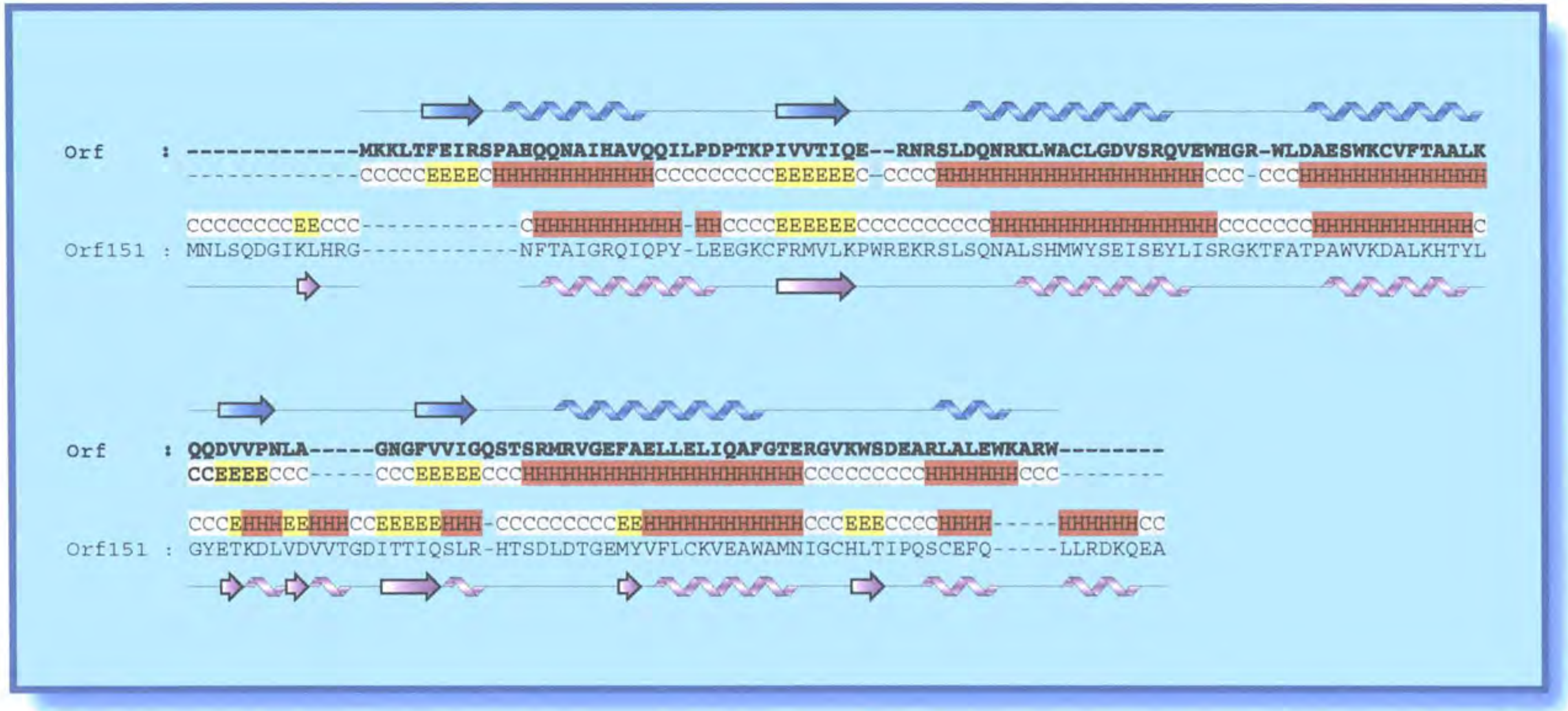

Figure 3.3. Alignment of Orf secondary structure with the predicted secondary structure of DLP12 Orf151.
Alignment and secondary structure prediction were performed using the Phyre protein homology recognition engine. Known and predicted secondary structure elements are indicated above and below each amino acid sequence in text and 2D form. Elements are highlighted: coil (grey C), β-strand (yellow E), and α-helix (red H). 2D secondary structure elements are coloured blue (Orf) and purple (Orf151)

protein families possessing sequence identity in the range of 10%-30%. It has also established that >50% secondary structure similarity between a pair of proteins is generally compatible with functional relatedness (Errami *et al.*, 2003; Geourjon *et al.*, 2001).

To help in our structural comparisons, we therefore submitted the Orf151 and ETA20 sequences into the Phyre (Protein homology/analogy recognition engine) protein fold recognition server, a recent replacement for 3D-PSSM (Kelley *et al.*, 2000). Phyre allows the identification of structurally related homologs based on known or predicted secondary structure. The query sequence (either Orf151 or ETA20 in this instance) is used initially to search protein databases (PSI-BLAST) for homologs. The multiple sequence alignments recovered are transferred to the SCOP (structural classification of proteins) database to form a multiple structural alignment of homologs retrieved from this database. Finally, secondary structure predictions are made by PSI-Pred. This last program helps identify remote homologs from a database of known protein structures based on the profiles obtained from the PSI-BLAST and SCOP searches.

Phyre analysis of Orf151 revealed that it shares significant structural relatedness to λ Orf with a precision of 95% (the estimated confidence of the relationship between the two predicted structures). It is evident from these secondary structure alignments that Orf151 shares many of the structural features present in Orf (Figure 3.3). Interestingly, the secondary structure predicted for λ Orf matches remarkably well that obtained from the crystal structure (Maxwell *et al.*, 2005)(Figure 3.2). Similarly, ETA20 showed a structural homology to Orf with a precision of 75% as depicted in Figure 3.4. This result was a little more surprising

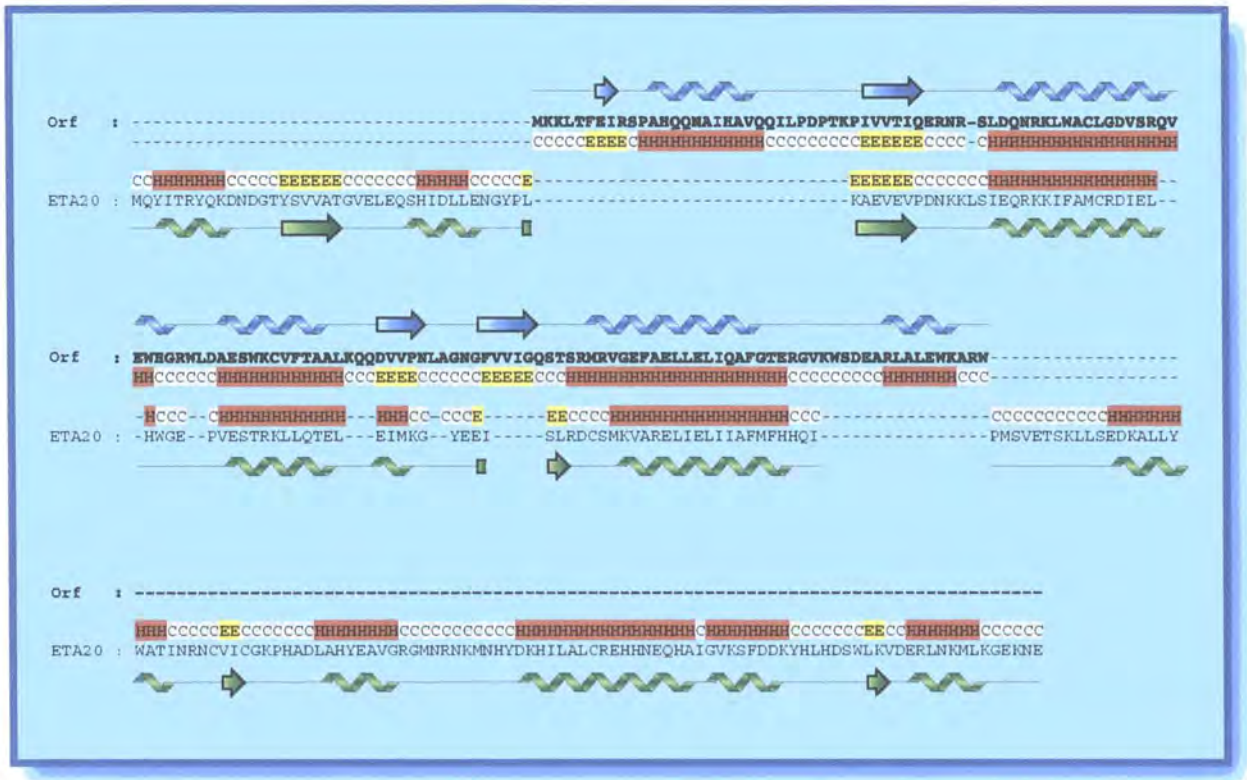

Figure 3.4. Alignment of Orf secondary structure with the predicted secondary structure of ϕETA ETA20.
Alignment and secondary structure prediction were performed using the Phyre protein homology recognition engine. Known and predicted secondary structure elements are indicated above and below each amino acid sequences in text and 2D form. Elements are highlighted: coil (grey C), β-strand (yellow E), and α-helix (red H). 2D secondary structure elements are coloured blue (Orf) and green (ETA20).

given the limited sequence identity between ETA20 and other members of the Orf family. The presence of several gaps in the ETA20-Orf alignment may indicate that the folds of these proteins do differ, with potentially important consequences for function. It is not uncommon for high confidence values (% precision) to be recovered from Phyre analysis, even when low percentage identities are noted between template and query sequences (in fact this is the point/power of fold recognition). In benchmark tests for example, 75% of query-template matches generated with an E-value of 0.94 were found to be correct, in terms of being homologous (Lawrence Kelley, personal communication). Hence, while the Phyre analysis cannot be used reliably to assert that Orf151 and ETA20 are genuine structural and functional homologs of Orf, it does confirm that the initial primary sequence alignments were valid in spite of the limited sequence identity. Moreover, both Orf151 and ETA20 can be considered as distantly related homologs of Orf, at least until experimental evidence proves otherwise.

3.3 Conserved genomic location of Orf proteins

The genetic context of *orf*-like sequences was examined, using the BLAST suite of programs together with information from GenBank, to assess whether gene synteny was maintained in phages containing homologs of Orf. A basic alignment for the relevant regions from each phage genome was assembled to show the gene organisation surrounding Orf-like sequences (Figure 3.5). *S. aureus* phages were omitted from this alignment as the majority of genes from these phages have yet to be assigned a function making inclusion of the data uninformative. As mentioned previously, Orf is located within the *ninR* region of phage λ, downstream of the

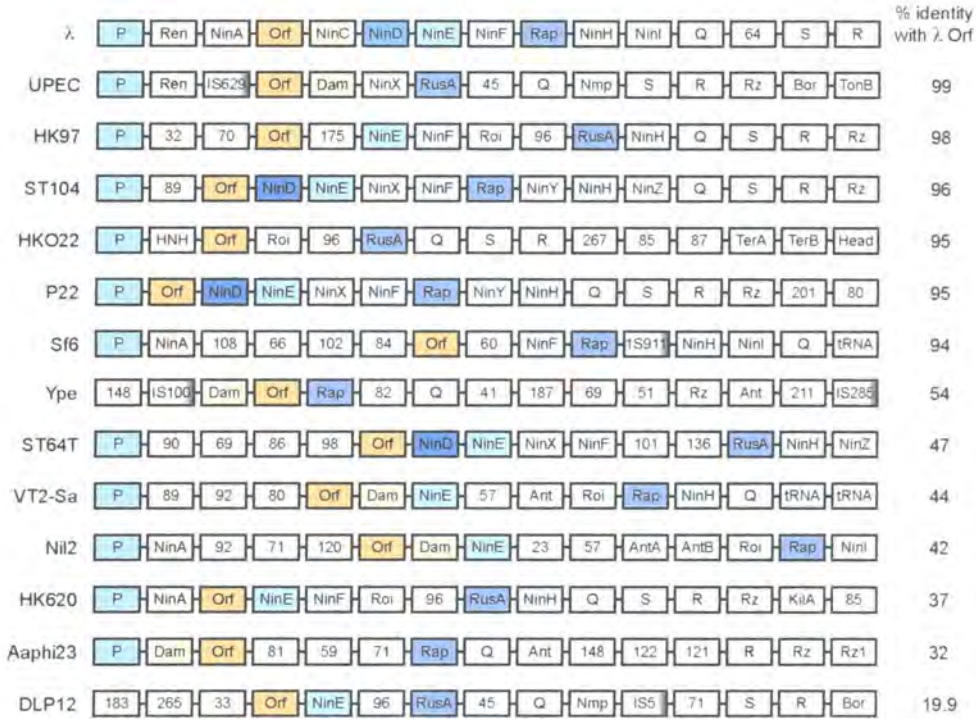

Figure 3.5. Conservation of λ ninR region genome organisation.
The gene arrangement in the regions surrounding the *orf* gene from various lambdoid phages and prophages was examined. Protein coding regions are indicated by boxes and coloured based on their position within the λ genome (matches to λ genes are coloured accordingly). Putative genes with no identified homologs from this section of λ are shown as white boxes and labelled according to their length in amino acids. Shaded grey boxes indicate transposons or insertion sequences. The representative genome regions come from *E. coli* UTI189, *E. coli* phage HK97, *S. typhimurium* phage ST104, *E. coli* phage HK022, *S. typhimurium* phage P22, *S. flexneri* phage Sf6, *Yersinia pestis* prophage Ype, *S. typhimurium* phage ST64T, *E. coli* 0157:H7 phage VT2-Sa, *E. coli* strain Nil653 phage Nil2, *E. coli* phage HK620, *A. actinomycetescomitans* phage Aaphi23 and *E. coli* K12 cryptic prophage DLP12.

replication protein P, and upstream of the Q antiterminator and S/R cell lysis proteins. A similar arrangement is found among most of the other phage genomes analysed (Figure 3.5). In addition, a number of open reading frames matching the λ *nin* genes, notably NinD, NinE, NinF and NinG (Kroger and Hobom, 1982), are frequently preserved in the intervening region and often in a similar order in each phage genome. Of these putative genes, NinG (Rap) is known to participate in phage recombination by supplying a branch-specific endonuclease with the capacity to resolve Holliday junctions and other DNA branched structures (Sharples et al., 2004; Sharples et al., 1998; Sharples et al., 1999a). The relative position of Rap mirrors that of the structurally-unrelated RusA Holliday junction resolvase (Mahdi *et al.*, 1996). In every case, either Rap or RusA is present at a similar place downstream of Orf (Figure 3.5). Significantly, Orf151 is found three genes upstream of RusA in DLP12 with a homolog of NinE located in the intervening region. The evidence provided by sequence similarity, coupled with the conserved genomic arrangements highlighted here, therefore supports the notion that DLP12 Orf151 and λ Orf do share common ancestry and are likely to play similar roles in phage recombination.

3.4 Discussion

This chapter summarises efforts to identify proteins homologous to λ Orf utilising currently available database search methods. Multiple homologs of Orf have been identified and all of these are associated with phage sequences. Those matching Orf most closely reside in lambdoid phages and are almost certainly genuine orthologs; the differences in amino acid sequence arising by evolutionary divergence from a common ancestor. For potential Orf homologs displaying a more distant relationship,

we employed secondary structure predictions and evaluated the results alongside data obtained from the crystal structure of λ Orf. *E. coli* cryptic prophage DLP12 Orf151 and *S. aureus* phage ϕETA Orf20 (ETA20) secondary structures were shown to share significant similarities to Orf. The genomic organisation of phages containing Orf homologs was also investigated and provided additional support for the similarities identified by sequence homology. Orf was reliably located between replication and lysis functions and upstream of alternative Holliday junction resolvases supplied by either Rap or RusA.

The conservation of gene sequence and gene arrangement among this group of phages suggests that Orf fulfils an important role in phage biology. We therefore decided to purify λ Orf and begin an analysis of its biochemical properties. To confirm the relationship between Orf and the distantly-related proteins identified here, we initiated a similar study of Orf151 and ETA20. *E. coli* DLP12 Orf151 shares only minimal sequence homology with Orf but does match well in its predicted secondary structure. ETA20 shows even less similarity with typical members of the Orf family. However, its fusion to an HNH nuclease domain and the fact that it comes from a Gram-positive phage makes it an interesting candidate for further study.

Chapter 4

Purification and quaternary structure of Orf proteins

4.1 Introduction

λ Orf protein appears to influence the initial stages of phage genetic recombination. In addition to substituting for the *E. coli* RecFOR complex in λ Red-mediated crosses, it can also promote bacterial recombination in cells expressing the Red system (Sawitzke and Stahl, 1992; Sawitzke and Stahl, 1994; Poteete, 2004). To gain a better understanding of the role of Orf in genetic recombination, we decided to purify the protein and examine its biochemical properties *in vitro*. The sequence analysis, described in Chapter 3, identified several distantly related Orf-like proteins among a diverse group of lambdoid phages. From this study, two putative Orf homologs, *E. coli* DLP12 Orf (Orf151) and *S. aureus* phage ϕETA Orf (ETA20) were selected for examination alongside the λ protein. Thus the activities of three diverse representatives of the Orf family could be compared and their functional relationship confirmed experimentally. In this chapter we describe the cloning, overexpression and purification of these three Orf proteins; the quaternary structure of each was also investigated.

4.2 Wild-type λ Orf

4.2.1 Cloning of wild-type *orf*

The λ *orf* gene was amplified by PCR from λ genomic DNA and inserted into the pT7-7 expression vector to give pGS903 (supplied by GJ Sharples). This placed *orf* immediately downstream of the T7 RNA polymerase promoter of phage T7 gene *10*, allowing selective high-level transcription of the wild-type *orf* gene. The integrity of the insert was verified by DNA sequencing, and found to be identical to that of the published genome.

4.2.2 Overexpression of wild-type Orf

E. coli BL21 (DE3) pLysS competent cells were transformed with pGS903 as described in the Chapter 2 (section 2.3.4). This strain is a T7 lysogen and expresses T7 RNA polymerase under *lac* control. The pLysS plasmid (Cm^r) carries T7 lysozyme that helps eliminate any T7 RNA polymerase that might be expressed in uninduced conditions. The improved regulation of target gene transcription was considered worthwhile as previous reports suggested that Orf expression from multicopy plasmids is deleterious to *E. coli* cells (Sawitzke and Stahl, 1994).

Transformants were grown in LB broth (2 litres) containing appropriate antibiotics and *orf* expression induced at A_{650nm} ~0.5 by addition of 1 mM IPTG. The cells were incubated for a further 3 hours with vigorous shaking. Increased aeration results in high oxygen transport, which can reduce the formation of inclusion bodies, therefore increasing protein yield in the lysate (Sambrook, 2001). We found however that the cells were less viable with too much aeration and the optimum for cells over expressing Orf protein was between 100 and 150 rpm in an orbital shaker. Cells were

harvested by centrifugation and overexpression of the ~16 kDa wt-Orf protein observed by SDS-PAGE analysis (Figure 4.1A, lanes 2 and 3).

4.2.3 Purification of wild-type Orf

4.2.3.1 Lysis

Cells, containing overexpressed λ Orf protein, were resuspended in Buffer A containing 1 M KCl and lysed by sonication (Chapter 2, section 2.3.10ii). A high salt concentration was used in order to disrupt potential protein:DNA interactions and increase protein solubility. PMSF was added to a final volume of 0.5 mM to inhibit proteolysis. Inclusion of salt and PMSF in the lysis buffer improved Orf solubility (Figure 4.1A, lanes 4 and 5). Previously, we found that almost all of the Orf protein remained in the pellet following sonication (data not shown). Growth conditions were also experimented with, as reduced induction temperature has been shown to improve protein solubility (Bishai *et al.*, 1987; Sambrook, 2001). However, little or no improvement in protein solubility was detected using reduced growth temperatures (30°C or 25°C) during induction. Further modifications including reducing IPTG concentration or length of induction and increasing cell density before addition of inducer, failed to increase Orf solubility. This is consistent with Orf being a relatively insoluble protein rather than excretion into inclusion bodies. Urea extraction (solubilization) of the sonication pellet could have been attempted, although small quantities of Orf protein could be recovered from the supernatant following lysis. Further studies revealed that Orf protein tends to precipitate below a salt concentration of 250 mM KCl. Subsequent purification steps therefore maintained the salt concentration above this level.

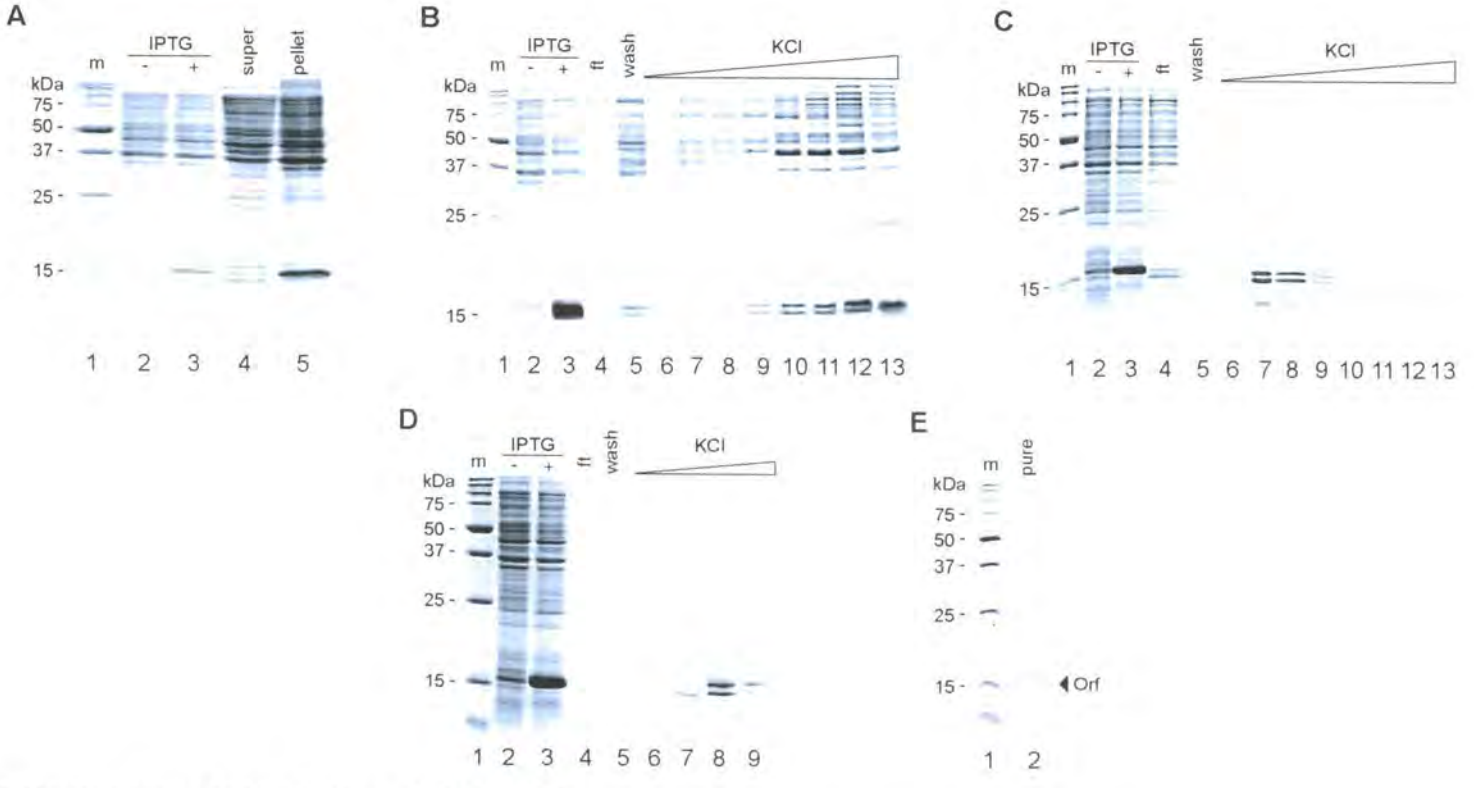

Figure 4.1. Overexpression and purification of wild-type Orf.
(A) Overexpression of Orf. Lane 1, molecular weight (MW) markers; Lane 2, uninduced cells; Lane 3, IPTG-induced cells; Lane 4, supernatant following cell lysis; Lane 5, pellet following cell lysis. **(B)** Purification of Orf by Q-sepharose ion affinity chromatography. Lane 1, MW markers; Lane 2 and 3, uninduced and induced cells; Lane 4, flow-through from column; Lane 5, wash fraction; Lane 6-13, consecutive fractions from the column, eluted on a 0.3 M-1.0 M KCl gradient. **(C)** Purification of Orf by ssDNA cellulose chromatography. Lanes as described for B. **(D)** Purification of Orf by heparin agarose chromatography. Lanes 1-5 as in part B; Lanes 6-9 consecutive fractions from the column, eluted on a 0.3 M-0.6 M KCl gradient. **(E)** Purified Orf at 0.6 mg/ml. Lane 1, MW markers; Lane 2, purified Orf. Proteins were visualised on 15% SDS-PAGE stained with Coomassie blue. The double Orf bands are thought to be a gel artifact as both bands migrate at the same position in gel filtration chromatography (see Figure 4.3A).

4.2.3.2 Q-sepharose ion exchange chromatography

A Q-sepharose Fast Flow column (10 ml) was equilibrated with Buffer A (pH 7.5) containing 250 mM KCl. This pH was found to be optimal for Orf binding and it eluted in a broad peak at 0.4-1.0 M KCl (Figure 4.1B, lanes 7-13). As the Q-sepharose column was used mainly to help reduce the viscosity of the supernatant, as well as removing several contaminants, all of these fractions were pooled. The Orf protein at this and subsequent stages appeared as a double band (Figure 4.1B). This appears to be a gel artefact, as size-exclusion chromatography (Figure 4.3A) failed to separate these species. Such gel artefacts are not unusual and may suggest the presence of intramolecular disulphide bonds that are not easily disrupted by boiling in the presence of reducing and denaturing agents.

4.2.3.3 ssDNA cellulose affinity chromatography

Because Orf is involved in the early stages of recombination, we anticipated that it might associate with ssDNA. We therefore employed a 3 ml ssDNA cellulose column, prepared to ensure removal of fine particles, pre-equilibrated in 250 mM KCl Buffer A. The pooled Orf fraction from the Q-sepharose column was applied to the column and bound proteins eluted with a salt (0.25–1.0 M KCl) gradient collecting 1 ml fractions. Most of the Orf protein bound this column, eluting at 0.4-0.6 M KCl (Figure 4.1C, lanes 7-9), and these fractions were dialysed against 2 x 1L 250 mM KCl Buffer A for 3 hours. Since Orf bound to ssDNA cellulose, and was sufficiently pure at this stage, we analysed its ability to bind ^{32}P-labelled ssDNA in a gel retardation assay (see Chapter 6). The protein did indeed associate with ssDNA in this assay, however, we could not exclude the possibility that a contaminant was

responsible for the gel shift. An additional chromatography step was therefore added to further purify the wt-Orf protein.

4.2.3.4 Heparin chromatography

2 ml of heparin agarose was equilibrated in 250 mM KCl Buffer A and the dialysed fractions from ssDNA cellulose loaded onto the column in the same buffer. Orf eluted from the column at 0.4-0.6 M KCl (Figure 4.1D, lanes 7-9). At this stage the protein appeared pure, no contaminants were visualised on 15% SDS-PAGE gels (Figure 4.1E), therefore the wt-Orf protein was deemed suitable for further biochemical analysis. Pooled wt-Orf fractions were dialysed against 500 mM KCl Buffer A containing 50% glycerol and stored in aliquots at -80°C. A total of 1.2 mg of purified wt-Orf was recovered at 0.6 mg/ml.

4.3 Histidine-tagged λ Orf (His-Orf)

Purification of the wild-type Orf protein was significantly hindered by insolubility problems, resulting in an extended purification protocol and reduced protein yield. The use of affinity tag fusions has become increasingly popular due the ease of protein isolation, purification and detection. One commonly used affinity tag utilises a tandem arrangement of six histidine residues (His_6), which binds tightly to nickel ions. The small size of the tag helps limit potential interference of protein function, or interactions with partner proteins. We decided to construct a His_6-tagged Orf fusion protein in an attempt to recover larger quantities of protein for biochemical studies. Use of a GST fusion was discounted, as its tendency to form dimers in solution would complicate an investigation of the quaternary structure of Orf.

4.3.1 Construction and overexpression of a His-Orf fusion

The pGS903 plasmid construct carrying the λ *orf* gene was digested with NdeI and BamHI and the released 438 bp insert introduced to pET14b cut with the same enzymes. The resulting clone (pPR100) generates an N-terminal His_6-Orf fusion containing a thrombin cleavage site between Orf and the tag.

BL21 (DE3) pLysS carrying pPR100 was grown in 2L of LB broth containing chloramphenicol and ampicillin and expression induced as described for wt-Orf. Similar growth problems were encountered as before and optimal induction conditions matched those employed with wt-Orf. Cells were harvested by centrifugation as before and samples of uninduced and induced cultures analysed by SDS-PAGE. Significant overexpression of the His-Orf fusion protein was evident in IPTG-induced cells (Figure 4.2A, lanes 2 and 3).

4.3.2 Purification of His-Orf

4.3.2.1 Lysis

The cell pellet was resuspended in 15 ml lysis buffer (50 mM sodium phosphate (NaH_2PO_4) pH 8.0, 300 mM NaCl, 10 mM imidazole). The lysis buffer contains 10 mM imidazole to minimise binding of untagged, contaminating proteins and to increase purity with fewer wash steps. Cells were lysed by sonication and experimentation with the salt concentration of the lysis buffer found that 300 mM NaCl was sufficient to limit protein:DNA interactions and maximise recovery of soluble His-Orf. Cell debris was removed by centrifugation and supernatant and pellet analysed by 15% SDS-PAGE (Figure 4.2A). As with wt-Orf, the majority of the His-Orf protein precipitated with the pellet (Figure 4.2A, lane 5). However, more

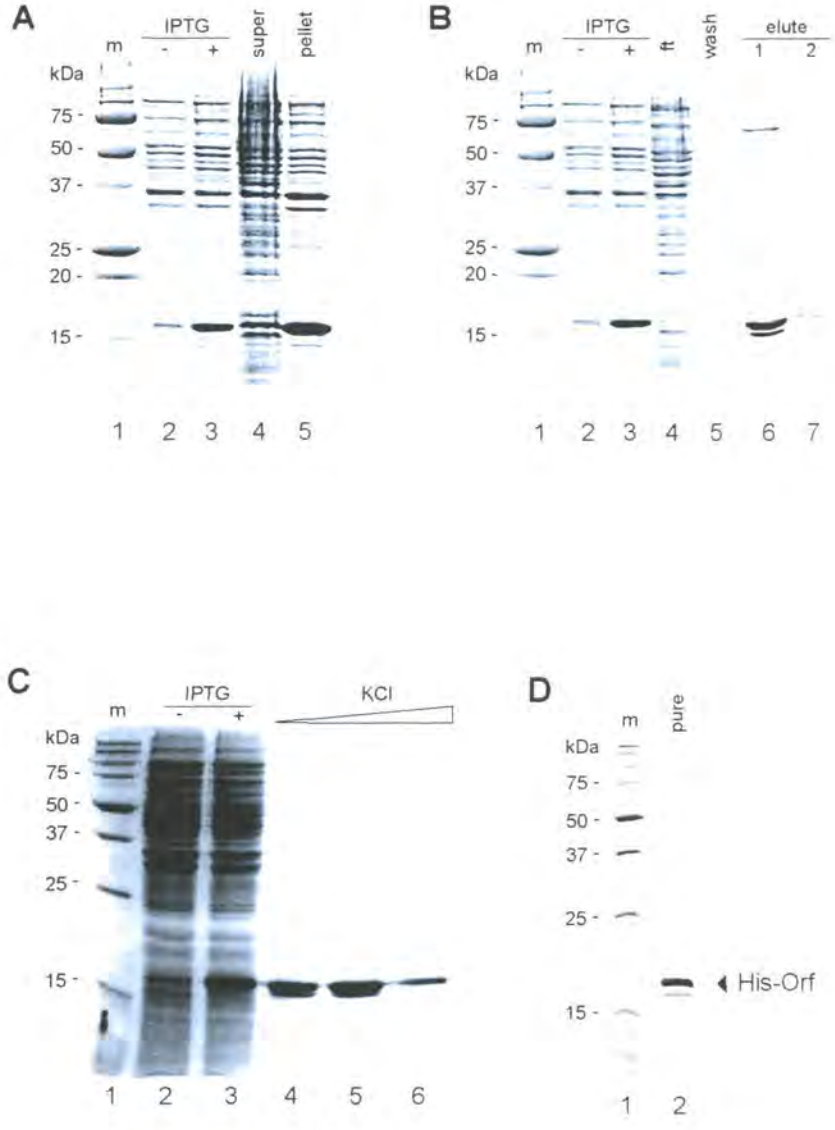

Figure 4.2. Overexpression and purification of His-Orf.
(A) Overexpression of His-Orf. Lane 1, molecular weight markers; Lane 2, uninduced cells; Lane 3, IPTG-induced cells; Lane 4, supernatant following cell lysis; Lane 5, pellet following cell lysis. **(B)** Purification of Orf by Ni^{2+} agarose affinity chromatography. Lane 1, MW markers; Lane 2 and 3, uninduced and induced cells; Lane 4, flow-through from column; Lane 5, wash fraction; Lane 6 and 7, consecutive fractions from the column, eluted with 250 mM imidazole **(C)** Purification of His-Orf by heparin agarose chromatography. Lanes 1-3 as in part A; Lanes 4-6, 0.5 M-0.7 M elute fractions from the column. **(D)** Purified His-Orf at 1.3 mg/ml. Lane 1, MW markers; Lane 2, purified His-Orf. Proteins were visualised on 15% SDS-PAGE stained with Coomassie blue. The double Orf bands are thought to be a gel artifact as both bands migrate at the same position in gel filtration chromatography (see Figure 4.3A)

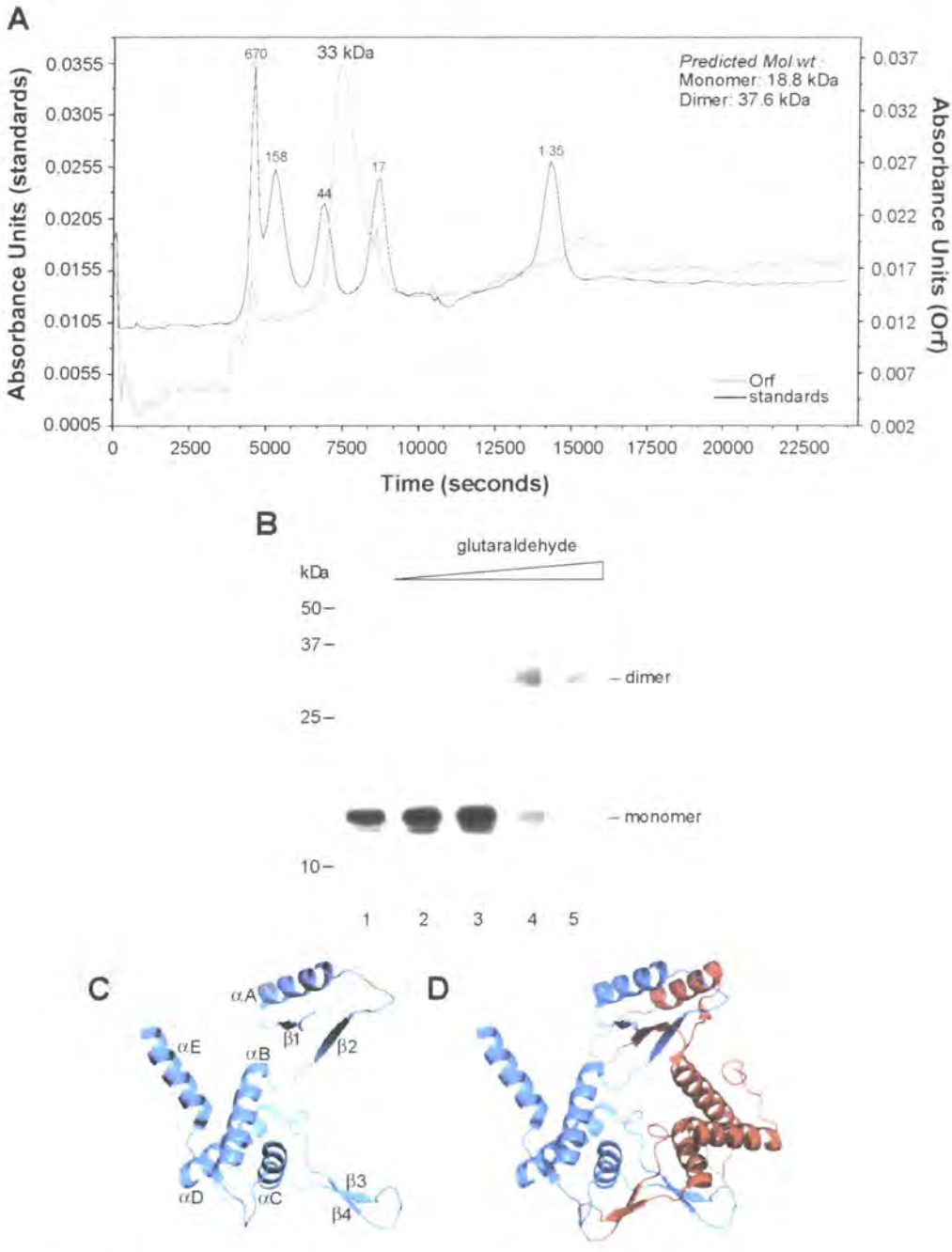

Figure 4.3. Quaternary structure analysis of His-Orf.
(A) Gel filtration chromatograph of purified His-Orf. The known MW of each marker protein is indicated. The position of the His-Orf peak correlates with a MW of 33 kDa, consistent with formation of a homodimer. **(B)** Gluteraldehyde crosslinking of His-Orf. The protein (0.1 mg) in 100 mM $NaHCO_3$ was mixed with 0.00025, 0.0025, 0.025, and 0.25% gluteraldehyde (lanes 2-5) or in its absence (lane 1) for 2 minutes and separated on 15% SDS-PAGE. A western blot was performed on the gel and monomeric and dimeric Orf species detected with anti-His antibodies. **(C)** Structure of the Orf monomer. Ribbon representation with N- and C-termini and secondary structure elements labelled. **(D)** Structure of the Orf dimer. Ribbon representation showing monomer A in blue and monomer B in red.

His-Orf was present in the soluble cell extract (Figure 4.2A, lane 4) than was recovered previously with wt-Orf (Figure 4.1A).

4.3.2.2 Nickel affinity chromatography

A 2 ml His-Select™ Nickel affinity gel column was equilibrated with lysis (equilibration) buffer and the clarified lysate applied. Bound proteins were eluted with 5 column volumes of elute buffer (lysis buffer containing 250 mM imidazole) and 2 ml fractions collected. The increased imidazole concentration causes His-tagged proteins to dissociate from the matrix, as they can no longer compete for binding sites on the Ni^{2+} resin. All of the His-Orf protein eluted in the first two fractions (Figure 4.2B, lanes 6 and 7) and these were pooled and dialysed against 250 mM KCl Buffer A. Several minor contaminants were present in these fractions.

4.3.2.3 Heparin chromatography

The dialysed His-Orf fractions were applied to a 1 ml of heparin agarose column and bound proteins eluted with a 0.25-1.0 M KCl gradient. Most of the His-Orf protein eluted at 0.5 M-0.7 M KCl fractions (Figure 4.2C, lanes 4-6) and appeared sufficiently pure on SDS-PAGE for biochemical studies (Figure 4.2). Peak fractions from the heparin column were pooled and dialysed against 500 mM KCl Buffer A + 50% glycerol and stored in aliquots at -80°C. A total of 2.6 mg of purified His-Orf was recovered at 1.3 mg/ml.

4.3.3 Quaternary structure analysis of His-Orf

Since His-Orf was recovered at a higher concentration than wt-Orf, the tagged protein was used to investigate the multimeric status of the polypeptide using gluteraldehyde cross-linking and size-exclusion chromatography. The crystal structure of λ Orf became available shortly after these studies.

4.3.3.1 Gluteraldehyde cross-linking of His-Orf

Chemical cross-linking with bifunctional reagents such as gluteraldehyde allow the recovery of stable oligomeric species and can indicate the quaternary structure of a protein in solution. Gluteraldehyde induces covalent cross-links between protein subunits in close proximity, allowing their visualisation by SDS-PAGE (Silva, 2004; Hermann *et al.*, 1979; Habeeb and Hiramoto, 1968; Sambrook, 2001).

To investigate the multimeric nature of λ Orf, we incubated His-Orf in the presence of increasing concentrations of gluteraldehyde as described in Chapter 2. Reaction products were analysed by SDS-PAGE followed by staining with Coomassie blue (data not shown). Monomeric His-Orf could be seen on these gels in the absence of the cross-linking agent, but it was difficult to detect protein on as the gluteraldehyde concentration increased. We therefore decided to analyse His-Orf by western blotting using monoclonal antibodies directed against the histidine tag. His-Orf without gluteraldehyde treatment migrated as a monomer as expected (Figure 4.3B, lane 1). Addition of gluteraldehyde above a concentration of 0.025% induced the formation of a stable His-Orf species of ~33 kDa, consistent with the formation of an Orf homodimer (Figure 4.3B, lane 3-5).

4.3.3.2 Gel filtration of His-Orf

To confirm that His-Orf exists as a dimer in solution we applied His-Orf at 0.5 mg/ml to a S200HR Sephacryl column in Buffer A containing 250 mM KCl. As gel filtration separates proteins according to their molecular mass, it was possible to investigate the oligomeric status of His-Orf by comparing its elution with known molecular weight standards analysed under the same conditions (see Chapter 2). The His-Orf protein eluted from the size exclusion column with a molecular weight of approximately 33 kDa (Figure 4.3 A). The His-Orf monomer has a predicted molecular weight of 18.8 kDa, suggesting that Orf forms a dimer.

These experiments were repeated with wt-Orf and yielded similar results (data not shown). Wild-type Orf eluted from the gel filtration column at ~29 kDa; the predicted molecular weight of an untagged Orf dimer is 33.6 kDa. Gluteraldehyde treatment also generated a cross-linked product with the correct molecular weight for an Orf dimer (data not shown). Taken together, the cross-linking and gel filtration data strongly favour the conclusion that Orf exists as a homodimer in solution.

4.3.3.3 Crystal structure of Orf

The 2.5 Å crystal structure of λ Orf was determined shortly after these studies were completed and revealed an α+β protein class with a novel fold (Maxwell *et al.*, 2005). The structure confirmed our findings that Orf exists as a dimer in solution. The protein monomers comprise 5 α helices and 4 β sheets (Figure 4.3 C), and interact to form an asymmetrical ring-like structure with a central funnel-like cavity (Figure 4.3D). Nine hydrophobic residues (Ile8, Val21, Ile36, Val95, Val33, Val75, Val86, Leu89, Ile97) from two discrete segments of Orf, namely β1-α1-β2 and α3-

β3-β4, contribute to the dimer interface (Maxwell *et al.*, 2005). The dimer exhibits asymmetry due to a twist in the backbone at residues Asn-40-Ser-42; however, the C-terminal tail of the dimer provides additional asymmetry. The last 20 residues of monomer A (Figure 4.3D) form an α-helix, which extends outwards from the body of the dimer.

4.4 MBP-Orf

During the purification of wild type and His_6-tagged Orf, we experienced difficulties with the solubility of the recombinant proteins. In both cases the majority of the overexpressed protein after cell lysis associated with the cell debris. Although it was possible to purify protein from the soluble fractions, the total yield was relatively poor. We therefore decided to construct a maltose binding protein (MBP)-fusion with Orf since enhanced solubility of MBP fusion proteins has been observed previously (Kapust and Waugh, 1999; Braun *et al.*, 2002; Sambrook, 2001). MBP-fused proteins can be purified in a single step using amylose affinity resin.

4.4.1 Construction of an MBP-Orf fusion

The *orf* gene was amplified from λ genomic DNA using oligonucleotides MBP-NinB1 and NinB2, designed to introduce an EcoRI site at the 5' end of the gene and a BamHI site at the 3' end. The PCR product was inserted into pMALc2 cut with these enzymes, resulting in fusion of MBP (the *malE* product) at the N-terminus of Orf (Figure 4.4). The *malE* gene has a deletion within its signal sequence leading to cytoplasmic expression of the MBP-Orf fusion protein. The resulting construct,

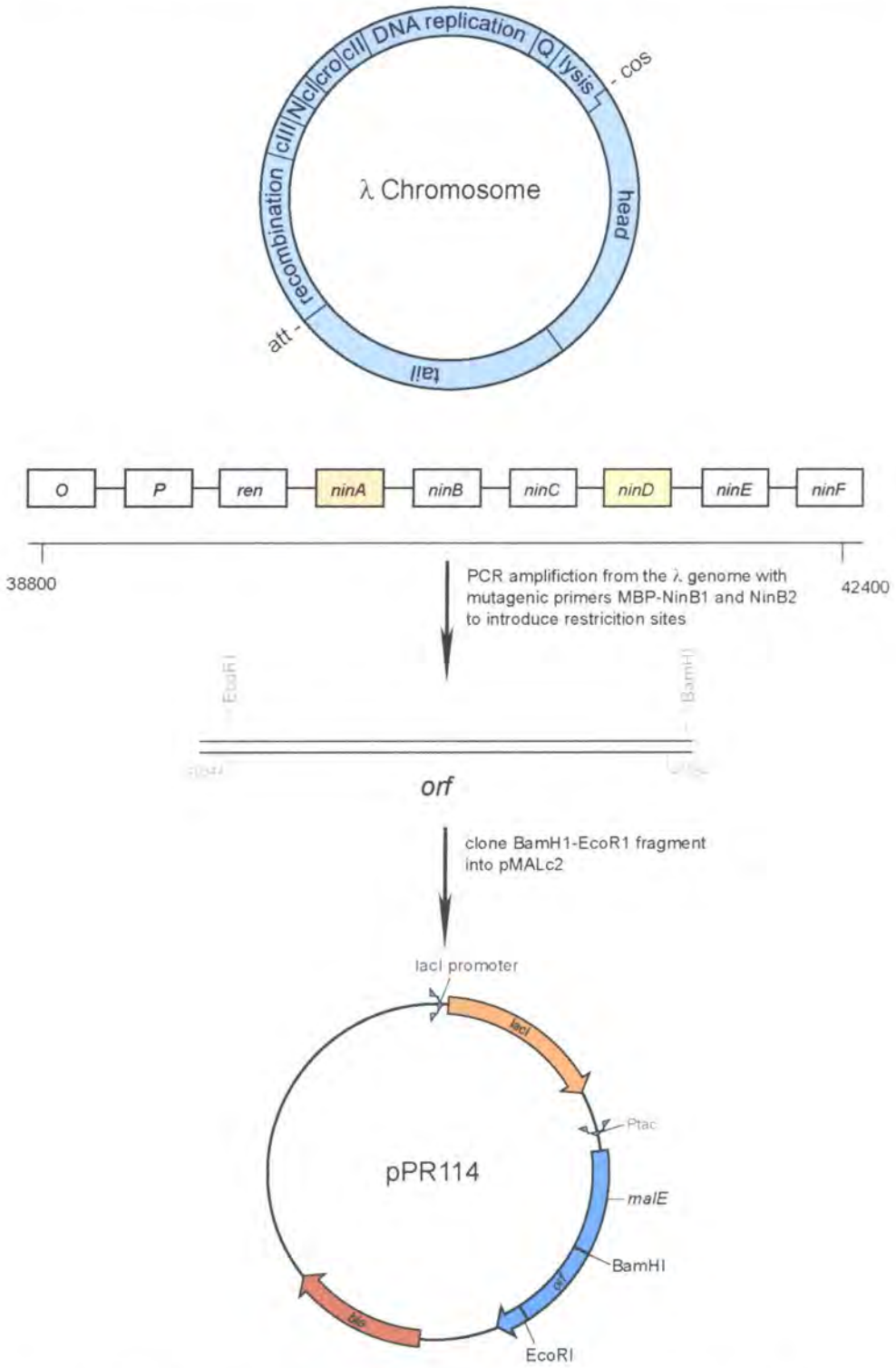

Figure 4.4. Construction of an MBP-Orf fusion from phage λ genomic DNA.
The *orf* gene was amplified by PCR from λ genomic DNA, using oligonucleotides to introduce restriction sites flanking the gene. The product was then inserted into the pMALc2 expression vector to give pPR114.

pPR114, was sequenced to check that no PCR-induced mutations had occurred within the *orf* gene.

4.4.2 Overexpression of MBP-Orf

E. coli BL21 (DE3) Codonplus was transformed with pPR114 as described in Chapter 2. This strain contains additional copies of tRNA genes that are considered to be rare in conventional *E. coli* backgrounds. By increasing their availability, a reduction in translation rates of overexpressed proteins due to rare tRNA depletion can be averted (Stratagene data sheet). Cells were grown in 2 x 1L of LB broth containing glucose and ampicillin. The inclusion of glucose eliminates the expression of amylase, which could degrade the amylose present in the affinity resin used for purifying MBP fusion proteins. Overexpression of MBP-Orf was induced at an A_{650nm} of 0.5 by addition of IPTG (0.3 mM) and cells incubated with aeration for a further 3 h at 37°C. Uninduced and induced samples were analysed by SDS-PAGE (Figure 4.5A, lanes 1 and 2) and a band of ~60 kDa corresponding to the MBP-Orf fusion protein was observed in lane 2.

4.4.3 Purification of MBP-Orf

4.4.3.1 Lysis

Overexpressed cells harvested by centrifugation (3 g) were resuspended in 30 ml of column buffer (contains EDTA to inhibit proteases that have a Ca^{+} cofactor) and lysed by sonication. Samples of the supernatant and pellet were analysed by SDS-PAGE (Figure 4.5B, lanes 4 and 5). As noted previously a large proportion of the Orf protein was retained in the post-sonicate pellet. However, a significant amount of the

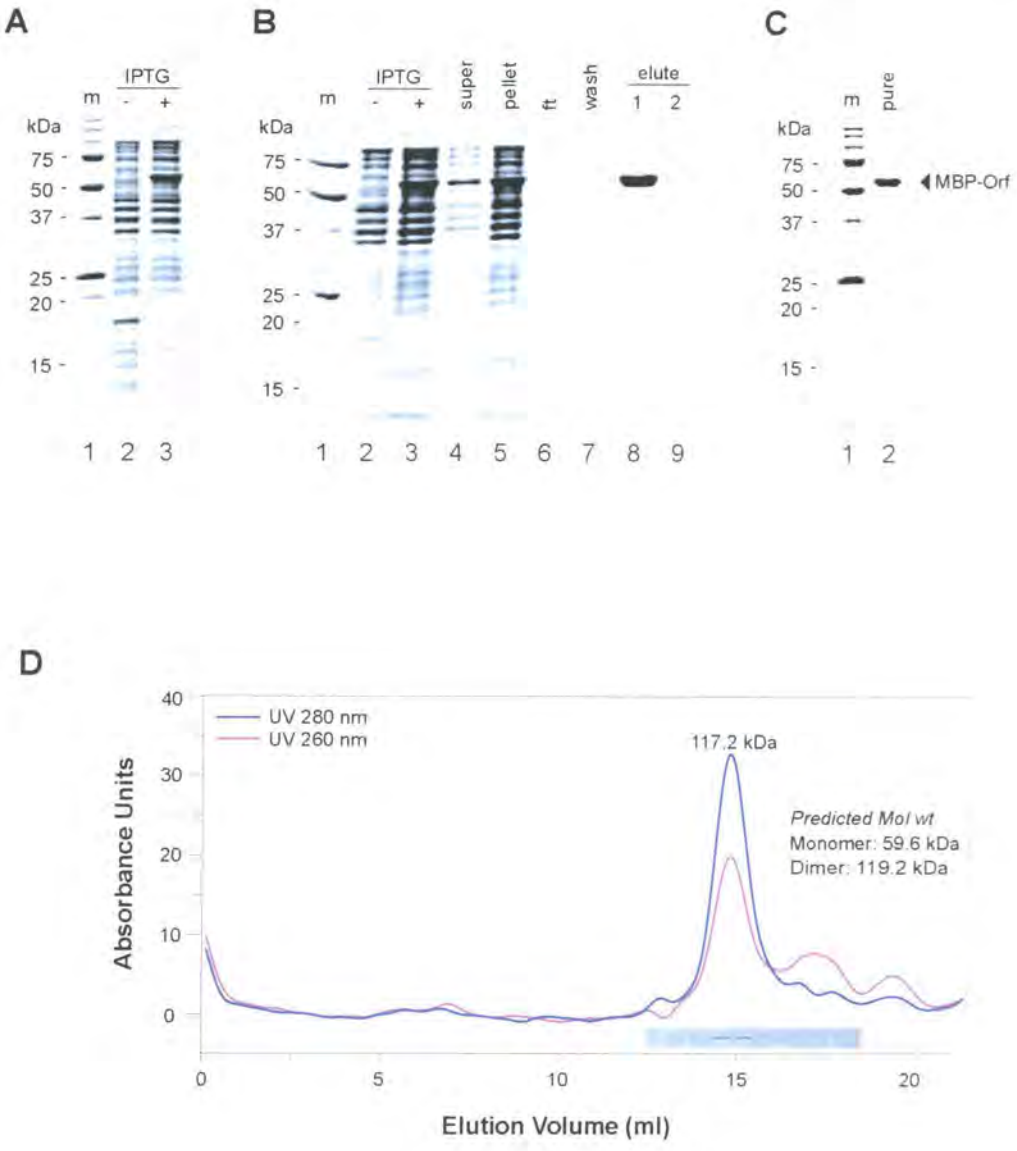

Figure 4.5. Overexpression and purification of MBP-Orf.
(A) Overexpression of MBP-Orf. Lane 1, MW markers; Lane 2, uninduced cells; Lane 3, induced cells. **(B)** Purification of MBP-Orf by amylose affinity chromatography. Lane 1, MW markers; Lane 2 and 3, uninduced and induced cells; Lane 4, supernatant following cell lysis; Lane 5, pellet following cell lysis; Lane 6, flow-through from column; Lane 7, wash fracion; Lane 8 and 9, consecutive fractions from the column, eluted with 10 mM Maltose. **(C)** Purified MBP-Orf at 4.77 mg/ml. **(D)** Gel filtration chromatogram of purified MBP-Orf. The protein eluted at ~117 kDa consistent with the formation of an MBP-Orf dimer. Molecular mass standards (BioRad) were loaded separately under the same conditions in order to allow MW estimations (not shown). Proteins were visualised on 15% SDS-PAGE stained with Coomassie blue.

fusion protein was present in the supernatant, suggesting that attachment of the MBP domain may have improved the recovery of soluble Orf protein.

4.4.3.2 Amylose affinity chromatography

Aproximately 2 ml of amylose resin was mixed with the clarified cell lysate containing MBP-Orf in column buffer and incubated with mixing on a rotation wheel at 4°C for 1 h. A column was poured with the mixture and the flow through collected before the matrix was washed with column buffer. Bound proteins were eluted from the amylose matrix in the presence of 10 mM maltose; most of the MBP-Orf protein appeared in the first fraction (Figure 4.5B, lane 8). The fusion protein was fairly dilute at this stage and so the sample was concentrated using Ym30 Centricon Concentrators. A total of 4.77 mg of MBP-Orf was recovered at a concentration of 4.77 mg/ml in storage buffer. The protein was divided into aliquots and stored at -80°C. A single band of ~60 kDa was visualised by SDS-PAGE after this single purification step (Figure 4.5C).

4.4.4 Quaternary structure analysis of MBP-Orf

To ensure that the presence of the relatively large N-terminal MBP tag (43 kDa) did not grossly affect Orf dimerisation, we analysed the MBP-Orf by chemical cross-linking and gel filtration as before.

4.4.4.1 Gluteraldehyde cross-linking of MBP-Orf

A gluteraldehyde cross-linking assay was performed with MBP-Orf as described previously with His-Orf. The formation of apparently dimeric MBP-Orf was noted

with increasing gluteraldehyde concentrations (data not shown). Larger, possibly multimeric forms, were seen at very high gluteraldehyde concentrations and these could be due to non-specific cross-linking of adjacent dimers. MBP is known to be a monomer in solution (Sharff *et al.*, 1992; Spurlino *et al.*, 1991), so the dimers seen here are most likely due to Orf subunit association.

4.4.4.2 Gel filtration of MBP-Orf

To verify the cross-linking data, MBP-Orf (1 mg/ml) was analysed on a 24 ml Superose-6 gel filtration column in 250 mM KCl Buffer A. The protein eluted from the gel filtration column at 117 kDa (Figure 4.5D). Since the MBP-Orf monomer has a predicted MW of 59.6 kDa, the elution profile corresponds well to a homodimeric protein. The presence of a 60 kDa protein in the elute fractions was confirmed by SDS-PAGE analysis of the relevant fractions.

4.5 λ Orf C-terminal deletions

As noted in section 4.3.3.3, the crystal structure of Orf reveals the presence of an α-helix at the C-terminus in one of the subunits; the same region is disordered in the other subunit (Maxwell *et al.*, 2005). The 20 residues containing this helix protrude from the dimer and are ideally positioned to be involved in DNA binding or protein: protein interactions. Removal of these C-terminal residues could potentially prove useful in determining their involvement in binding DNA and/or other recombination enzymes. Two C-terminal deletions were therefore constructed; the first, OrfΔC6 removes a flexible tail of 6 residues extending from the end of the helix, the second, OrfΔC19, eliminates the entire region including the α-helix.

4.5.1 Cloning of *orfΔC6* and *orfΔC19*

The truncated genes *orfC6* and *orfC19* were amplified by PCR from λ genomic DNA using mutagenic primers to generate an NdeI site at the start codon and a BamHI site downstream of the required coding sequence. The PCR products were inserted into pT7-7 using these restriction enzymes and the resulting constructs pPR109 (*orfΔC6*) and pPR110 (*orfΔC19*) were confirmed by DNA sequencing. The inserts from these constructs were also transferred into pET14b to give pPR111 (*orfΔC6*) and pPR112 (*orfΔC19*), allowing expression with a histidine tag at the N-terminus. Unfortunately, due to time constraints, further study of these mutant proteins was not possible. However, recent studies have revealed that the OrfΔC6 mutant displays reduced ssDNA binding, while removal of 19 residues from the C-terminus completely abolishes DNA binding (FA Curtis and GJ Sharples, personal communication).

4.6 *Escherichia coli* prophage DLP12 Orf151

As discussed in Chapter 3, Orf151 appears to be a distantly related homolog of λ Orf, located in the *E. coli* prophage DLP12 (Mahdi *et al.*, 1996). This similarity encouraged us to purify and analyse the biochemical properties of Orf151 in parallel with λ Orf. In this section we describe the cloning and purification of wild-type, His_6-tagged and MBP-fused Orf151 and investigate its multimeric status.

4.6.1 Cloning of *E. coli orf151*

An *orf151* pT7-7 construct (pGS904), amplified from the *E. coli* K12 genome, was available at the outset of this project (supplied by GJ Sharples). The insert from pGS904 was released by cleavage with NdeI and BamHI and transferred into

pET14b. The resulting plasmid, pPR101, provided a His_6-tagged version of the Orf151 protein. An MBP-Orf151 fusion was also constructed as part of a final year undergraduate project (LY Bowers). The *orf151* gene was amplified from *E. coli* K12 genomic DNA and recovered in pMALc2, as with the MBP-Orf construct to generate pPR116. The integrity of all of these genes was verified by DNA sequencing.

4.6.2 Orf151 overexpression

BL12 (DE3) pLysS competent cells were transformed with each of the three-overexpression plasmid constructs. Growth and expression conditions were as described in previous sections of this chapter for the Orf versions of these constructs. All of the recombinant proteins were highly expressed following induction with IPTG (Figure 4.6A and Figure 4.7A; note that the wt-Orf151 data is not included). Cells carrying pPR116 yielded very high levels of overexpressed MBP-Orf151 (Figure 4.7A). In light of the difficulties encountered with λ Orf, we did not attempt the purification of untagged Orf151. Cells expressing His-Orf151 and MBP-Orf151 were harvested by centrifugation as described previously (section 4.2.2).

4.6.3 Purification of *E. coli* Orf151 proteins

4.6.3.1 Lysis

As with His-Orf, cells overexpressing the His-tagged Orf151 protein were resuspended in lysis buffer and lysed by sonication. Cells expressing MBP-Orf151 were resuspended in column buffer prior to disruption by sonication. In both cases the lysate was cleared by centrifugation and pellet and supernatant samples were

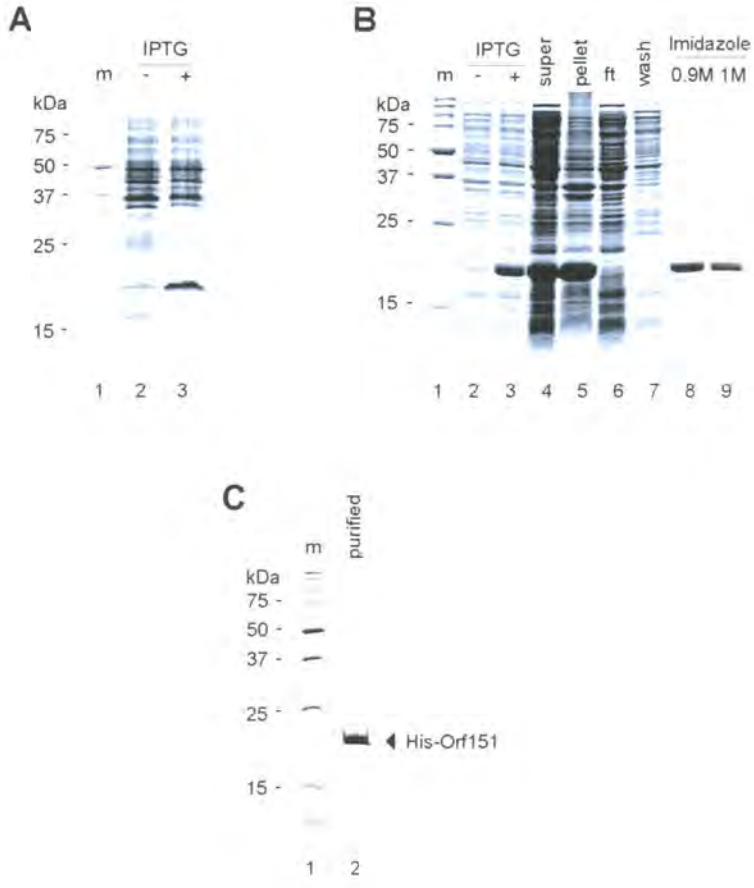

Figure 4.6. Overexpression and purification of His-Orf151
(A) Overexpression of His-Orf151. Lane 1, MW markers; Lane 2, uninduced cells; Lane 3, induced cells. **(B)** Purification of His-Orf151 by Ni^{2+} -agarose affinity chromatography. Lane 1-3, as in part A; Lane 4, supernatant following cell lysis; Lane 5, pellet following cell lysis; Lane 6, flow-through from column; Lane 7, wash fraction; Lane 8 and 9, consecutive fractions from the column, eluted with 0.9 M and 1.0 M imidazole. **(C)** Purified His-Orf151 at 5.2 mg/ml. A number of contaminants are visible in this sample suggesting the protein is only 90-95% pure. Proteins were visualised on 15% SDS-PAGE stained with Coomassie blue.

examined by SDS-PAGE (Figure 4.6B and Figure 4.7B). A significant proportion of both His-Orf151 and MBP-Orf151 samples remained in the pellet (Figure 4.6B, lanes 4 and 5, and Figure 4.7B lanes 4 and 5). However sufficient quantities of each was present in the clarified lysate for use in subsequent purification steps.

4.6.3.2 Purification of His-Orf151

The supernatant containing His-Orf151 was applied to 1 ml of Ni^{2+}-agarose (as in Section 4.3.3.2) column. After washing, bound proteins were eluted with increasing imidazole concentration. His-Orf151 bound to the matrix extremely tightly, failing to elute at 250 mM imidazole, as was the case with His-Orf. An imidazole gradient from 0.25-1.0 M was employed. Although some of the protein eluted at 300 mM imidazole, the majority was released between 500 and 800 mM imidazole (data not shown), along with some contaminating proteins. The pooled fractions were dialysed in multiple steps to try to remove the imidazole, however, the protein consistently precipitated at lower imidazole concentrations. The 900 mM and 1.0 M fractions from the column appeared to contain pure His-Orf151 protein as judged by SDS-PAGE (Figure 4.6B, lanes 8 and 9). These two fractions were pooled and dialysed against 500 mM KCl Buffer A + 50% glycerol and stored at -80°C for future use. A number of contaminating proteins were evident in this purified sample (Figure 4.6C) suggesting that the His-Orf has a tendency to precipitate out of solution. A total of 4.68 mg of protein recovered at 5.2 mg/ml, however, doubts remain over the solubility of this preparation.

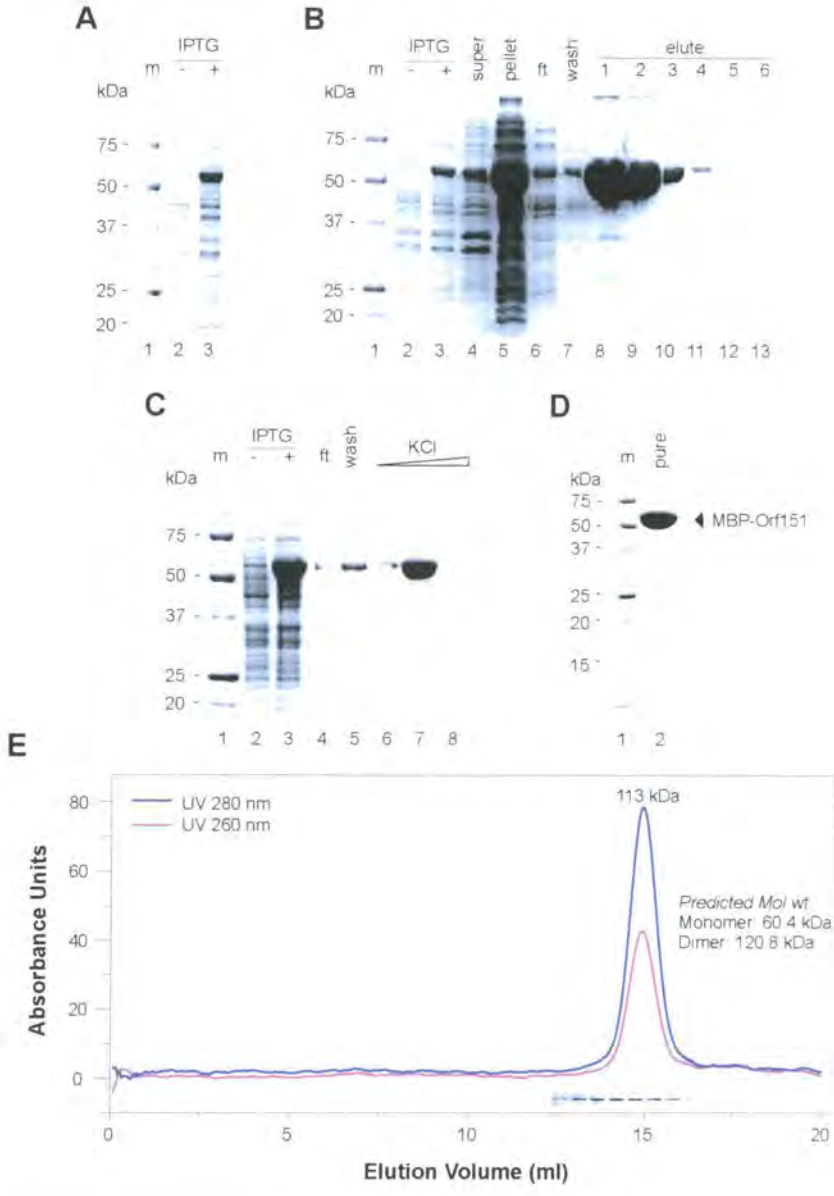

Figure 4.7. Overexpression and purification of MBP-Orf151
(A) Overexpression of MBP-Orf151. Lane 1, MW markers; Lane 2, uninduced cells; Lane 3, induced cells **(B)** Purification of MBP-Orf151 by amylose affinity chromatography, Lane 1, MW markers; Lane 2 and 3, uninduced and induced cells; Lane 4, supernatant following cell lysis; Lane 5, pellet following cell lysis; Lane 6, flow-through from column; Lane 7, wash fracion; Lane 8-13, consecutive fractions from the column, eluted with 10 mM Maltose. **(C)** Purification of pooled amylose fractions from by heparin agarose chromatography, Lane 1, MW markers; Lanes 2 and 3, uninduced and induced cells; Lane 4, flow-through from column; Lane 5, wash fraction; Lane 6-8, 0.8-1 M KCl elute fractions from the column. **(D)** Purified MBP-Orf151 at 17.5 mg/ml. Lane 1, MW markers; Lane 2, purified MBP-Orf151. **(E)** Gel filtration chromatogram of purified MBP-Orf151. The protein eluted at ~113 kDa, consistent with formation of an MBP-Orf151 dimer. Molecular mass standards (BioRad) were loaded separately under the same conditions in order to allow MW estimations (not shown). Proteins were visualised on 15% SDS-PAGE stained with Coomassie blue.

4.6.3.3 Purification of MBP-Orf151

The cleared lysate containing MBP-Orf151 (Figure 4.7B, lane 4) was mixed with ~5 ml of amylose resin slurry prepared as described in section 4.3.3.2. The mixture was incubated with stirring at 4°C for 1 hour prior to pouring of the column. The column was washed and bound proteins eluted with 10 mM maltose as before. Significant quantities of MBP-Orf151 eluted from the column in the first six 1 ml fractions, although some contaminants were visible upon SDS-PAGE analysis (Figure 4.7B lanes 8-13). The elute fractions were pooled and dialysed against 200 mM KCl Buffer A before loading on a 2 ml heparin agarose column in the same buffer. MBP-Orf151 bound tightly to this column, eluting between 800mM and 1M KCl (Figure 4.7C, lanes 6-8). The ~60 kDa protein (consistent with its predicted molecular weight) was dialysed against 500 mM KCl Buffer A + 50% glycerol and appeared pure when analysed by SDS-PAGE (Figure 4.7D). A total of 26.25 mg of purified MBP-Orf151 was recovered at a concentration of 17.5 mg/ml.

4.6.3.4 Quaternary structure analysis of MBP-Orf151

MBP-Orf151 (1 mg/ml) was analysed by gel filtration in 250 mM KCl Buffer A as described previously for MBP-Orf (section 4.4.4.2). A protein peak, corresponding to a MW of 113kDa, fits with the formation of a dimer (Figure 4.7 E).

MBP-Orf151 dimers were also seen following exposure to gluteraldehyde at concentrations above 0.025% (data not shown). Furthermore, Orf151 self-association was detected in the Matchmaker III yeast-two hybrid system (GJ Sharples and MD Watson, personal communication). We conclude that DLP12 Orf exists as a homodimer in solution in keeping with its predicted similarity to λ Orf.

4.7 *Staphylococcus aureus* phage ϕETA Orf20

In addition to DLP12 Orf151, another distantly-related homolog of Orf, *Staphylococcus aureus* phage ϕETA Orf20, was identified in Chapter 3 (Yamaguchi *et al.*, 2000). Unlike Orf and Orf151, ETA20 carries an additional conserved C-terminal domain that resembles HNH family nucleases. In this section we describe the cloning and purification of this protein and an investigation of its quaternary structure. We also attempted the construction of a truncated protein, eliminating 82 C-terminal residues containing the HNH domain, to probe the putative nuclease activity of ETA20.

4.7.1 Cloning of *ETA20* and *ETA20ΔC82*

ETA20 and *ETA20ΔC82* were amplified from *S. aureus* phage ϕETA genomic DNA, kindly provided by M Sugai (Department of Microbiology, Hiroshima University). PCR products were cloned into pT7-7 and pET14b vectors using NdeI and HindIII restriction enzymes. FA Curtis and LY Bowers generated these constructs. Overexpression and purification of the wild-type and His_6-tagged versions of ETA20 and ETA20ΔC82 was attempted but with little success. Since MBP fusions were successful in obtaining purified Orf and Orf151 proteins, we utilised the same approach with the *S. aureus* phage protein.

The *ETA20* gene was successfully inserted into pMALc2 and the resulting construct designated pPR113. The integrity of the cloned gene was verified by sequencing and found to be identical to the published genomic sequence. Despite several attempts we were unable to clone *ETA20ΔC82*. Time constraints prevented

further cloning attempts and so studies continued using only the full-length protein fused to MBP.

4.7.2 Overexpression and purification of MBP-ETA20

Overexpression of the MBP-ETA20 fusion protein was observed following induction of BL21 (DE3) pLysS carrying pPR113 (Figure 4.8A). The ~70 kDa correlates with the molecular weight of MBP-ETA20 (69 kDa) predicted from its sequence. Harvested cells (4 g) were resuspended in 40 ml column buffer and lysed by sonication. The cell debris was sedimented by centrifugation and a significant amount of the fusion protein remained in the soluble fraction (Figure 4.8, B lanes 4 and 5).

The supernatant was mixed with 1 ml of amylose resin and the mixture prepared and the column poured as described (section 4.4.3.2). Flow through and wash samples were collected for analysis (Figure 4.8 B, lanes 6 and 7) and bound proteins eluted with 10 mM maltose. All of the bound MBP-ETA20 eluted from the column in the first fraction (Figure 4.8B, lane 8). The protein appeared pure enough at this stage for biochemical analysis and was stored in 500 mM KCl Buffer A + 50% glycerol at -80°C in 20 µl aliquots (Figure 4.8C). A total of 1.6 mg of MBP-ETA20 was recovered at 1.6 mg/ml.

4.7.3 Quaternary structure analysis of MBP-ETA20

The purified MBP-ETA20 protein, at 1 mg/ml, was analysed by gel filtration to investigate its ability to form oligomers. The protein eluted at 6.7 ml close to the void position (Figure 4.8 D). Thyroglobulin, the largest of the molecular weight

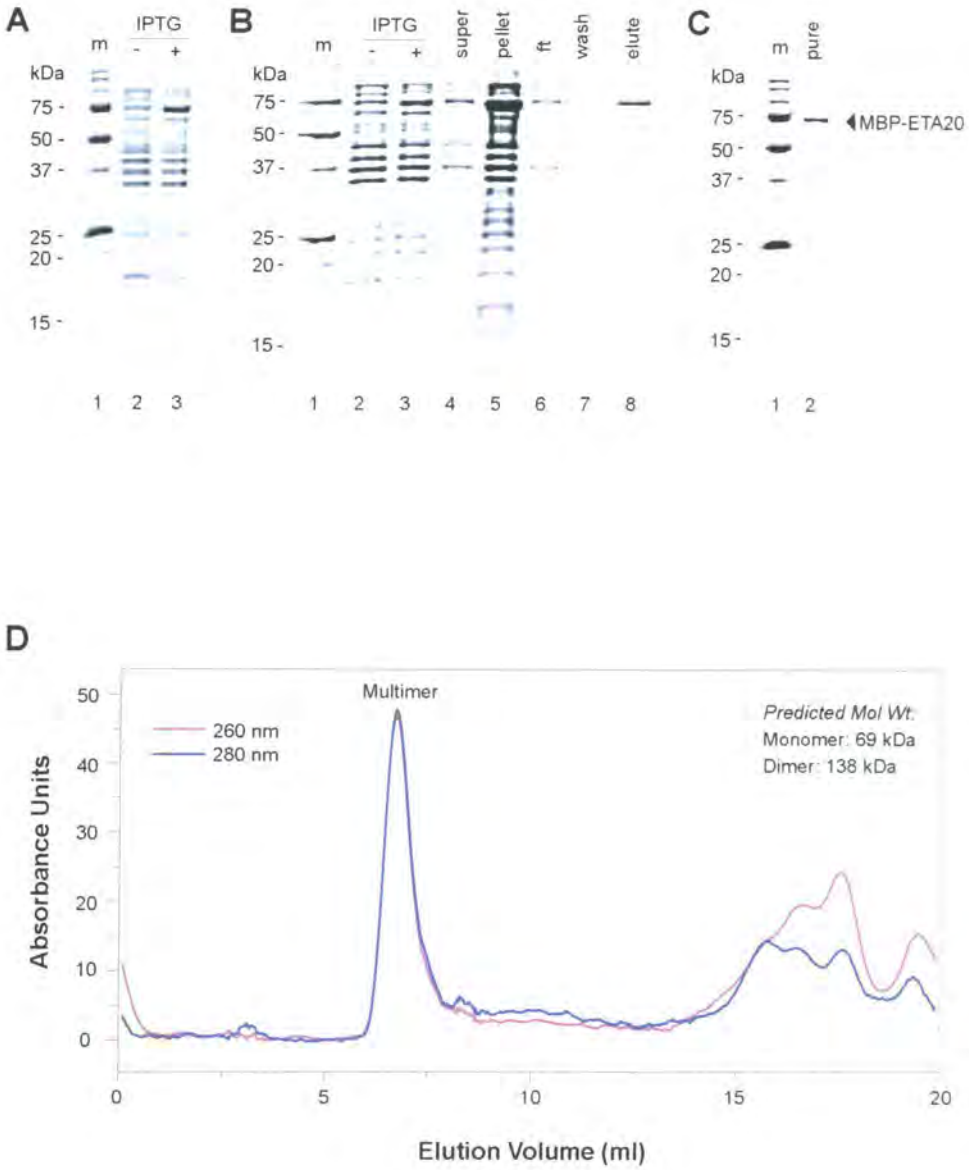

Figure 4.8. Overexpression and purification of MBP-ETA20
(A) Overexpression of MBP-ETA20. Lane 1, MW markers; Lane 2, uninduced cells; Lane 3, induced cells **(B)** Purification of MBP-ETA20 by amylose affinity chromatography, Lane 1, MW markers; Lane 2 and 3, uninduced and induced cells; Lane 4, supernatant following cell lysis; Lane 5, pellet following cell lysis; Lane 6, flow-through from the column; Lane 7, wash fracion; Lane 8, fraction from the column, eluted in 10 mM Maltose. **(C)** Purified MBP-ETA20. Lane 1, MW markers; Lane 2, purified MBP-ETA20 at 1.6 mg/ml. **(D)** Gel filtration chromatogram of purified MBP-ETA20. The protein eluted as a large complex, much larger than any of the molecular weight standards run under the same conditions. Proteins were visualised on 15% SDS-PAGE stained with Coomassie blue.

standards, eluted at 12 ml with an expected molecular weight of 670 kDa. It appears that MBP-ETA20 forms large multimeric complexes in solution.

A gluteraldehyde cross-linking assay was performed with MBP-ETA20 following the same protocol as outlined in section 4.3.4.1. In the absence of gluteraldehyde MBP-ETA20 appears as a monomer on SDS-PAGE. However, addition of gluteraldehyde, even at very low concentrations, resulted in smearing at the top of SDS polyacrylamide gels (data not shown). Unlike Orf and Orf151, MBP-ETA20 does not appear to form dimers in solution. Instead it has a tendency to form large protein aggregates. The differences in quaternary structure between these proteins could be a consequence of the extra C-terminal residues found in ETA20 and certainly warrants further investigation.

4.8 Discussion

The main objective of this study was to purify λ Orf and two potential orthologs, *E. coli* DLP12 Orf151 and *S. aureus* ϕ ETA Orf20 (ETA20). Orf was successfully purified in its native form, as a histidine-tagged protein and as an MBP-fusion protein. In all three cases, cells grown to an A_{650nm} 0.5 with 3 hours incubation at 37°C after addition of IPTG afforded the best conditions for overexpression. A substantial proportion of the native and His-tagged Orf proteins precipitated with the cell debris following cell lysis. This hampered the recovery of large quantities of these proteins and led us to purify an MBP-fusion version, which proved successful. Precipitation of MBP-Orf after cell disruption was also observed, but its improved overexpression and increased solubility allowed the recovery of significantly more purified Orf protein than had been possible with the native and His-tagged forms.

Gel filtration and gluteraldehyde cross-linking of Orf showed that it exists as a homodimer in solution. The crystal structure of Orf, solved by a structural genomics consortium at the University of Toronto, revealed that Orf does indeed exist as a dimer, confirming our results.

Two C-terminal deletion mutants of Orf were constructed with the intention of overexpressing and purifying the truncated proteins and investigating their DNA binding abilities and/or proteins interactions. The overexpression constructs were successfully engineered and would have produced both native and histidine-tagged forms of the mutant proteins. Unfortunately due to time constraints these proteins were not overexpressed and purified at this stage.

During the purification of Orf it was noted that the protein bound to both ssDNA cellulose and heparin agarose matrices. As heparin is an analogue of DNA, binding to this matrix and to ssDNA cellulose is clearly indicative of DNA binding activity. As Orf is critical for the early stages of recombinational exchange (Sawitzke and Stahl, 1992; Sawitzke and Stahl, 1994) and supplies an activity analogous to *E. coli* RecFOR proteins that assist strand exchange reactions mediated by RecA (Webb *et al.*, 1997; Poteete, 2004; Kowalczykowski *et al.*, 1994), it is expected that Orf will interact with DNA and other recombination proteins. The genetic evidence, coupled with information drawn from analysis of the crystal structure, encouraged us to investigate these possibilities using biochemical assays. The results of these experiments are described in subsequent chapters of this thesis.

The Orf homologs, Orf151 and ETA20, were also successfully purified in milligram quantities from histidine-tagged and MBP-fusion expression constructs. Gel filtration and gluteraldehyde-crosslinking indicated that Orf151, like Orf, exists

as a dimer in solution. In contrast, ETA20 apparently formed a multimeric complex, indicating that its divergence from Orf, perhaps *via* the C-terminal HNH extension, confers a structurally distinct function from that proposed for Orf. Attempts to construct a C-terminal deletion mutant of ETA20, ETA20ΔC82, were unsuccessful. Further investigation of Orf151 and ETA20 is required to elucidate their roles, if any, in genetic exchange, and their functional similarities to λ Orf.

Chapter 5

Purification of other recombination proteins

5.1 Introduction

As mentioned previously, phage λ encodes a third protein, Orf (NinB), in addition to Exo and β, influential in the initial stages of recombination (Sawitzke and Stahl, 1992; Sawitzke and Stahl, 1994; Sawitzke and Stahl, 1997; Tarkowski *et al.*, 2002; Poteete, 2004). Orf can replace the function of the *E coli* RecFOR complex during Red-mediated recombination (Poteete, 2004; Sawitzke and Stahl, 1994). λ Exo is a highly processive 5' to 3' exonuclease that binds to dsDNA termini and degrades one chain of the duplex to yield 3' single-stranded tails (Kovall and Matthews, 1998). The ssDNA generated serves as the substrate for its molecular partner β required for annealing homologous sequences (Passy *et al.*, 1999). As described in the introduction, significant progress has been made in our understanding of Exo structure and activity (Little, 1967; Kovall and Matthews, 1997), DNA binding properties of β (Kmiec and Holloman, 1981; Muniyappa and Radding, 1986), interactions between Exo and β (Radding, 1971) and how the Red pathway mediates recombination (Stahl *et al.*, 1997). Despite these advances, we still do not know how Exo and β activities are coupled, what regulates Exo degradation and how Orf is accommodated within the pathway. In order to investigate the potential involvement and interactions between Orf and Red recombinases both Exo and β were purified as N-$His_{(6)}$ fusions.

Orf is able to substitute for the *E. coli* RecF, RecO and RecR recombination proteins (see Introduction). Individually and corporately, RecFOR display multiple *in vitro* activities; notably, the coordinated action of all three proteins mediates loading of RecA onto SSB-coated ssDNA to accelerate DNA strand exchange (Morimatsu and Kowalczykowski, 2003). SSB disrupts RecO-RecR complexes to form RecO-SSB complexes (Umezu and Kolodner, 1994), RecO promotes RecF interactions with SSB and RecR proteins, and RecF acts to stabilize RecO-RecR complexes in the presence of SSB (Hegde et al., 1996a). Although only RecO seems to directly interact with SSB, the entire RecFOR complex is associated with SSB. Given the ability of Orf to replace the functions of the *E. coli* RecFOR it was important that we investigate the interaction of λ Orf with *E. coli* SSB. We decided, therefore, to purify SSB and a C-terminal deletion mutant, lacking the last ten C-terminal amino acids, SSBΔC10. The acidic C-terminal domain of *E. coli* SSB is disordered (Savvides *et al.*, 2004) and serves as a binding site for many proteins participating in DNA metabolism (Umezu and Kolodner, 1994; Kantake *et al.*, 2002; Handa *et al.*, 2001; Genschel *et al.*, 2000).

In this chapter, we describe the overexpression and purification of λ Exo and β proteins and *E. coli* SSB and SSBΔC10. Analysis of protein: protein interactions between λ and bacterial proteins will be reported in subsequent chapters of this thesis.

5.2 Overexpression and purification of *E. coli* SSB and SSBΔC10

Overexpression constructs of *E. coli* SSB (pCC146) and SSBΔC10 (pCC180) were created as described (Cadman and McGlynn, 2004), and were kindly provided by Peter McGlynn (University of Aberdeen, Aberdeen, UK). Both harbour the relevant gene/gene fragment in pET22b (minus the optional C-terminal His-Tag) and regulated by the vector T7 promoter. The constructs were transformed with BL21 (DE3) pLysS (pCC146) and BL21-SI (pCC180) competent cells as described in the materials and methods. The BL21-SI cells contain a chromosomal insertion of the T7 RNA polymerase gene under control of the salt-inducible *proU* promoter. Addition of 0.3 M NaCl to the growth medium induces expression of the polymerase and hence any genes cloned downstream of the T7 promoter.

Cells containing recombinant clones were inoculated as small-scale cultures for overnight growth. BL21-SI carrying pCC180 (SSBΔC10) was grown in LBON (LB without NaCl) broth supplemented with ampicillin at 150 μg/ml at 30°C. BL21 (DE3) pLysS pCC146 were grown in LB broth supplemented with ampicillin at 150 μg/ml and chloramphenicol at 75 μg/ml at 37°C. The overnight cultures for each clone were used to inoculate 1 litre of fresh medium (LB or LBON) containing the relevant antibiotics, and cultures incubated at 30°C and 37°C respectively, with aeration until A_{650nm} of 0.5. Expression was induced with 1 mM IPTG in BL21 (DE3) pLysS cells or with 300 mM sterile NaCl in BL21-SI cells, and incubation continued for 3 hours at the relevant temperatures. Cells were harvested as previously described (Section 4.2.2) and stored at 4°C. Overexpression of both SSB and SSBΔC10 was observed at this time by Coomassie blue staining of SDS-PAGE gels (data not shown).

Both recombinant SSB proteins were purified following a published protocol (Cadman and McGlynn, 2004), although the final gel filtration step was omitted. Cell pellets were resuspended in 1.0 M NaCl Buffer A, lysed by sonication as described (Section 2.3.10ii) and the lysate cleared by centrifugation (JA-20 rotor, 13 000 rpm, 20 minutes, 4°C). Polymin P was added to the supernatant at a final concentration of 0.4% (v/v). Ammonium sulphate solution was added slowly up to a concentration of 19% (v/v) (as described by Cadman and McGlynn, 2004). These steps were performed at 4°C to limit protein denaturation. After precipitation the pellet was separated by centrifugation (13 000 rpm, 10 mins, 4°C) and resuspended in 100 mM NaCl Buffer A for further purification. Supernatant and pellet samples were analysed by SDS-PAGE at this stage and SSB proteins of the expected mass were observed in each precipitate (data not shown).

SSB and SSBΔC10 supernatants were applied to 3 ml Q-sepharose anion exchange columns, washed with 100 mM NaCl Buffer A, and eluted from the column with increasing salt concentration. Both proteins eluted between 200 mM and 1.0 M NaCl (data not shown) and recovered fractions were pooled and dialysed back into 100 mM NaCl Buffer A. Supernatants were mixed with 1 ml pre-equilibrated heparin agarose resin on an orbital shaker at 175 rpm for 15 minutes at 4°C before pouring of the column. Bound proteins were eluted from the heparin matrix between 0.4 M and 1.0 M NaCl. Elute fractions were stored in 500 mM NaCl Buffer A + 50% (v/v) glycerol at -80°C and appeared >99% pure on SDS-PAGE (Figure 5.1). A total of 9.6 mg of SSB at 3.2 mg/ml and 11 mg of SSBΔC10 at 4.4 mg/ml were obtained. As noted previously (Cadman and McGlynn, 2004), the wild-type *ssb* allele present in BL21 strains will result in contamination of the SSBΔC10

mutant sample with trace amounts of SSB. Differences in mobility between full-length and truncated SSB proteins analysed by SDS-PAGE (Figure 5.1), allowed us to confirm that no wild-type SSB is apparent in the SSBΔC10 sample. We conclude that any contamination of SSBΔC10 is negligible and unlikely to affect subsequent biochemical assays.

5.3 Overexpression and purification of λ Exo

The λ *exo* gene was amplified by PCR from λ genomic DNA using Pfx polymerase and oligonucleotides Exo-1 and Exo-2 that introduce NdeI and BamHI restriction sites on either side of the coding region. The PCR product was digested with these enzymes and inserted into pET14b to give pFC150 (generated by F.A. Curtis). This construct expresses an N-terminal $His_{(6)}$-Exo fusion. The integrity of the cloned gene was confirmed by DNA sequencing.

BL21 (DE3) pLysS pFC150 cells were used to inoculate small-scale overnight cultures, which were incubated at 37°C with aeration. These cultures were used to inoculate 1 litre of fresh LB broth supplemented with 150 μg/ml ampicillin and 50 μg/ml chloramphenicol. Expression of the plasmid-encoded His-Exo was induced with 1 mM IPTG when the culture reached an A_{650nm} of 0.5 and cells incubated for a further 3 hours at 37°C. High level expression of Exo protein was visualised by SDS-PAGE (Figure 5.2).

Cells were harvested by centrifugation (JA-10 rotor, 6000 rpm, 10 mins, 4°C), resuspended in 10 ml of lysis buffer and disrupted by sonication as described (Section 2.3.10ii). A large proportion of the protein was found in the pellet, suggesting that His-Exo has a tendency to form insoluble inclusion bodies or that cell

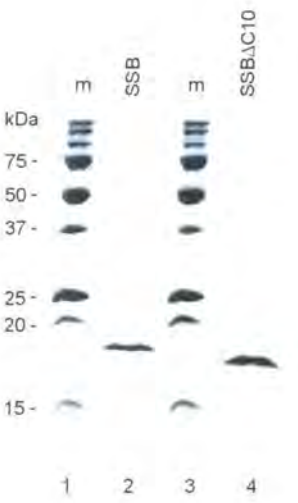

Figure 5.1. Purification of *E. coli* SSB and SSBΔC10.
15% SDS-PAGE showing the pure protein samples. Lane 1 and 3, molecular weight markers; Lane 2, Pure SSB at 3.2 mg/ml; Lane 4, Pure SSBΔC10 at 4.4 mg/ml. SSB protein migrates at its predicted molecular weight of 18.97 kDa. The difference in mobility between SSB and the C-terminal mutant SSBΔC10 is evident.

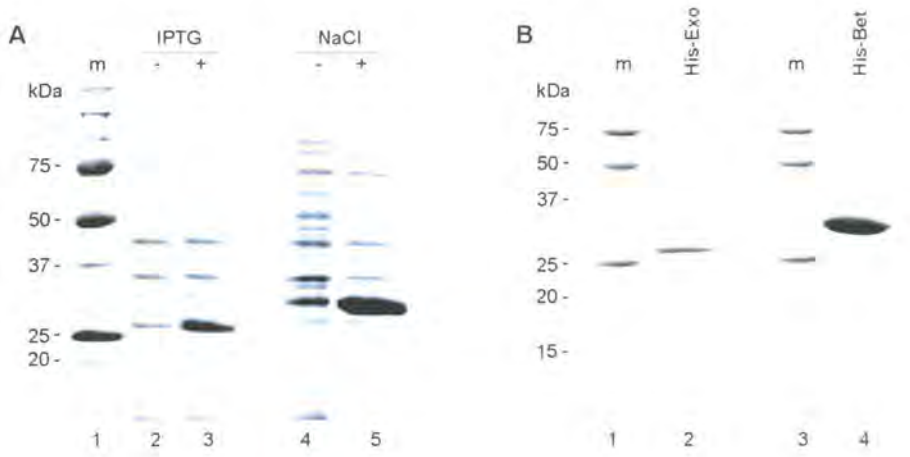

Figure 5.2. Overexpression and purification of His-tagged λ Exo and Beta proteins.
(A) 15 % SDS-PAGE showing the overexpression of His-Exo and His-Beta in BL21 (DE3) pLysS and BL21-SI competent cells respectively. Lane 1, molecular weight markers; Lane 2, uninduced BL21 (DE3) pLysS cell lysate; Lane 3, ITPG-induced BL21 (DE3) pLysS cell lysate, overexpression of His-Exo (26 kDa) can be observed; Lane 4, uninduced BL21-SI cell lysate; Lane 5, NaCl-induced BL21-SI cell lysate, overexpression of His-Beta (30 kDa) can be seen. **(B)** 15% SDS-PAGE showing the pure proteins. Lane 1 and 3, molecular weight markers; Lane 2, pure His-Exo at 0.74 mg/ml; Lane 4, pure His-Beta at 1.73 mg/ml.

disruption had been ineffective. However, there was sufficient Exo protein in the supernatant to proceed with the purification protocol (data not shown). The cleared lysate was incubated with 400 µl of pre-equilibrated Ni^{2+} agarose slurry for 15 minutes at 4°C (as described in section 4.3.3.2) prior to pouring of the mixture into a column. The column was treated with 3 ml of wash buffer prior to elution of the His-Exo with elute buffer. The high concentration (250 mM) of imidazole in this buffer caused the protein to disassociate from the matrix in three 500 µl fractions. The protein appeared pure at this stage as judged by SDS gel electrophoresis (Figure 5.2), and was dialysed against 500 mM KCl Buffer A + 50% (v/v) glycerol and stored at -80°C. The recovered protein was concentrated using Centricon concentrators and 1 ml of pure His-Exo was recovered at a final concentration of 0.74 mg/ml.

5.4 Overexpression and purification of λ β protein

The overexpression construct, pFC149 (λ *bet* in pET14b) was made by PCR amplification of the λ *bet* gene from λ genomic DNA using oligonucleotides (Bet-1 and Bet-2) to introduce NdeI and BamH1I restriction sites essentially as described above (Section 5.3). This clone was generated by Dr F.A. Curtis.

BL21-SI competent cells were transformed with pFC149 on media containing ampicillin and overnight cultures incubated at 30°C for 16 hours. Fresh LBON (1 litre) containing 150 µg/ml ampicillin was inoculated from the overnight cultures and grown to A_{650nm} 0.5 at 30°C. Target gene expression was induced by addition of sterile NaCl solution to a final concentration of 300 mM. Large quantities of His-β were produced under these conditions and clearly visible as a 30 kDa band on SDS-PAGE (Figure 5.2). Cells were harvested as previously described (Section 5.2), resuspended in 10 ml of lysis buffer and disrupted by sonication (Section 2.3.10ii).

Approximately equal amounts of His-β protein were observed in the pellet and supernatant.

The supernatant was applied to 1 ml of pre-prepared Ni^{2+} agarose and the column prepared and washed as described for His-Exo (Section 5.3). His-β was eluted from the column in 1 ml fractions under an increasing imidazole concentration gradient. The majority of the His-β protein eluted from the column between 250 mM and 500 mM imidazole. Fractions were pooled and dialysed sequentially with Buffer A supplemented with (i) 500 mM KCl and 500 mM imidazole, (ii) 400 mM KCl and 300 mM imidazole, (iii) 300 mM KCl and 100 mM imidazole, (iv) 200 mM KCl, in order to prevent protein precipitation. Minor contaminants were present and so the His-β sample was applied to a 1 ml Q-sepharose anion exchange column (pre-equilibrated with 200 mM KCl Buffer A). Bound protein was washed with 10 ml wash buffer and eluted in 500 µl fractions on a step-wise KCl gradient (0.2 M-1.0 M, in 50 mM increments). His-β eluted from the Q-sepharose matrix between 350 mM and 600 mM KCl and appeared >99% pure on SDS-PAGE (Figure 5.2). Purified protein was stored in aliquots at -80°C in 500 mM KCl Buffer A + 50% (v/v) glycerol. A total of 2.6 mg His-β was obtained at a concentration of 1.73 mg/ml.

5.5 Discussion

To investigate possible interactions between λ Orf and other λ and host recombination proteins *in vitro*, we obtained purified samples of the relevant protein partners. *E. coli* SSB and SSBΔC10 were successfully purified in milligram quantities using a published protocol (Cadman and McGlynn, 2004). In addition, N-terminal-$His_{(6)}$ fusions of λ Exo and β proteins were purified using standard protocols

devised for histidine-tagged proteins (Novagen). Untagged versions of these proteins had previously been purified using FPLC chromotography (Kovall and Matthews, 1997; Rybalchenko *et al.*, 2004; Subramanian *et al.*, 2003; Karakousis *et al.*, 1998). We discovered that it was possible to obtain purified His-Exo in a single step by affinity chromatography using Ni^{2+} agarose. Purification of His-β required the inclusion of a Q-sepharose anion exchange column, after the Ni^{2+} agarose step, to obtain a sample free of contaminants. The His-Exo recovered was at a relatively low concentration and had time permitted overexpression and purification would have been repeated to obtain a more concentrated sample. Analysis of protein: protein interactions between λ and *E. coli* proteins purified here is documented in Chapter 7.

Chapter 6

DNA binding properties of Orf proteins

6.1 Introduction

During the purification of λ Orf protein, both the wild type and His-tagged versions of the protein were found to bind heparin-agarose (Chapter 3). Heparin is a highly sulphated glycosaminoglycan, mimicking the polyanionic structure of nucleic acid and can therefore be used as a ligand for purifying DNA binding proteins. The affinity of Orf for heparin suggested that it may bind DNA; this was substantiated by its association with ssDNA cellulose. An interaction with DNA would not be unexpected as Orf substitutes for *E. coli* RecFOR, which facilitate loading of RecA onto gapped DNA (Morimatsu and Kowalczykowski, 2003), and genetic evidence indicating that λ Orf influences the initial phase of genetic exchange (Poteete, 2004; Sawitzke and Stahl, 1992; Sawitzke and Stahl, 1994; Kowalczykowski, 2000; Tarkowski *et al.*, 2002). RecFOR proteins display a variety of DNA binding activities *in vitro*. *E. coli* RecR does not appear to bind either ssDNA or dsDNA (Umezu and Kolodner, 1994), however, the RecR protein from *Bacillus subtilis* interacts with both (Alonso, 1993). RecF binds linear ssDNA independently of ATP (Griffin and Kolodner, 1990; Madiraju and Clark, 1991), and forms stable complexes with dsDNA in the presence of non-hydrolysable ATP (ATPγS) or RecR (Madiraju and Clark, 1992; Webb *et al.*, 1995; Webb *et al.*, 1999). The RecF protein shows a marked preference for binding to gapped DNA substrates in the presence of ATP, and it loads at dsDNA-ssDNA junctions (Hegde et al., 1996b). *E. coli* RecO can bind

both ssDNA and dsDNA in the presence or absence of ATP (Luisi-DeLuca and Kolodner, 1994). RecO also binds to SSB-coated ssDNA and can anneal complementary ssDNA strands (Kantake *et al.*, 2002; Luisi-DeLuca, 1995).

The affinity of Orf for heparin and ssDNA cellulose matrices and its ability to replace *E. coli* RecFOR in recombination, prompted us to probe its ability to bind DNA. ^{32}P-labelled DNA substrates designed to mimic the early intermediates of genetic recombination were assembled and binding, annealing and cleavage assays were conducted in the presence of Orf. Similar assays were performed using the Orf-like proteins, Orf151 and ETA20, to examine their functional relatedness to λ Orf. The analyses performed and results obtained form the basis of this chapter.

6.2 λ Orf binds preferentially to ssDNA

Electrophoretic mobility shift assays (EMSAs) were used to assess the ability of Orf to bind different DNA substrates. A 50-base oligonucleotide NH7 and a 25-nt ssDNA were 5'-end labelled with ^{32}P-ATP as described in Chapter 2. A 50 bp duplex was assembled by annealing labelled NH7 to its complement NH11. A flayed duplex (fork substrate, fdDNA) containing a 30 bp duplex with 20 nt ssDNA arms was generated by annealing NH7 and NH8. Annealed DNA substrates were purified by separation on 10% polyacrylamide gels in TBE buffer. These three substrates represent some of the intermediates likely to be encountered during phage recombination or replication *in vivo*.

Initial assays were performed with 0.3 nM ^{32}P-labelled DNA substrates in binding buffer with purified wt-Orf and His-Orf at 10, 100 and 1000 nM. Binding mixtures were incubated on ice for 15 min before loading on 4% PAGE in LIS buffer

and separation by electrophoresis as described in Chapter 2 (section 2.8.3iii). Gels were dried onto filter paper and analysed by autoradiography. Orf proteins efficiently retarded the migration of both ssDNA and fork substrates, but failed to form stable complexes with the duplex (Figure 6.1). The N-His_6 fusion and wild-type proteins showed very similar DNA binding profiles on these three substrates (Figure 6.1A and B) as did the MBP-Orf fusion (data not shown). The presence of either of the fusion tags did not appear to adversely affect the ability of Orf to bind DNA. The proteins also readily bound to the 25nt ssDNA substrate with affinities similar to those seen with the 50-nt ssDNA (data not shown). The band smearing observed when the protein was incubated with dsDNA suggests that Orf forms unstable complexes which dissociate during passage through the gel (Figure 6.1A and B, lanes 5-8). A similar binding pattern was observed with ssDNA and forked substrates, indicating a preference for loading onto ssDNA; with Orf protein binding the 20-nt ssDNA arms of the flayed-duplex rather than the dsDNA region (Figure 6.1A and B, lanes 9-12).

The gel shift assays were repeated with His-Orf using a broader range of protein concentrations to see if a clear preference for one of these substrates could be distinguished. The reactions were performed as described above and the dried gels analysed by phosphorimaging, using Fuji Image Gauge software to quantify DNA binding. At low protein concentrations His-Orf binds well to both ssDNA and fork substrates (Figure 6.1C). Similar binding profiles were obtained and almost all of the substrate was complexed with protein at 1000 nM His-Orf (Figure 6.1C). In contrast, His-Orf bound poorly to duplex DNA (Figure 6.1C); even at the highest protein

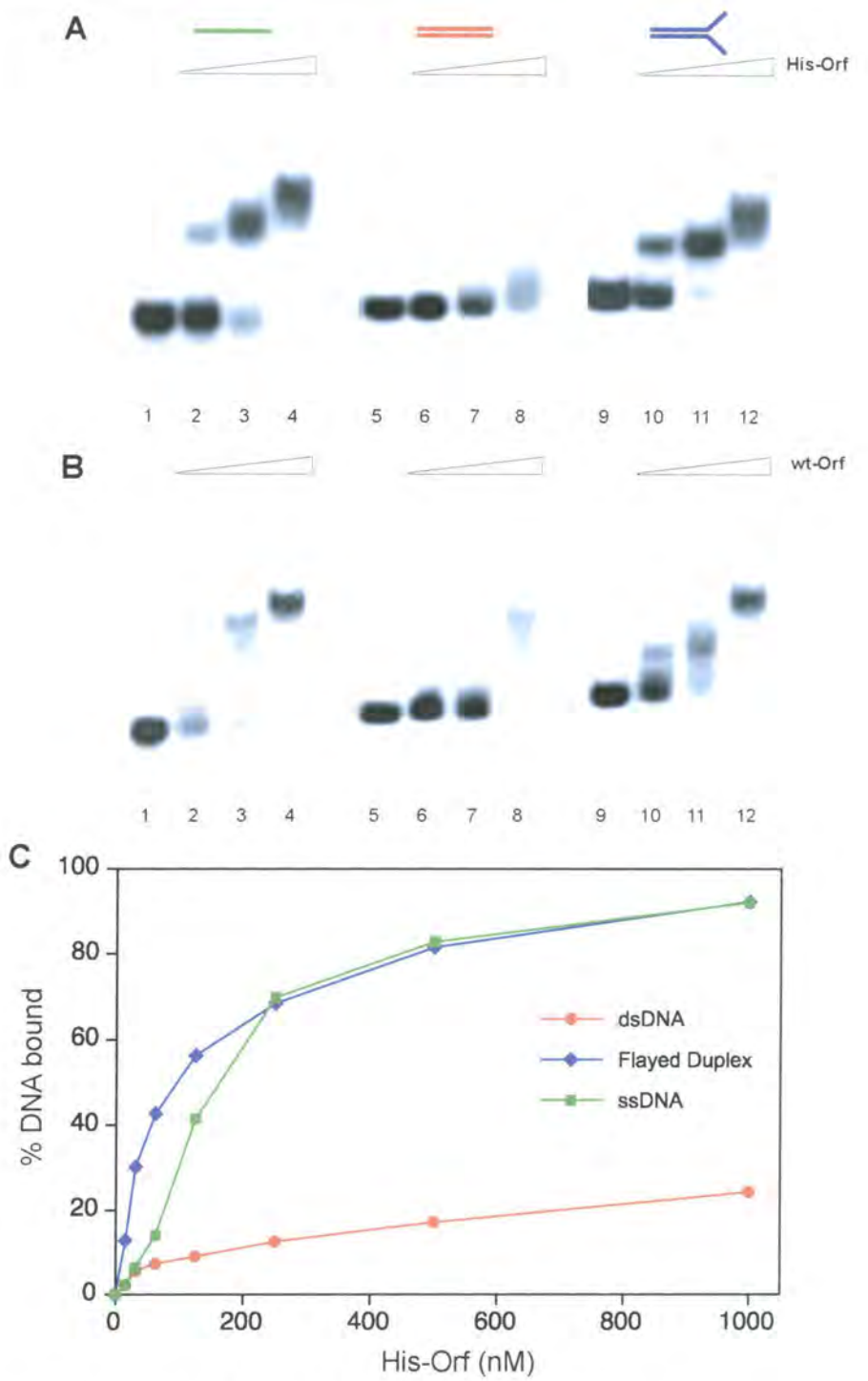

Figure 6.1. His-Orf and wt-Orf binding to DNA in gel retardation assays.
(A) Binding mixtures contained 0.3 mM ^{32}P-labelled ssDNA (lanes 1-4), dsDNA (lanes 5-8) or flayed duplex (fd) DNA (lanes 9-12) with 0, 10, 100, and 1000 nM His-Orf protein. **(B)** Binding mixtures contained 0.3 nM ^{32}P-labelled ssDNA, dsDNA or fdDNA as in **(A)** with 0, 10, 100, and 1000 nM wild-type Orf. **(C)** Quantification of DNA binding by His-Orf protein was performed with band shift assays containing 0.3 nM ^{32}P-labelled ssDNA, dsDNA and fdDNA with 0,15.625, 32.25, 62.5, 125, 250, 500 and 1000 nM protein. Data are the mean of two independent experiments. Lines are coloured according to the substrates represented in **(A)**: ssDNA (green), dsDNA (red) and fdDNA (blue).

concentration it associated with only about one-fifth of the substrate and the complexes detected were unstable (see band smearing in Figure 6.1A, lane 8).

Stopped-flow fluorescence spectroscopy was used to further probe the interaction between Orf and DNA, using an Applied Photophysics SX.18 mV stopped flow instrument. This method can be used to evaluate the interactions between proteins and their substrates by measuring the change in fluorescence of tryptophan residues (if they are situated close to the ligand binding site). Assays performed by GJ Sharples, using the purified proteins obtained during this study, revealed a quench in protein fluorescence when Orf was mixed with DNA (Maxwell *et al.*, 2005). Orf contains 7 tryptophans and one or more of these residues must be occluded when DNA is bound by the protein. Using the substrates described above, a similar fluorescence quench was detected when Orf was mixed with either ssDNA or fdDNA, suggesting that the protein binds the two substrates with a similar affinity (data not shown). The percentage change in fluorescence was much less with the 50 bp dsDNA substrate, indicating the Orf does not bind well to this substrate (data not shown). This data is in accordance with that obtained in the band shift assays and supports a preference for λ Orf in binding substrates containing a ssDNA component rather than to DNA duplexes.

6.3 Defining the Orf DNA binding site

When the crystal structure of λ Orf was determined (Maxwell *et al.*, 2005), two putative DNA binding sites were identified. The protein forms an asymmetric dimer with a tapered central cavity, which could accommodate dsDNA at the wider end and allow passage of ssDNA through the central channel. The interior of this cavity is

highly positively charged and the aromatic residue, Trp50, present within this region and could act to stabilise complexes by stacking against nucleotide bases. The protein also has a U-shaped cleft which traverses the top (wide end) of the central cavity in which DNA could bind. The cleft contains numerous basic residues and Trp50 is again positioned appropriately to help stabilise protein:DNA interactions within this shallow cleft (Maxwell *et al.*, 2005). In an attempt to define the Orf DNA binding site, assessing whether ssDNA is threaded through the central cavity or across the surface cleft, we designed oligonucleotides with a central ssDNA gap, and 3'- or 5'-ssDNA overhangs. If ssDNA passes through the central channel of the dimer then Orf should be unable to bind the gapped duplex since dsDNA cannot be accommodated within the hole. The 3' and 5' tailed substrates served as controls and to investigate any preference for binding to 3' or 5' ended dsDNA:ssDNA junctions, as noted with RecF and RecFOR (Hegde et al., 1996b; Morimatsu and Kowalczykowski, 2003).

A 60-nt ssDNA substrate (FAC6) was 5'-end labelled with ^{32}P-ATP and annealed to a 60-base homologous ssDNA (FAC7) to give a duplex. The 17-nt oligonucleotides (FAC8 and FAC9) were also annealed to FAC6, producing a gapped duplex substrate (FAC6+8+9) containing a 26-nt ssDNA stretch flanked by 17-bp duplexes. The 3'-tailed ssDNA (FAC6+8) and 5'-tailed ssDNA (FAC6+9) substrates were also made at the same time. His-Orf at increasing concentrations was mixed with each of these ^{32}P- labelled DNA substrates and binding analysed on 4% PAGE. His-Orf bound the 60-mer ssDNA with a comparable affinity to that seen earlier with 50- and 25-nt ssDNA molecules (Figure 6.2A). His-Orf bound poorly to the 60 bp duplex (Figure 6.2A), consistent with our earlier findings using a 50 bp

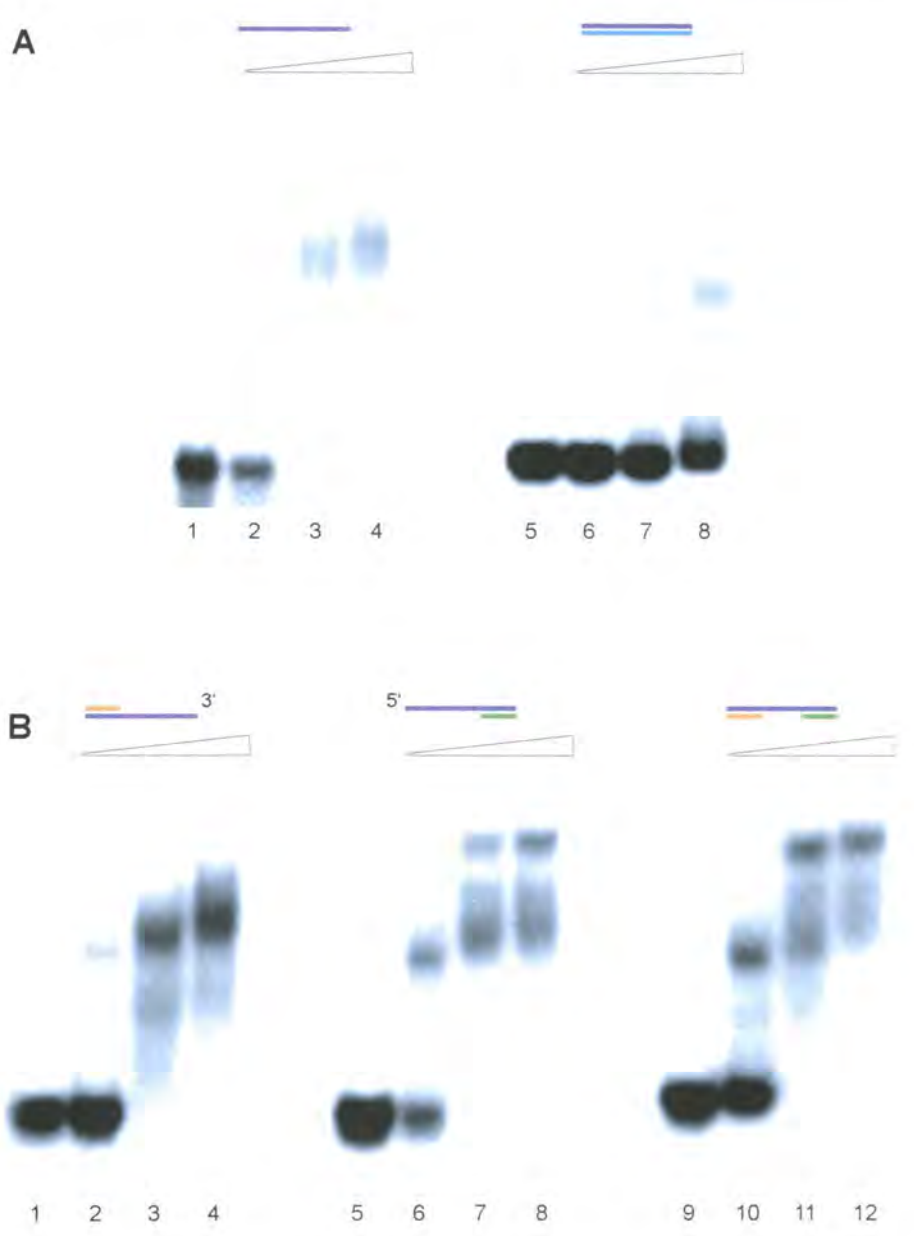

Figure 6.2. Defining the Orf DNA-binding site.
(A) Gel-shift assays of His-Orf binding to 0.3 nM ^{32}P-labelled 60-nt ssDNA (lanes 1-4) and duplex of the same sequence (lanes 5-8) with 0, 10, 100, and 1000 nM protein. **(B)** His-Orf binding to 0.3 nM ^{32}P-labelled 60-nt 3' overhang (lanes 1-4), 5' overhang (lanes 5-8) and gapped duplex (lanes 9-12) with 0, 10, 100, and 1000 nM protein. The substrates used are indicated above each panel and are coloured according to their composition; 60-nt labelled ssDNA (purple), 60-mer duplex (purple and blue), 3' overhang (purple and orange), 5' overhang (purple and green) and gapped duplex (purple, orange and green).

dsDNA. We then examined binding to the tailed and gapped duplex DNA substrates under the same conditions. His-Orf bound to all three with similar affinities, forming at least two distinct complexes with each (Figure 6.2B). The slower migrating complex was more obvious at higher protein concentrations, coinciding with a reduction in the amount of the faster-migrating complex. This is most likely due to assembly of more than one Orf dimer onto the substrate, although we cannot exclude the possibility of some conformational change within the protein:DNA complex (Figure 6.2B). Some differences in the number of the complexes and their relative gel mobilities are evident between these three substrates, although no firm conclusions can be made about what these might mean. Stopped-flow fluorescence experiments were conducted with all of these DNA substrates and the results mirrored those obtained by gel retardation (Maxwell *et al.*, 2005). In particular, the fluorescence change profiles detected with ssDNA, gapped duplex or the 5' and 3' overhangs were identical.

λ Orf protein clearly binds preferentially to substrates containing ssDNA. The fact that Orf bound equally well to either of the tailed and the gapped duplex substrates, suggests that DNA is bound within the U-shaped cleft upon the surface of the Orf dimer. However, binding to the gapped duplex could still be explained by Orf acting as a hinge protein, capable of opening to allow binding of ssDNA within the central channel. This hypothesis would be in agreement with a proposed action at dsDNA-ssDNA gaps, as seen with RecFOR, the face of the dimer where the central cavity is widest interacting with duplex DNA while ssDNA is accomodated within the channel itself. This explanation requires some mechanism for separating either of the two segments of Orf responsible for stable dimer formation. The similar binding

affinities for ssDNA and the 5' or 3' tailed suggests that Orf does not preferentially load at a particular ssDNA:dsDNA junction. Further discussion of these results can be found in Chapter 9.

6.4 His-Orf does not act as a ssDNA annealing protein

The *E. coli* RecO protein has the capacity to anneal homologous ssDNA molecules in the presence or absence of SSB protein (Luisi-DeLuca and Kolodner, 1994; Kantake *et al.*, 2002). We assayed His-Orf for strand annealing activity using 25-base and 50-base oligonucleotides matching ϕX174 circular ssDNA. Reactions contained 1, 10, 100 and 1000 nM protein with 0.1 ng/μl ^{32}P-5'-end labelled ϕX174-25 or ϕX174-50 oligonucleotides and 20 ng/μl ϕX174 ss-circular virion DNA in buffer (50 mM Tris-HCl pH 8.0, 1 mM DTT, 100 μg/ml BSA). His-Orf was mixed on ice with the labelled oligonucleotides and the virion ssDNA added last. Reactions were incubated at 37°C for 10 minutes and stopped by the addition of buffer containing proteinase K and incubation continued for 10 minutes. Suitable controls containing labelled substrate and circular ssDNA but no protein were subjected to the same incubation conditions to measure annealing in the absence of protein. An annealed sample was prepared by heating at 95°C for 2 minutes and cooling to room temperature. A third control, containing only the labelled oligonucleotide was incubated under the same conditions as protein-containing reactions. Annealing reactions were loaded onto a 10% polyacrylamide gel and run at 190 V for 1 h in TBE buffer. The gel was dried and the labelled DNA analysed by autoradiography.

His-Orf protein failed to stimulate strand annealing under these conditions, suggesting it lacks this activity (data not shown). Further experiments (performed by

GJ Sharples), with homologous 50- and 25-nucleotide substrates analysed on polyacrylamide gels or stopped-flow assays using DAPI to measure duplex formation also failed to provide any evidence that Orf can promote strand annealing (data not shown). These results were not too surprising as in the Red recombination system of phage λ, β protein supplies a single-stranded DNA annealing function (Kmiec and Holloman, 1981; Muniyappa and Radding, 1986).

6.5 λ Orf does not display nuclease activity

Potential nuclease activity of His-Orf was assayed using the same three substrates utilised in the initial DNA binding studies. The 50-nt ssDNA, dsDNA and forked substrates at 0.3 nM were mixed with 10, 100 and 1000 nM His-Orf in cleavage buffer (50 mM Tris-HCl pH 8.0, 1 mM DTT, 100 μg/ml BSA, 1 mM $MgCl_2$). Reactions were incubated at 37°C for 1 h before inactivation of the protein by addition of proteinase K. Samples were incubated for a further 10 minutes at 37°C before separation of DNA on 10% PAGE in TBE buffer at 190 V for 2 h. Gels were dried and analysed by autoradiography as described previously.

No degradation of the DNA was detected under these conditions, suggesting λ Orf protein does not function as a nuclease.

6.6 *E. coli* DLP12 Orf151 preferentially binds ssDNA

The DNA binding experiments performed with Orf were repeated with Orf151 to determine whether the *E. coli* protein is functionally analogous to the λ recombinase. Initial binding experiments were conducted with MBP-Orf151 using 50-nt, ssDNA, linear duplex (dsDNA) and flayed duplex (fdDNA) substrates. The protein bound all

three substrates (Figure 6.3A), however, slightly higher protein concentrations than those employed with His-Orf were needed. MBP-Orf151 formed two discrete complexes with the 50-nt ssDNA substrate. A similar binding affinity for 25-nt and 60-mer ssDNA substrates was observed, correlating with the results obtained with His-Orf on these substrates (data not shown). Unusually, the faster-migrating complex was more evident at higher protein concentrations, potentially indicating that the protein undergoes a conformational change or that the ssDNA can adopt an alternative conformation within the protein complex. Another possibility is that the appearance of this smaller complex at higher protein concentrations is a consequence of electrophoresis, with a pair of MBP-Orf151 dimers being displaced from the main complex resulting in the two species seen (Figure 6.3A). MBP-Orf151 formed a single stable complex with a proportion of the dsDNA and appeared to bind poorly to the forked substrate (Figure 6.3A). However, MBP-Orf151 formed two discrete complexes on fdDNA, the faster-migrating complex again being more evident at higher protein concentrations.

A larger range of protein concentrations was used to analyse the preferred substrates and facilitate a comparison with Orf DNA binding activity (Figure 6.3B). MBP-Orf151 displayed a distinct preference for the ssDNA substrate relative to either dsDNA or forked DNA (Figure 6.3B). At 1000 nM, the protein bound over 50% of the ssDNA, while only 30% dsDNA and 20% fdDNA were bound by MBP-Orf151 at the same concentration. MBP-Orf151 appears to have a lower affinity for ssDNA relative to Orf, which bound over 90% of the substrate at this protein concentration (Figure 6.1B). The *E. coli* protein displays a slightly higher affinity for dsDNA based on the percentage DNA bound; it also formed more stable complexes

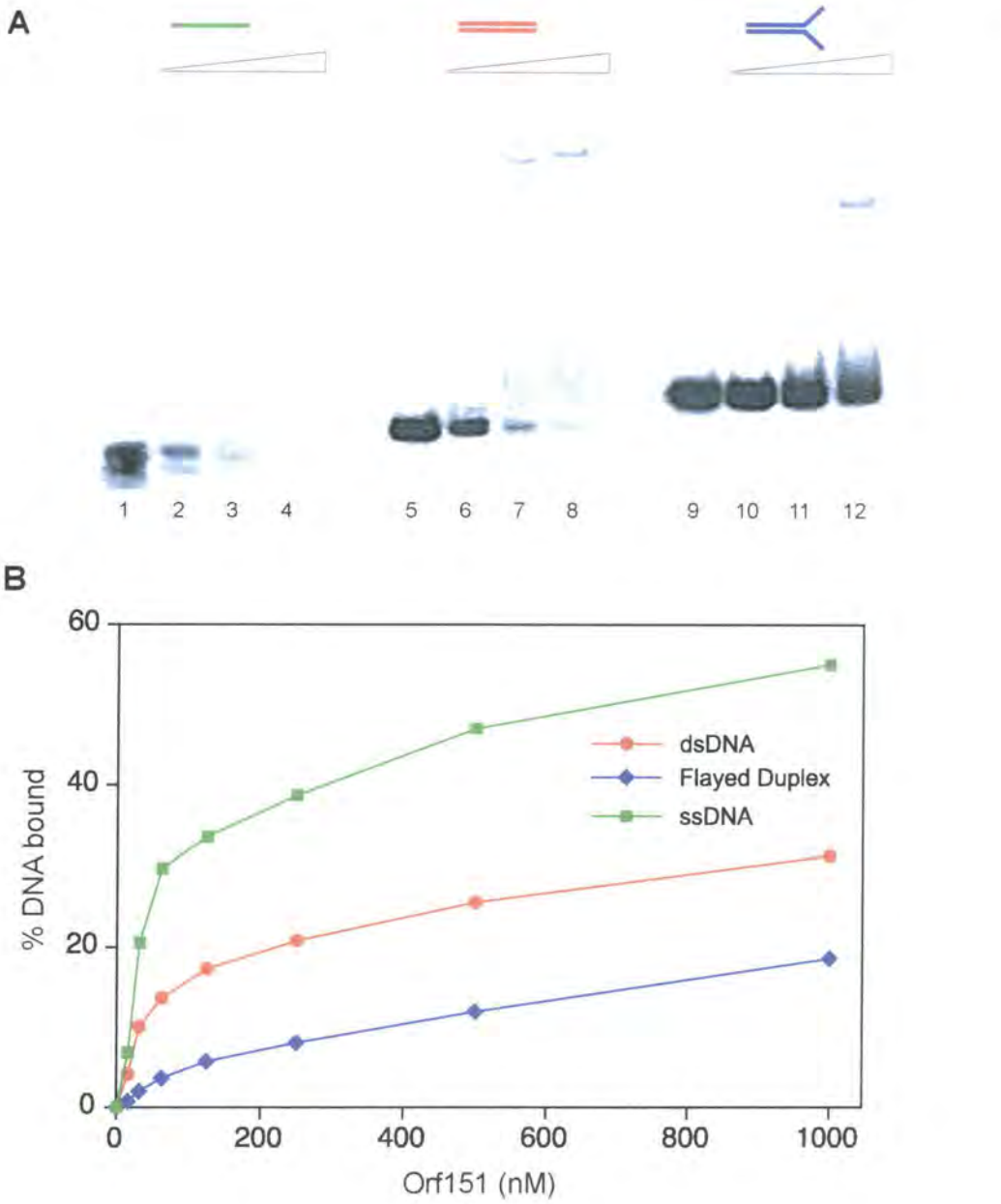

Figure 6.3. MBP-Orf binding to DNA in band shift assays.
(A) Binding reactions contained 0.3 nM ^{32}P-labelled ssDNA (green; lanes 1-4), dsDNA (red; lanes 5-8) and fdDNA (blue; lanes 9-12) with 0, 20, 200, and 2000 nM MBP-Orf151 protein. **(B)** Quantification of DNA binding by MBP-Orf protein was performed with band shift assays containing 0.3 nM ^{32}P-labelled DNA substrates with 0, 15.625, 32.25, 62.5, 125, 250, 500, and 1000 nM MBP-Orf151. Data are the mean of two independent experiments. DNA substrates are coloured as in **(A)**: green (ssDNA), red (dsDNA) and blue (fdDNA).

with the substrate in comparison to the unstable complexes formed by Orf (see band smearing in Figure 6.1A and B). In contrast to Orf, which bound fdDNA and ssDNA with similar affinities, MBP-Orf151 bound poorly to the forked substrate (Figure 6.1). It is possible that MBP-Orf151 requires ssDNA arms of longer than 20 nt to assemble on this substrate, or that binding is hindered by the branch-point.

We also investigated MBP-Orf151 binding to the gapped duplex and tailed substrates described in Section 6.3. MBP-Orf151, like Orf, appeared to bind equally well to all three DNA substrates, although a slight preference for 3'-tailed DNA was noted (Figure 6.4). Three distinct complexes were observed with the 3'-tailed and gapped duplex substrates, the fastest migrating of these being more obvious at higher protein concentrations. In comparison, only two distinct complexes were seen with the 5'-tailed substrate, corresponding to the two faster-migrating complexes, and less DNA was retarded. This may indicate that the protein recognises 3'-ssDNA ends more efficiently than 5'-ssDNA ends. Further analysis is required to confirm the reproducibility of complex formation and determine whether MBP-Orf151 has a specific preference for certain DNA substrates.

MBP-Orf151 failed to cleave ssDNA, dsDNA, or fdDNA substrates when assayed for nuclease activity in a similar manner to Orf (Section 6.5). The protein also failed to display any annealing activity with homologous oligonucleotides or those matching ϕX174 circular ssDNA (data not shown).

Although there were subtle differences in the way Orf and Orf151 bound the DNA substrates analysed, these studies clearly show that both bind DNA species with a significant preference for ssDNA. Neither shows a particular affinity for branched DNA, nor do they display any nuclease or strand-annealing activity. These

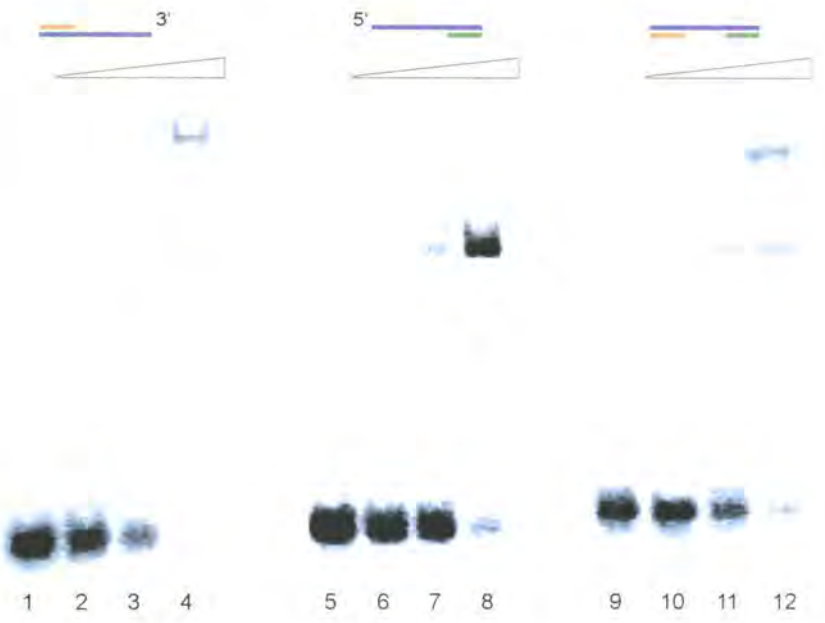

Figure 6.4. MBP-Orf151 binding to gapped and tailed DNA substrates.
Gel mobility shift assays were performed with 0.3 nM ^{32}P-labelled 60-nt 3' overhang (lanes 1-4), 5' overhang (lanes 5-8) and gapped duplex (lanes 9-12) substrates with 0, 10, 100, and 1000 nM MBP-Orf151. The substrates used are indicated above each panel and are coloured according to their composition: purple (^{32}P-5'-end labelled 60-nt ssDNA), orange (17-nt strand complementary to the 5'-end of labelled ssDNA), and green (17-nt strand homologous to the 3'-end of labelled ssDNA).

functional similarities help to validate the relationship between Orf and Orf151 identified by primary and secondary structure analysis, though further investigation is warranted.

6.7 Potential nuclease activity of *S. aureus* ϕETA Orf20

Electrophoretic mobility shift assays using the same range of ssDNA, dsDNA, gapped and branched substrates were repeated with the putative λ Orf homolog from *S. aureus* phage ϕETA. The ETA20 protein failed to bind any of the substrates under these experimental conditions (data not shown). The protein did, however, cleave 50-nt ssDNA, dsDNA and fdDNA at high concentrations in the presence of 1 mM $MgCl_2$ (Figure 6.5A). No significant preference for single-strand, duplex or forked DNA was observed, with the protein cleaving approximately 80% of all three substrates at 1000 nM concentration (Figure 6.5B). The protein encodes an additional C-terminal extension, absent from Orf and Orf151, which incorporates a zinc-finger motif belonging to the HNH family nuclease domain (Mehta *et al.*, 2004). The presence of this motif fits with the nuclease activity identified here. However, further investigation is required to ensure that the observed activity, only evident at high protein concentrations, is not due to a contaminating *E. coli* nuclease. Comparisons with the ETA20ΔC82 protein, which is missing the HNH nuclease domain, may help in this regard.

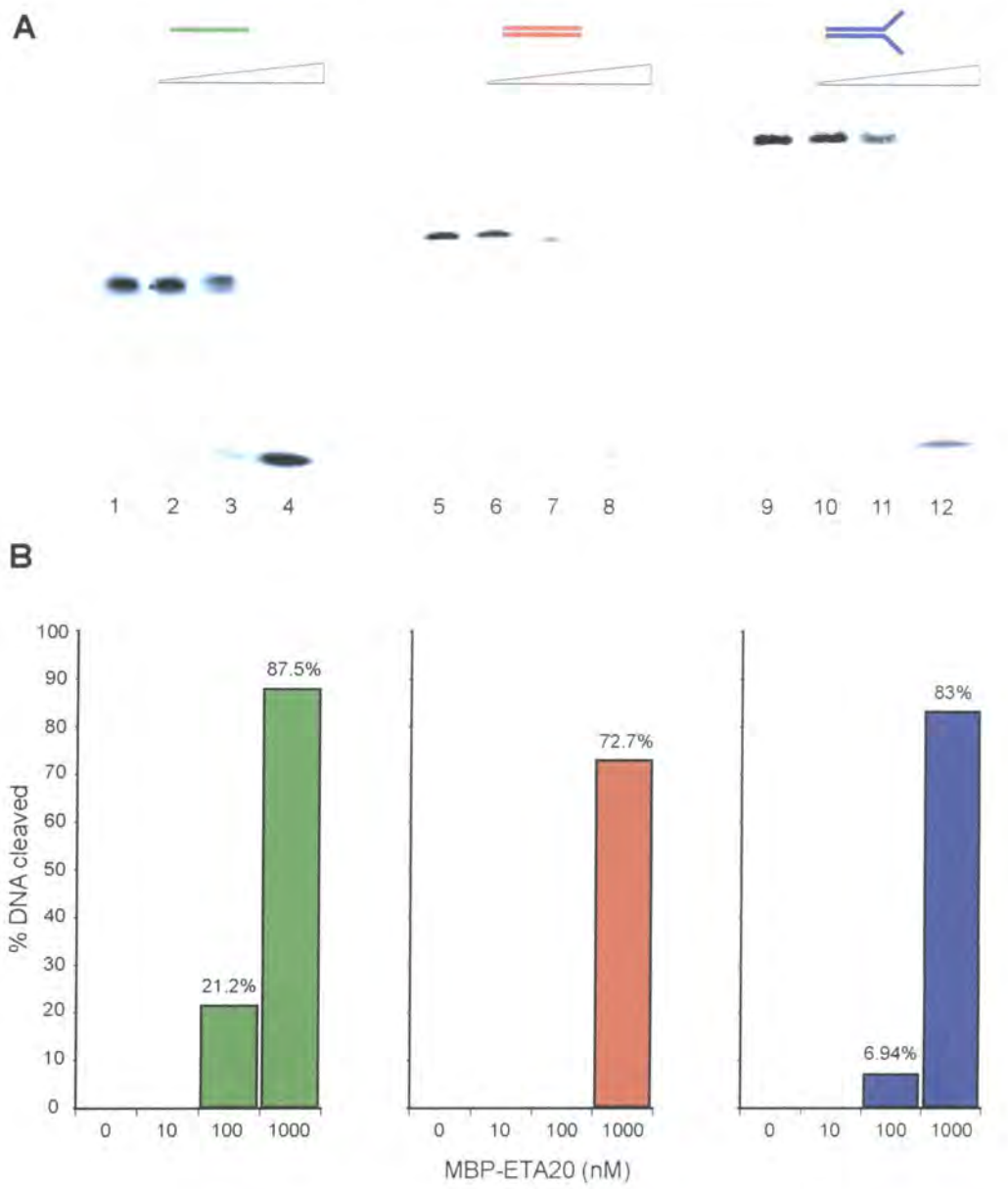

Figure 6.5. MBP-ETA20 nuclease activity.
(A) Reactions contained 0.3 nM ^{32}P-labelled ssDNA (lanes 1-4), dsDNA (lanes 5-8) and flayed duplex (lanes 9-12) with 0, 10, 100, and 1000 nM MBP-ETA20 protein. **(B)** The percentage DNA cleaved was calculated from the mean of two independent experiments. Bars charts are aligned with the appropriate gel data. Substrates are indicated by colour: green (ssDNA), red (dsDNA) and blue (fdDNA).

6.8 Discussion

In this chapter we investigated the ability of λ Orf and two putative orthologs, Orf-151 and ETA20, to interact with a variety of DNA species, designed to mimic the intermediates likely to be encountered during the early stages of genetic exchange. The Orf and Orf151 proteins showed a preference for binding to ssDNA over duplex DNA, and both could form complexes with small (25nt) ssDNA species. Subtle differences in their binding properties were noted, for instance Orf formed a single complex on ssDNA while Orf151 produced two distinct complexes. λ Orf showed limited binding to dsDNA and an affinity for the forked substrate comparable to that seen with ssDNA. In contrast, Orf151 could form stable complexes with duplex DNA, unlike Orf, but bound only poorly to the forked substrate. Both proteins showed generally similar affinities for 3'- and 5'-tailed DNA and a 26-nt ssDNA flanked by 17-nt duplexes. Neither protein showed any degradative activity on DNA or the capacity to anneal homologous sequences, suggesting they do not possess these activities.

The ETA20 protein did not display any DNA binding activity in these assays, but did show nuclease activity upon ssDNA, dsDNA and fdDNA. However, the possibility that this was due to a contaminating nuclease has not been eliminated. The lack of DNA binding activity with ETA20, coupled with its limited primary and secondary sequence similarity with λ Orf, may suggest that the *S. aureus* phage protein is not a genuine member of the Orf family.

The similarities between the DNA binding properties of λ Orf and *E. coli* DLP12 prophage Orf151 are, however, suggestive of a functional relationship. These studies provide the first evidence that the proposed sequence similarities identified in

Chapter 3 are authentic. Furthermore, they strengthen the notion that Orf family recombinases, conserved across a diverse number of lambdoid phages, may all function in a similar manner in the early stages of recombination.

Chapter 7

Interactions between Orf and other recombinases

7.1 Introduction

The *E. coli* SSB protein is directly involved in recombination, disrupting the secondary structure of ssDNA and protecting the labile strand from nucleases (Meyer and Laine, 1990). By disrupting the complicated secondary structure, SSB makes the ssDNA accessible to recombination and replication enzymes. For example *in vitro*, SSB can stimulate strand-exchange reactions promoted by RecA, but this interaction is dependent upon the concentration and order of addition of RecA and SSB (Meyer and Laine, 1990; Cox and Lehman, 1987; Muniyappa *et al.*, 1984). *In vivo*, SSB has been shown to compete with RecA and other proteins for binding to ssDNA (Tsang *et al.*, 1985; Menetski and Kowalczykowski, 1985; Kowalczykowski *et al.*, 1987; Kowalczykowski and Krupp, 1987). The rate limiting step in the formation of the RecA presynaptic filament is the initial binding of RecA to ssDNA to form a nucleation site, which once formed, allows RecA to cooperatively polymerise onto ssDNA and displace the bound SSB (Thresher *et al.*, 1988). As previously mentioned, the *E. coli* RecFOR complex counters this inhibitory effect of SSB, by loading RecA specifically at the 5' end of gapped duplex DNA, thereby accelerating DNA strand exchange (Morimatsu and Kowalczykowski, 2003). In brief, the RecO protein interacts directly with SSB *in vitro* (Umezu and Kolodner, 1994) and can load RecA protein onto SSB-coated ssDNA (Umezu *et al.*, 1993; Bork *et al.*, 2001).

The RecOR complex assists RecA to utilise the SSB-ssDNA complex in joint molecule formation assays and forms a strong physical interaction with SSB *in vitro*, although no direct interaction with RecA has been demonstrated (Umezu *et al.*, 1993; Umezu and Kolodner, 1994). The RecR and RecF proteins do not interact directly with SSB, but the RecFOR complex as a whole can form interactions with SSB, mediated by RecO (Hegde et al., 1996a).

As previously discussed in this work, Orf is required for λ recombination in the absence of any of the three *E. coli* recombination proteins RecF, RecO, and RecR that assist in RecA-mediated strand exchange reactions (Sawitzke and Stahl, 1992; Sawitzke and Stahl, 1994; Kowalczykowski *et al.*, 1994; Webb *et al.*, 1997; Poteete, 2004). The ability to replace the RecFOR function in λ recombination suggests that this protein functions in the loading or unloading of β protein, RecA or SSB at the ssDNA:dsDNA interface. Initial investigation of potential protein:protein interactions, performed by MD Watson and D Hart, employed the Matchmaker III yeast-two hybrid system to identify enzymes that work together in driving recombination reactions. A number of likely partners were screened and potential interactions between Orf, β and SSB, and Orf151, Exo and SSB were uncovered (Table 7.1). In this chapter we describe attempts to authenticate the yeast-two hybrid data, using purified proteins (Chapters 4 and 5) in a number of biochemical assays. The ability of Orf, or Orf151, to interact directly with Exo, β and SSB was examined by far-western blotting, affinity chromatography and size exclusion chromatography. The assays performed and results obtained are described in this chapter and discussed in more depth in Chapter 9.

	RecA	SSB	Exo	Beta	Orf	Orf151
RecA	++++					
SSB		++++			++	
Exo			++++	++		++
Beta			++	++		
Orf		++		++	++++	
Orf151		++++	++			

Table 7.1. Yeast two-hybrid analysis of *E. coli* and λ proteins implicated in genetic recombination.
Experiments were conducted by MD Watson and D Hart and are based on the results from SD and α-gal plates using pGADT7 and pGBKT7 constructs containing the relevant genes. Strong (++++) or weak (++) interactions are indicated.

7.2 Far-western analysis confirms an interaction with SSB

Western blot analysis involves the use of an antibody to detect the presence of a particular protein bound to a suitable membrane (e.g. PVDF). In a far western (overlay assay), a second protein is added, frequently a recombinant fusion protein containing a suitable tag. Interactions between the fusion protein and the target protein bound to the membrane can be detected using antibodies to the fusion-tag (Sambrook, 2001). Biotinylation or radioactive labelling of the fusion protein can also be used for detection. The method was originally developed to screen protein expression libraries with a ^{32}P-labelled GST-fusion protein (Blackwood and Eisenman, 1991; Kaelin *et al.*, 1992). However, it can also be used for membranes generated by transfer of proteins following SDS-PAGE, and with any number of fusion tags (e.g. GST, MBP or His).

Purified prey proteins (SSB, SSBΔC10, His-Exo and His-β) at 2.5 μg were separated on 15% SDS-PAGE with pre-stained molecular weight markers and transferred to a PDVF membrane using a transblotter (BioRad). Negative (His-Orf or His-Orf151 at 2.5 μg) and positive (MBP-Orf or MBP-Orf151 at 2.5 μg) controls

were analysed in parallel. The membrane was incubated with the bait protein (MBP-Orf or MBP-Orf151 at 5 μg) in blocking buffer, to prevent non-specific binding, for 4 hours at 4°C with gentle agitation. Tris-Tween buffer was used to wash the membrane before addition of monoclonal antibodies directed against the fusion tag (see Section 2.4.6). Secondary antibody (rabbit anti-mouse-IgG peroxidase conjugate) was added and the membrane washed in Tris-Tween buffer, with frequent buffer changes, prior to detection with ECL reagents (Amersham Pharmacia) and exposure to X-ray film.

Using MBP-Orf as a bait protein, an interaction between MBP-Orf and SSB was detected, although the signal was extremely weak (data not shown). This was possibly an indication that: (i) the probe protein concentration was insufficient, (ii) an increased incubation time with the probe protein was required, or (iii) the protein interactions are not genuine. It was difficult to discern whether MBP-Orf interacted with any of the other prey proteins due to the high background signals evident on X-ray films. No interaction with SSB was detected when MBP-Orf151 was used as the bait protein (data not shown). It is possible that MBP-Orf or MBP-Orf151 do not interact with prey proteins under these conditions. Further assays are needed to determine optimal protein concentration and incubation periods before the possibility of interactions in this assay can be ruled out. Coomassie blue staining of the polyacrylamide gel used for blotting confirmed that the majority of the prey proteins had been successfully transferred to the membrane.

These assays have since been repeated by FA Curtis using MBP-Orf (20μg) to probe blotted SSB protein (0.5, 1.9 and 3.8 μg) followed by detection with anti-MBP antibodies. The results showed a clear interaction between MBP-Orf and SSB

under the same conditions (Maxwell *et al.*, 2005). The use of 20 μg of pure MBP-Orf in this assay in comparison to only 5 μg in the initial study could explain the lack of a clear signal. Subsequent studies have also confirmed that MBP-Orf151 and SSB interact in far-western blots. Although an interaction between SSB and MBP-Orf was indicated in the initial experiments, and has since been confirmed using this assay, the experiments need to be repeated with His-Exo or His-β to validate potential interactions suggested by the yeast-two hybrid data (Table 7.1).

7.3 Affinity chromatography analysis of protein:protein interactions

Interactions between the RecF, RecO and RecR proteins with SSB were previously analysed by affinity chromatography (Hegde et al., 1996a). SSB protein was coupled to Affiprep-10 matrix (BioRad) and various combinations of the RecF, RecO and RecR proteins were loaded onto the column and bound proteins eluted with a linear salt gradient.

To probe the interaction of MBP-Orf and MBP-Orf151 with SSB and SSBΔC10, we performed similar affinity chromatography experiments. MBP-tagged proteins were allowed to bind amylose resin followed by addition of purified SSB or SSBΔC10. Purified MBP-Orf (0.5 mg/ml) or MBP-Orf151 (0.5 mg/ml) was incubated with 200 μl of amylose resin at 4°C for 15 minutes before washing in column buffer and pouring of the column. SSB or SSBΔC10 (0.5 mg/ml) were loaded and the column washed to remove unbound proteins. MBP fusion proteins, together with any bound SSB protein, were eluted with column buffer containing 10 mM maltose and fractions analysed by SDS-PAGE. Control assays with purified

SSB or SSBΔC10 applied directly to the amylose resin showed that neither protein binds the matrix under these experimental conditions.

A small proportion of SSB protein co-eluted from the amylose column with the MBP-Orf and MBP-Orf151 proteins, however, the eluted SSB protein was barely visible on Coomassie blue stained gels (data not shown). Much of the SSB protein recovered failed to bind the affinity column indicating that any interaction is relatively weak. Using anti-SSB antibodies to detect the protein in western blots would help to confirm this interaction, as would repetition of the assay with higher SSB concentration, or perhaps longer incubation with the MBP-Orf-bound resin.

No binding of amylose coupled-MBP-Orf or MBP-Orf151 was observed when SSBΔC10 was applied to the column. All of the probe protein appeared in the flow through and wash fractions (data not shown). However, recent studies have shown that MBP-Orf and MBP-Orf151 can still bind SSBΔC10 in far-western blots (FA Curtis and GJ Sharples, personal communication).

Unfortunately, this assay failed to yield convincing results and further investigations are required. The use of an FPLC system to detect protein elution by UV absorption and the addition of higher concentrations of probe SSB protein may provide more reliable data. It is possible however that the Orf proteins require accessory factors to form strong interactions with SSB protein. RecF and RecR proteins require the presence of RecO in order to interact with SSB (Hegde et al., 1996a). Furthermore, many recombination mediator proteins only recognise their cognate SSB when complexed with ssDNA (Beernink and Morrical, 1999).

7.4 Gel filtration analysis of protein complex formation

Size exclusion chromatography was also employed in order to investigate interactions among the RecF, RecO, RecR and SSB proteins (Hegde et al., 1996a). To further examine the interactions between Orf, Orf151, SSB, Exo and β indicated by the yeast-two hybrid data (Table 7.1) and by far-western blotting, we analysed complex formation by size-exclusion chromatography. A 24 ml Superose-6HR gel filtration column was equilibrated in 250 mM KCl Buffer A. A relatively high salt concentration was used to help alleviate any potential problems with Orf/Orf151 solubility (as discussed in Chapter 4). Elution profiles were analysed with respect to known molecular weight standards exposed to the same conditions (Figure 7.1A). The molecular weight of each complex was calculated using the calibration curve produced by plotting $_{log}$MW against the elution volume of the standards (Figure 7.1B). Proteins (1 mg/ml unless stated otherwise) were mixed on ice and incubated for 15 minutes before loading on the column. Protein was detected by UV absorption at 260 nm and 280 nm and peak fractions collected and analysed by SDS-PAGE to confirm the presence of each protein.

7.4.1 Orf and Orf151 fail to associate with SSB

MBP-Orf eluted with a relative molecular weight of 117 kDa in agreement with the 119 kDa predicted for a homodimer (Figure 7.2A and Figure 4.5D) as noted previously with His-Orf (Figure 4.3). SDS-PAGE analysis of the peak fractions confirmed the presence of MBP-Orf protein (Figure 7.2A). Similarly, MBP-Orf151 eluted as an apparent dimer of 113 kDa (Figure 7.2C; predicted MW of 120 kDa) as previously described (Figure 4.7E). These results confirm that both proteins exist as

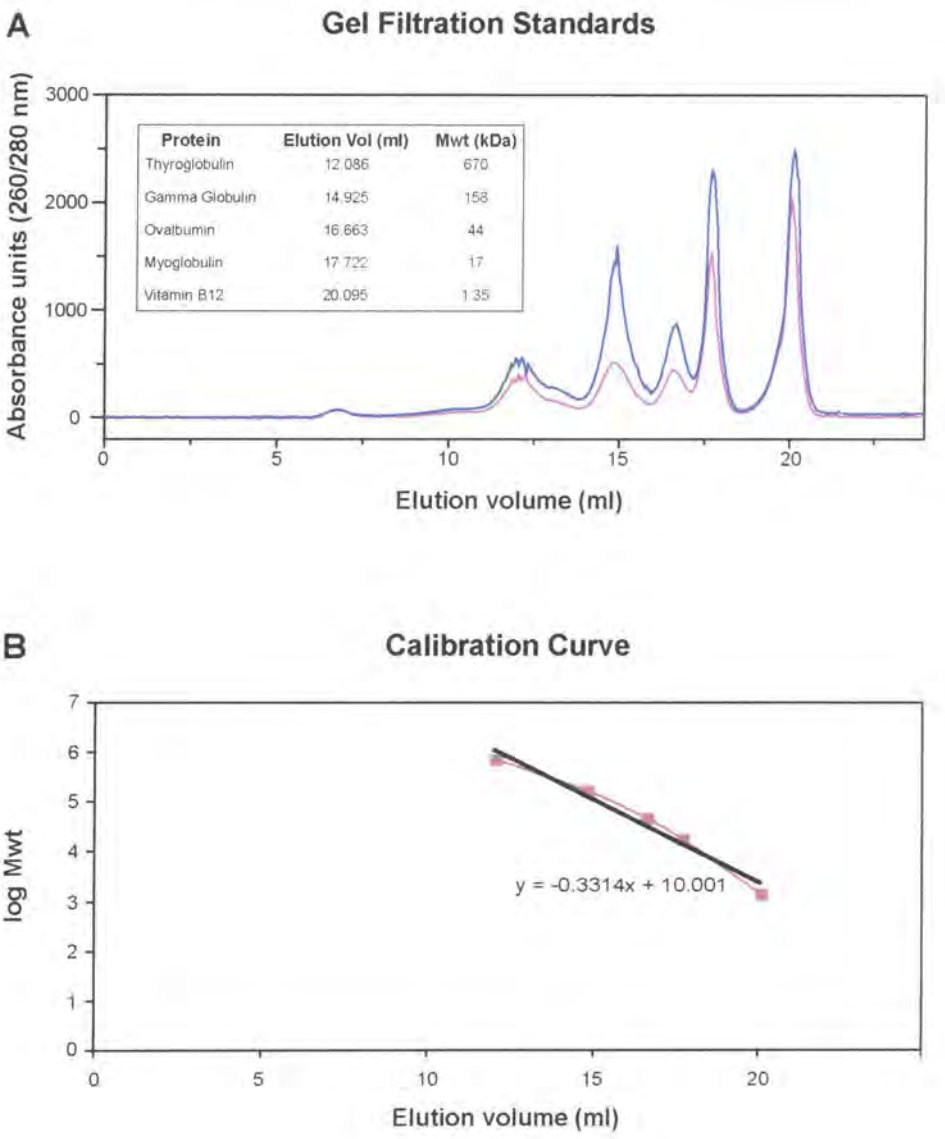

Figure 7.1. Elution profiles of the molecular weight standards and calibration curve.
(A) Chromatograph showing the elution profiles of the molecular weight standards; 0.125 ml of standards in H_2O were loaded on a Superose 6HR 10/30 column and eluted at a flow rate of 0.3 ml/min in 250 mM KCl Buffer A. **(B)** Calibration curve; elution volumes for each standard were recorded and plotted against $_{log}$MW to produce the linear plot used for calculating the mass of proteins/protein complexes on the column.
Absorption at 260 nm (pink) and 280 nm (blue) was used to detect proteins eluting from the column.

homodimers in solution, supporting the yeast-two hybrid results showing that the proteins self-associate (Table 7.1). *E. coli* SSB protein eluted as an apparent tetramer of 84 kDa (Figure 7.2E; predicted MW of 76 kDa), consistent with its known oligomeric state. When MBP-Orf and SSB were mixed and incubated on ice before application to the column, a single broad peak was observed, initially indicating the formation of an MBP-Orf-SSB complex (Figure 7.2B). However, when the peak fractions were analysed by SDS-PAGE, the proteins did not appear to co-elute; most of the MBP-Orf appeared in the early fractions with SSB eluting slightly later (Figure 7.2B). The elution volume of the single peak correlates with a molecular weight of 85 kDa appropriate to the position of the SSB tetramer. As the elution volumes of MBP-Orf (15.298 ml) and SSB (15.32 ml) are similar, the two separate profiles obscure one another resulting in the appearance of a single peak. When MBP-Orf151 and SSB were mixed (Figure 7.2D), a single broad peak with a distinct shoulder was observed. SDS-PAGE analysis of the appropriate fractions confirmed that MBP-Orf151 and SSB do not coelute (Figure 7.2D). The results suggest that neither MBP-Orf nor MBP-Orf151 is capable of binding SSB under the conditions employed.

The experiments were repeated using SSBΔC10, in place of SSB, to see if loss of the C-terminal tail affects Orf/Orf151 binding (Figure 7.3). MBP-Orf and MBP-Orf151 elution profiles are reproduced here for reference (Figure 7.3A and C). SSBΔC10 eluted at 73 kDa, in agreement with the 71 kDa predicted for a tetramer carrying a 10 residue C-terminal truncation (Figure 7.3E). When MBP-Orf and SSBΔC10 reaction mixtures were analysed, a single broad peak with a distinct shoulder was seen (Figure 7.3B). SDS-PAGE analysis of the eluted fractions

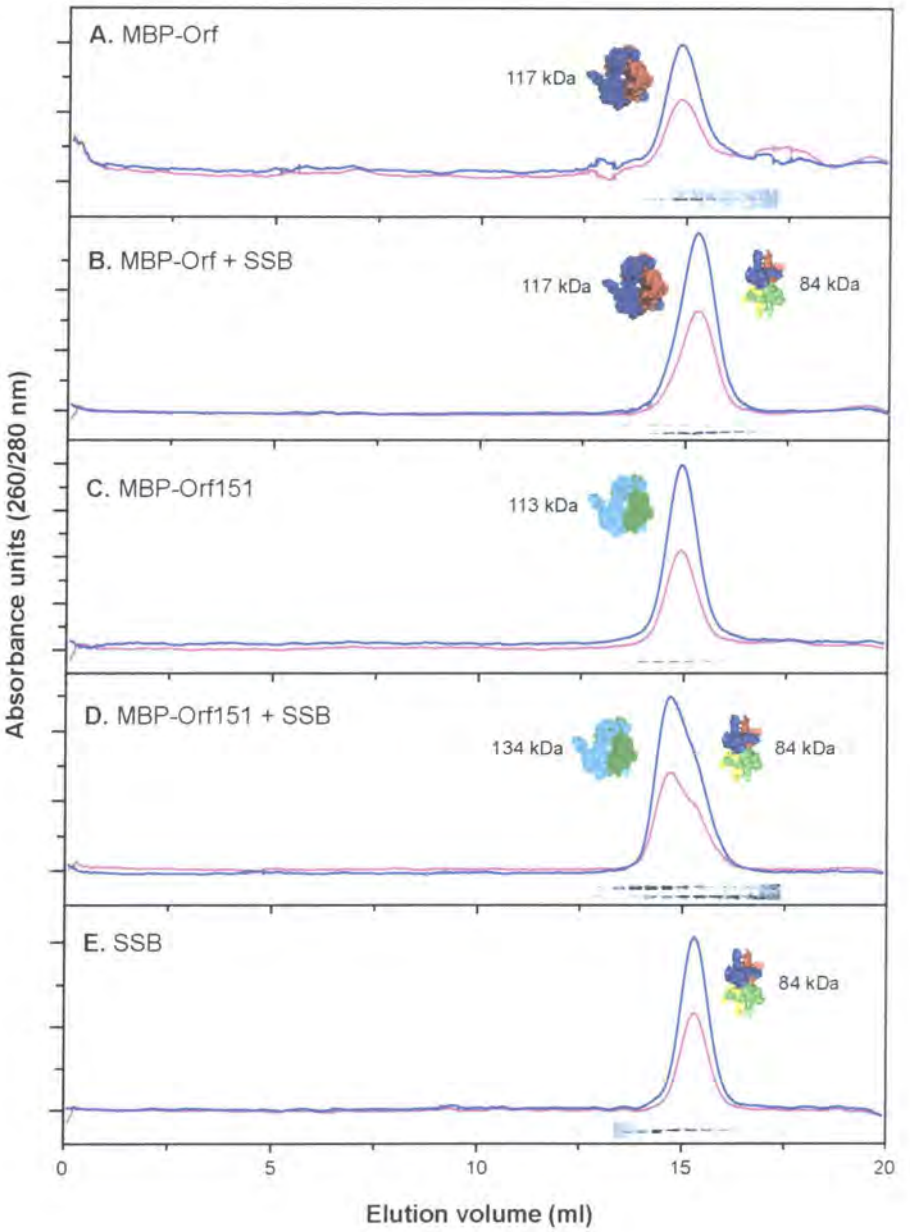

Figure 7.2. Size exclusion chromatography of MBP-Orf, MBP-Orf151 and SSB. Proteins (1 mg/ml) were loaded on a 24 ml Superose-6HR 10/30 column in 250 mM KCl Buffer A and eluted at a flow rate of 0.3 ml/min. Elution volume is shown on the x-axis and eluted proteins were collected by absorption at 260 nm (pink) and 280 nm (blue). Peak fractions were separated on 15% SDS-PAGE to confirm the presence of protein, and shown below the relevant peak in each panel. The predicted molecular weight of each protein monomer is: MBP-Orf (59 kDa), MBP-Orf151 (60 kDa) and SSB (19 kDa). **(A)** MBP-Orf eluted as an apparent dimer of 117 kDa, **(B)** MBP-Orf and SSB did not co-elute, **(C)** MBP-Orf151 forms a dimer of 113 kDa, **(D)** when mixed MBP-Orf151 and SSB do not interact, **(E)** SSB elutes as an apparent tetramer of 84 kDa. Structures of each protein are displayed for reference purposes and are not intended to indicate the precise conformation or molar ratio of proteins present.

confirmed that the two proteins eluted separately. Similar results were obtained when MBP-Orf151 and SSBΔC10 proteins were mixed (Figure 7.3D). The results suggest that neither Orf nor Orf151 is able to bind SSBΔC10 under these specific experimental conditions.

7.4.2 Orf does not interact with Exo or β

When applied to the gel filtration column, His-Exo eluted as a complex of 50 kDa (Figure 7.4A), which correlates with the approximate size of a dimeric species (58 kDa). Previous studies, using X-ray crystallography and native polyacrylamide gel electrophoresis, revealed that His-tagged λ Exo trimers have a molecular weight of 85.4 kDa (Kovall and Matthews, 1997; Subramanian *et al.*, 2003). It is possible that the toroidal structure of the Exo trimer causes it to migrate anomalously in gel filtration. Alternatively, the salt concentration of the buffers used (250 mM KCl) could prevent formation of the trimeric complex. The exonuclease activity of Exo is known to be inhibited at relatively low salt concentrations (200 mM NaCl) and this could be due to disruption of subunit assembly (Little *et al.*, 1967). The protein was used at only 0.74 mg/ml and the UV absorption values for this protein were extremely low (~25 mAu). The assay should, therefore, be repeated with a more concentrated sample of purified Exo and at lower salt concentrations.

When MBP-Orf and Exo were mixed, two distinct protein peaks were observed, indicating that there is no interaction between these proteins (Figure 7.4B). This was confirmed by SDS-PAGE analysis of the elute fractions (Figure 7.4B). However, because Exo may not be present in its biologically active form (a trimer), no firm conclusions can be drawn from these results.

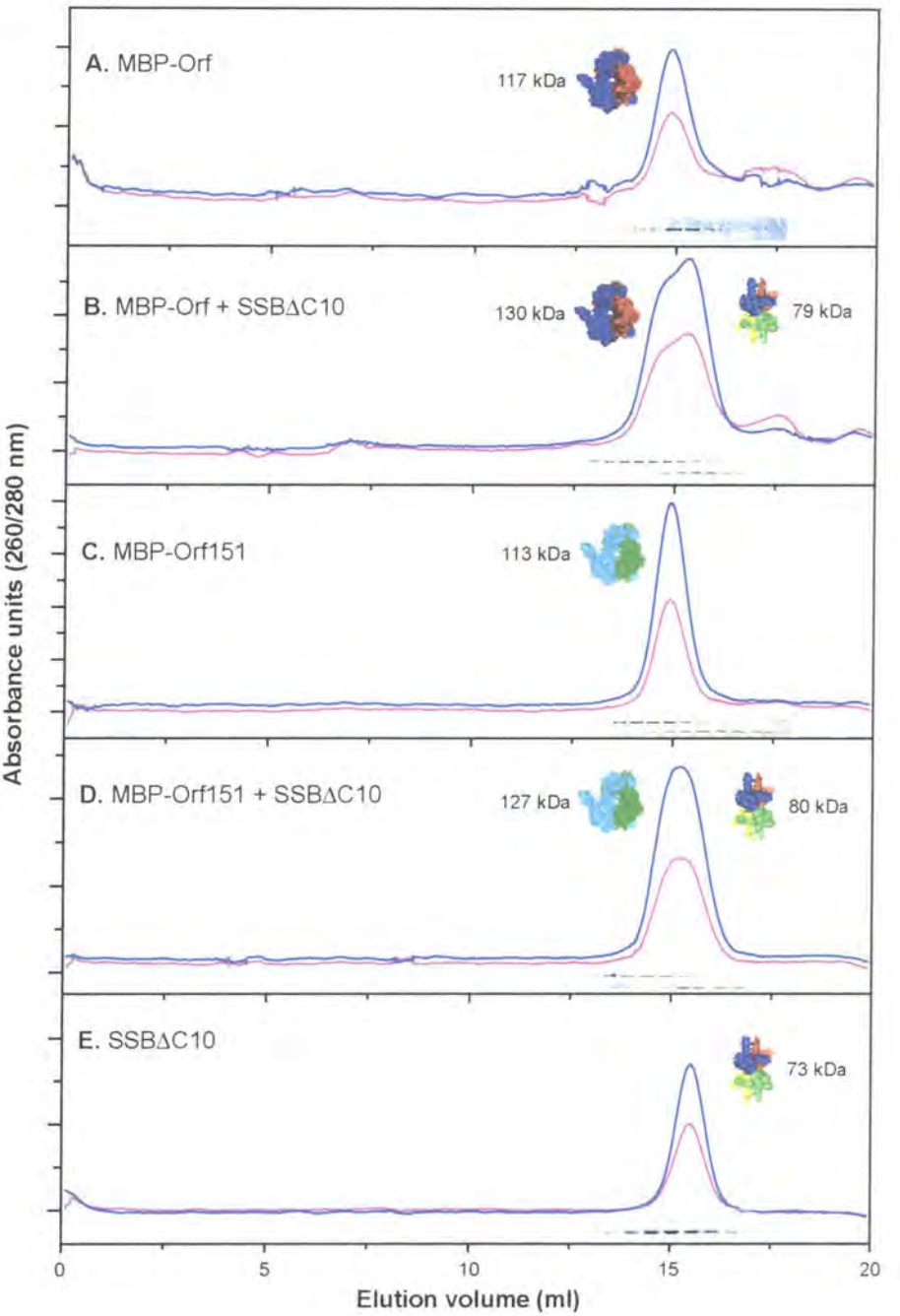

Figure 7.3. Size exclusion chromatography of MBP-Orf, MBP-Orf151 and SSBΔC10. Proteins (1 mg/ml) were loaded on a 24 ml Superose-6HR 10/30 column in 250 mM KCl Buffer A and eluted at a flow rate of 0.3 ml/min. Elution volume is shown on the x-axis and eluted proteins were collected by absorption at 260 nm (pink) and 280 nm (blue). Peak fractions were separated on 15% SDS-PAGE to confirm the presence of protein, and shown below the relevant peak in each panel. The predicted molecular weight of each protein monomer is: MBP-Orf (59 kDa), MBP-Orf151 (60 kDa) and SSBΔC10 (17.8 kDa). **(A)** MBP-Orf eluted as an apparent dimer of 117 kDa, **(B)** MBP-Orf and SSBΔC10 did not co-elute, **(C)** MBP-Orf151 forms a dimer of 113 kDa, **(D)** when mixed MBP-Orf151 and SSBΔC10 do not interact, **(E)** SSBΔC10 elutes as an apparent tetramer of 73 kDa. Structures of each protein are displayed for reference purposes and are not intended to indicate the precise conformation or molar ratio of proteins present. The structure of SSB is used to indicate SSBΔC10 here.

The elution profile of His-β compared with the standards suggested that it exists in a range of forms, with most of the protein eluting in a broad peak (Figure 7.4C). β protein (29 kDa) is known to exist in numerous states, forming multisubunit rings and helical filaments in the presence of Mg^{2+} and ssDNA (Passy *et al.*, 1999). Species with estimated molecular weights of 45 kDa and 182 kDa (probable monomer and tetramer, respectively) were observed. However, the absorption values were extremely low (5 mAu and 3 mAu, respectively), indicating that only a small fraction of the protein exists in these forms under these experimental conditions. No protein could be visualised in the elute fractions from these peaks (data not shown). When His-β was mixed with MBP-Orf only a single peak at 131 kDa was seen, corresponding to the latter protein (Figure 7.4D). On this occasion, β protein could be visualised on SDS gels (as an apparent monomer), however, the two protein peaks did not coincide (Figure 4.7D).

The above experiments could be repeated using native or histidine-tagged Orf and Orf151 and these might provide clearer data. However, as we have failed to show any direct interactions between Orf, Orf151 and SSB, SSBΔC10, Exo or β under the conditions employed here, it may also be necessary to use alternative methods to substantiate the interactions indicated by the yeast-two hybrid assay.

7.4.3 Orf and Orf151 bind SSB-ssDNA complexes

Gel filtration chromatography has previously been used to investigate complexes, formed between SSB interacting partners in the presence and absence of ssDNA (Sandigursky and Franklin, 1994; Kelman *et al.*, 1998; Umezu and Kolodner, 1994). Since Orf, Orf151 and SSB bind ssDNA in gel shift assays, the experiments

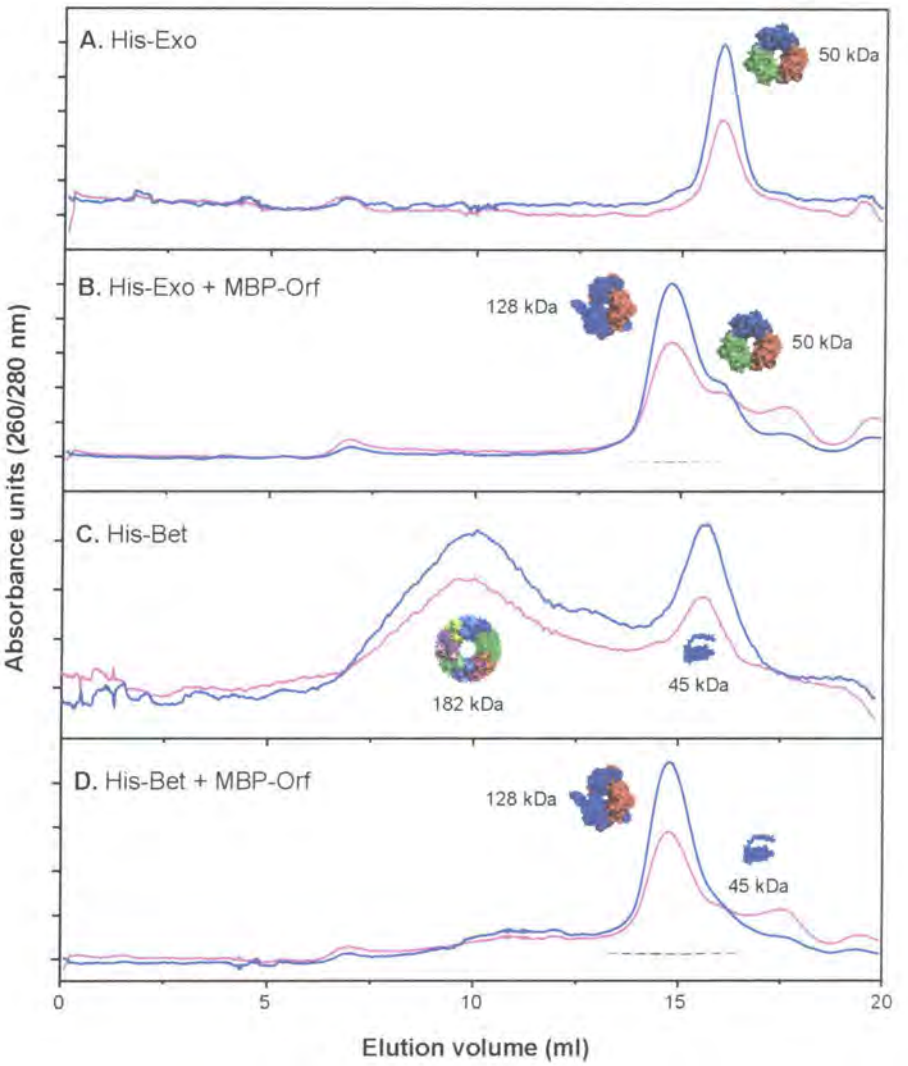

Figure 7.4. Size exclusion chromatography of MBP-Orf, His-Exo and His-β.
MBP-Orf (1 mg/ml) was mixed with His-Exo (0.74 mg/ml) or His-β (1 mg/ml) and incubated on ice for 15 minutes before loading on a 24 ml Superose-6HR column in 250 mM KCl Buffer A. Proteins were eluted in the same buffer at a flow rate of 0.3 ml/min, and detected by absorption at 260 nm (pink) and 280 nm (blue). Peak fractions were separated on 15% SDS-PAGE, as shown below the relevant peaks. His-β could not be visualised by SDS-PAGE analysis of either peak fraction suggesting most of the protein eluted in the void volume of the column. The predicted molecular weight of each protein monomer is; MBP-Orf (59 kDa), His-Exo (29 kDa) and His-β (32 kDa). **(A)** His-Exo eluted at a volume calculated to indicate a molecular weight of 50 kDa, suggesting dimer formation or a conformational change causing the trimer to migrate faster **(B)** His-Exo and MBP-Orf did not interact on the column and eluted separately, **(C)** most of the His-β protein eluted in the void volume, possibly as a mulisubunit complex, a small proportion eluted as a multimeric complex of 182 kDa and an apparent monomer (45 kDa), **(D)** MBP-Orf and His-β did not form a complex and eluted separately.

described above were repeated with the inclusion of a 51-base oligonucleotide, under the same binding conditions. Proteins (1 mg/ml) were mixed with ssDNA (0.1 mg/ml) and incubated on ice for 15 minutes prior to loading on to the column. The relative proportion of absorbance at 260 and 280 nm allows the discrimination of protein and DNA peaks; hence both wavelengths were used in these studies. Neither MBP-Orf (Figure 7.5A) nor MBP-Orf151 (Figure 7.5C) bound the ssDNA substrate, indicated by the presence of two distinct peaks; both eluted as apparent dimers (131 kDa and 114 kDa) and were visualised in the collected fractions corresponding to these peaks (Figure 7.5A and C). The increased absorption at 260 nm of one of the peaks indicated the presence of DNA, which alone elutes at 33 kDa (Figure 7.5F). The predicted molecular weight of the oligonucleotide is 15.75 kDa, however it eluted at almost double this value, probably because of its flexibility. The presence of DNA within these fractions was confirmed by agarose gel electrophoresis and SYBR green staining. It is possible that the relatively high salt concentration used in these assays disrupts electrostatic interactions between Orf/Orf151 and DNA, preventing complex formation. Alternatively, the protein:DNA interactions may be relatively unstable and subject to disruption under gel filtration conditions, which would also be exacerbated by the presence of salt.

In contrast, SSB bound the ssDNA substrate under these conditions (Figure 7.5E) and the estimated size of the complex (150 kDa) suggested that two SSB tetramers were assembled on the DNA. The fact that DNA is wrapped tightly around the SSB tetramer and does not significantly alter its size or shape, could explain why the molecular weight of the free ssDNA is not apparent in this value. SSB protein was evident in the peak fractions analysed on SDS-polyacrylamide gels (Figure

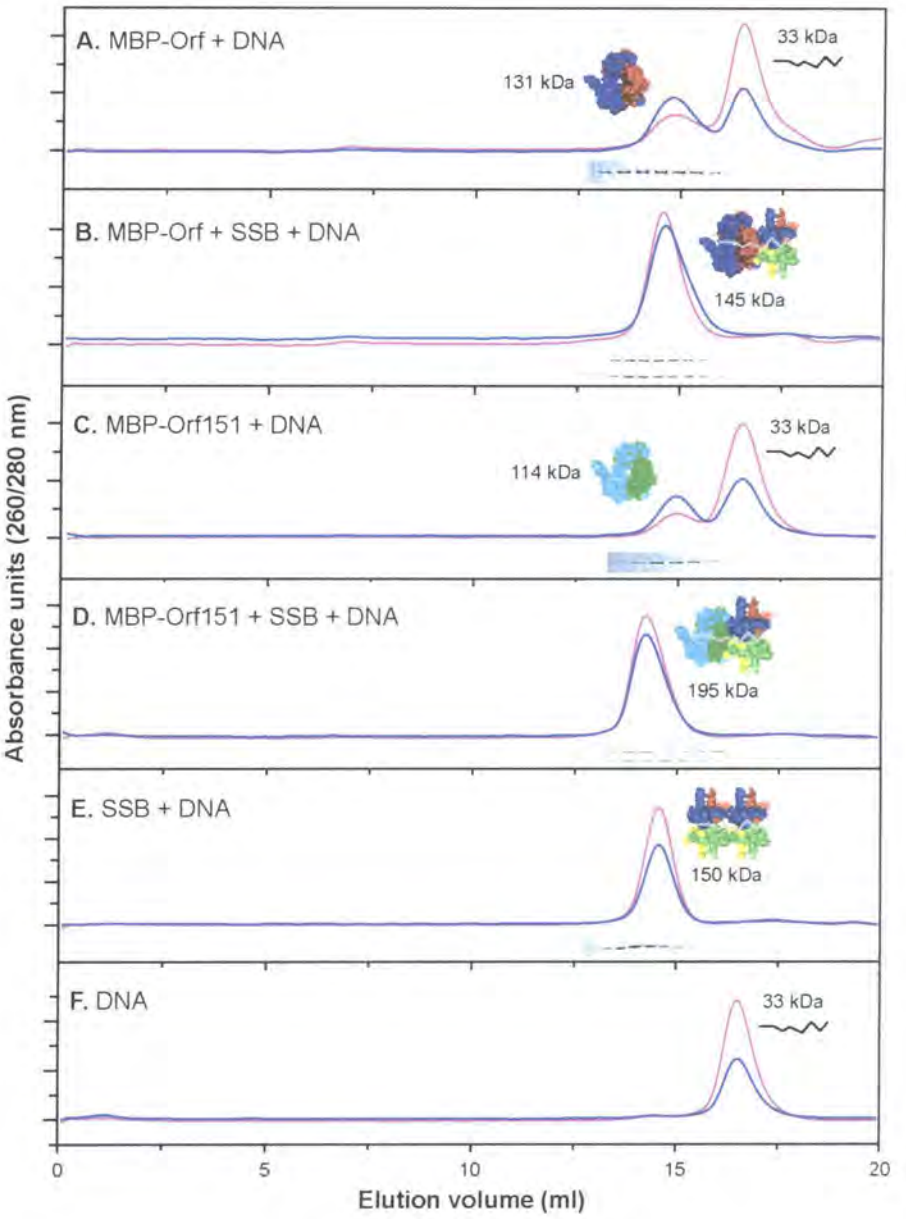

Figure 7.5. Size exclusion chromatography of MBP-Orf, MBP-Orf151 and SSB in the presence of ssDNA. Proteins (1 mg/ml) and ssDNA (0.1 mg/ml) in the indicated combinations, were mixed in 250 mM KCl Buffer A, incubated on ice for 15 minutes, and loaded on a 24 ml Superose-6HR column. Proteins and DNA were eluted in the same buffer, detected by absorpion at 260 nm (pink) and 280 nm (blue) and separated on 15% SDS-PAGE. DNA presence was indicated by the high absorption at 260 nm (pink) and confirmed by agarose gel electrophoresis and Sybr green staining (data not shown). Predicted molecular weights of protein monomers were as described in the legend to Figures 7.2 and 7.3. **(A)** MBP-Orf did not interact with ssDNA under these conditions, eluting separately **(B)** in the presence of ssDNA, MBP-Orf and SSB formed a complex of 145 kDa, as indicated by the single peak and co-elution as visualised by SDS-PAGE, **(C)** MBP-Orf151 and ssDNA eluted separately, two distinct elution profiles were seen and ssDNA was absent from the peak fraction containing MBP-Orf151, **(D)** MBP-Orf151 and SSB form a 195 kDa complex with ssDNA, and the proteins co-elute, **(E)** SSB bound ssDNA and the comlpex eluted with a calculated molecular weight of 150 kDa, **(F)** ssDNA (51nt) alone eluted with an estimated molecular weight of 33 kDa. Assumed complexes formed are displayed diagramatically in each panel, but are not intended to indicate the precise conformation of the complexes or the exact molar ratio of species present.

7.5E) and DNA was detected by SYBR green staining of agarose gels (data not shown). Under these experimental conditions, with a salt concentration of 250 mM (KCl), we would expect SSB to adopt the $(SSB)_{65}$ mode of binding (Lohman and Overman, 1985). This binding mode is favoured at NaCl concentrations above 200 mM and involves all four SSB subunits contacting the ssDNA and protecting 65 nt of ssDNA. Since only a 51 nt ssDNA was used here, it seems likely that the complex identified here is a dimer of tetramers (octamer), suggesting SSB is present in the highly cooperative $(SSB)_{35}$ mode, where only two subunits of the tetramer contact the DNA, or in one of the transitional modes observed at salt concentrations between 100 and 200 mM (Lohman and Overman, 1985). It is generally agreed that SSB under physiological salt and SSB concentrations binds ssDNA in the $(SSB)_{65}$ mode, although it is also thought that at high SSB or low ssDNA concentrations (which can occur in the cell) that the $(SSB)_{35}$ mode can be adopted (Meyer and Laine, 1990). The estimated physiological concentration of SSB is 0.5 μM (Lohman and Overman, 1985), however, in this assay SSB was used at a concentration of 1 mg/ml (52.7 μM) which could explain the binding of two SSB tetramers apparently in the $(SSB)_{35}$ mode upon the ssDNA substrate. Attempts to quantify the molar ratio of proteins present in the peaks corresponding to complex profiles, by staining of polyacrylamide gels with Sypro orange, were unsuccessful. A single fraction corresponding to the crest of the elution peak was utilised and the protein concentration was too low due to the dilution factor. Pooling and subsequent concentration of all protein containing fractions would be required for reliable quantification. The precise details of whether one or a pair of tetramers can assemble on the 51-nt ssDNA is not crucial for the experiments described here.

When MBP-Orf was mixed with DNA and SSB, a single peak of 145 kDa was observed and the proteins were seen to co-elute when analysed by SDS-PAGE (Figure 7.5B). The high absorbance at 260 nm indicated the presence of DNA within this peak and this was confirmed by analysis of elute fractions by agarose gel electrophoresis and Sybr green staining (data not shown). This result could suggest the simultaneous binding of an MBP-Orf monomer and an SSB tetramer to the DNA. However we know MBP-Orf is a dimer in its biologically active form and that SSB binds ssDNA as a tetramer. No peak corresponding to an MBP-Orf monomer was seen, making the dissociation of the dimer an unlikely explanation for the smaller than expected complex size. It is possible that the interaction of an MBP-Orf dimer with SSB or ssDNA and SSB stimulates a conformational change in the protein making it more compact and causing it to migrate much more slowly. It is the shape of a molecule that determines how it elutes from gel filtration columns, thus a difference in the overall shape of the MBP-Orf-SSB-ssDNA complex could account for the smaller than expected mass observed.

A similar result was obtained when MBP-Orf151, SSB and ssDNA were mixed, with a single peak of 195 kDa containing both proteins and the DNA as determined by SDS-PAGE and Sybr green staining (Figure 7.5D). The simplest explanation for this result based upon the mass of the complex, involves a dimer of MBP-Orf151 and a tetramer of SSB binding simultaneously to DNA. If the ssDNA is bound tightly around the proteins, as seen with SSB alone, the individual MW of the DNA may not be apparent. An MBP-Orf151 dimer (114 kDa) in association with an SSB tetramer (84 kDa) would account almost exactly for the size of the complex seen.

The above assays were repeated using SSBΔC10, in place of SSB, and produced remarkably similar results (Figure 7.6). The elution profiles suggest that both MBP-Orf (Figure 7.6B) and MBP-Orf151 (Figure 7.6D) interact with SSBΔC10 in the presence of the 51nt ssDNA substrate. This indicates that any direct interaction between Orf and Orf151 with SSB does not occur via the C-terminus of SSB.

Subsequent experiments, performed by FA Curtis, examined interactions between purified MBP and SSB in the presence or absence of the 51-nt ssDNA. The two proteins failed to associate in both cases suggesting that any interaction between MBP-Orf/MBP-Orf151 and SSB in the presence of ssDNA is not due to the MBP domain (FA Curtis, personal communication).

The formation of MBP-Orf151-SSB-ssDNA complexes was also examined using a 25 base oligonucleotide (ϕX174-25) under the same conditions. The oligonucleotide alone eluted at a molecular weight of 15 kDa (Figure 7.7C), consistent with a high 260 nm absorbance and confirmed by Sybr green staining of agarose gels. In mixed binding reactions, neither MBP-Orf151, SSB nor SSBΔC10 bound this 25-nt ssDNA (Figure 7.7A and B). Formation of MBP-Orf151-SSB or MBP-Orf151-SSBΔC10 complexes was not detected in the presence of this shorter ssDNA substrate; the proteins formed dimers and tetramers, respectively, and failed to co-elute (Figure 7.7 A and B). These results suggest that the SSB tetramer requires a substrate of greater than 25-nt in order to bind and accommodate Orf151.

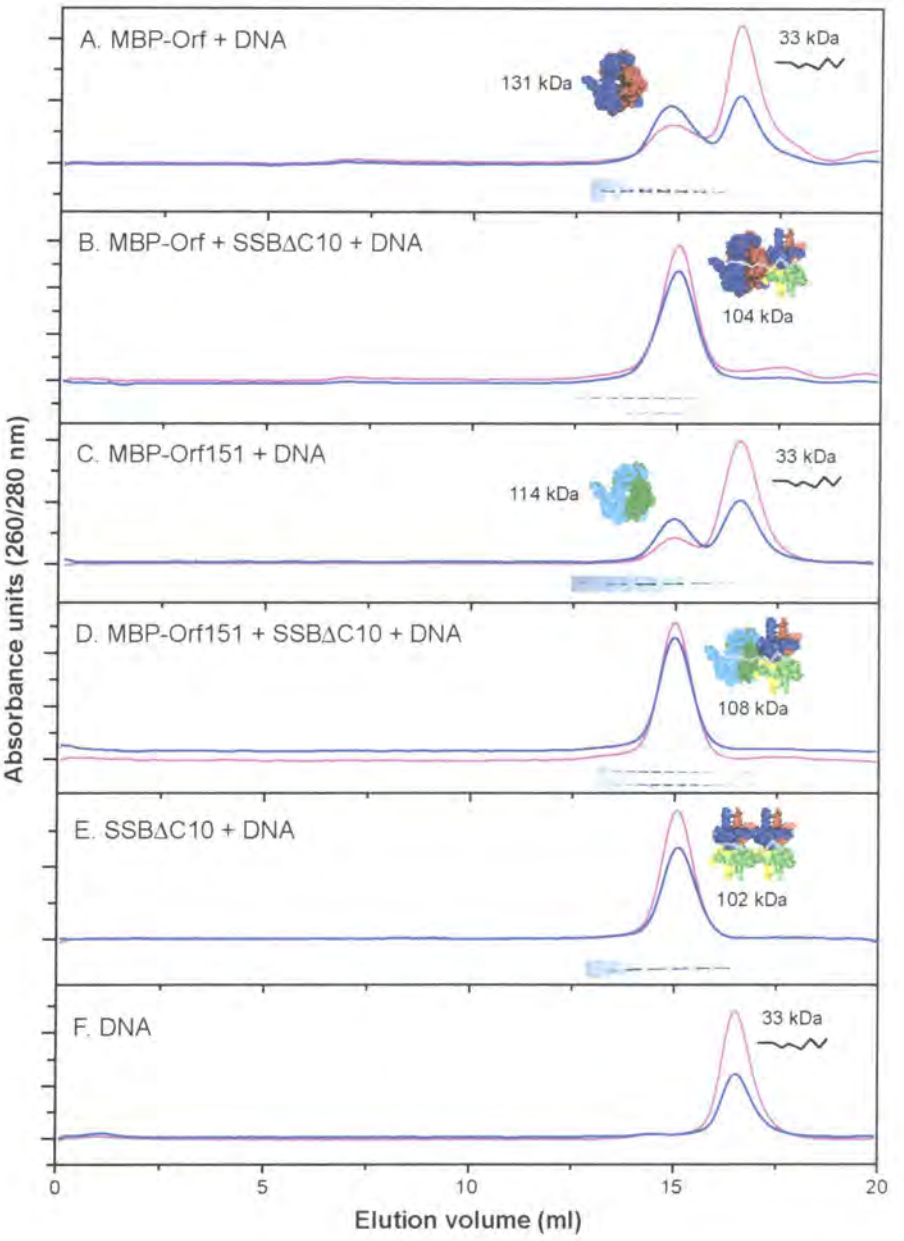

Figure 7.6. Size exclusion chromatography of MBP-Orf, MBP-Orf151 and SSBΔC10 in the presence of ssDNA. Proteins (1 mg/ml) and ssDNA (0.1 mg/ml) in the indicated combinations, were analysed as described in the legend to Figure 7.5. Proteins were detected by absorpion at 260 nm (pink) and 280 nm (blue) and separated on 15% SDS-PAGE. DNA presence was indicated by the high absorption at 260 nm (pink) and confirmed by agarose gel electrophoresis and Sybr green staining (data not shown). Predicted molecular weights of protein monomers were as described in the legend to Figures 7.2 and 7.3. **(A)** MBP-Orf did not interact with ssDNA under these conditions, **(B)** in the presence of ssDNA, MBP-Orf and SSBΔC10 formed a complex of 104 kDa, and co-elution was visualised by SDS-PAGE, **(C)** MBP-Orf151 and ssDNA did not interact and eluted separately, **(D)** MBP-Orf151 and SSBΔC10 form a 108 kDa complex with ssDNA, and the proteins co-elute, **(E)** SSBΔC10 bound ssDNA and the complex eluted with a calculated molecular weight of 102 kDa, **(F)** ssDNA (51-nt) alone eluted with an estimated molecular weight of 33 kDa. Assumed complexes formed are displayed diagramatically in each panel, but are not intended to indicate the precise conformation of the complexes or the exact molar ratio of species present.

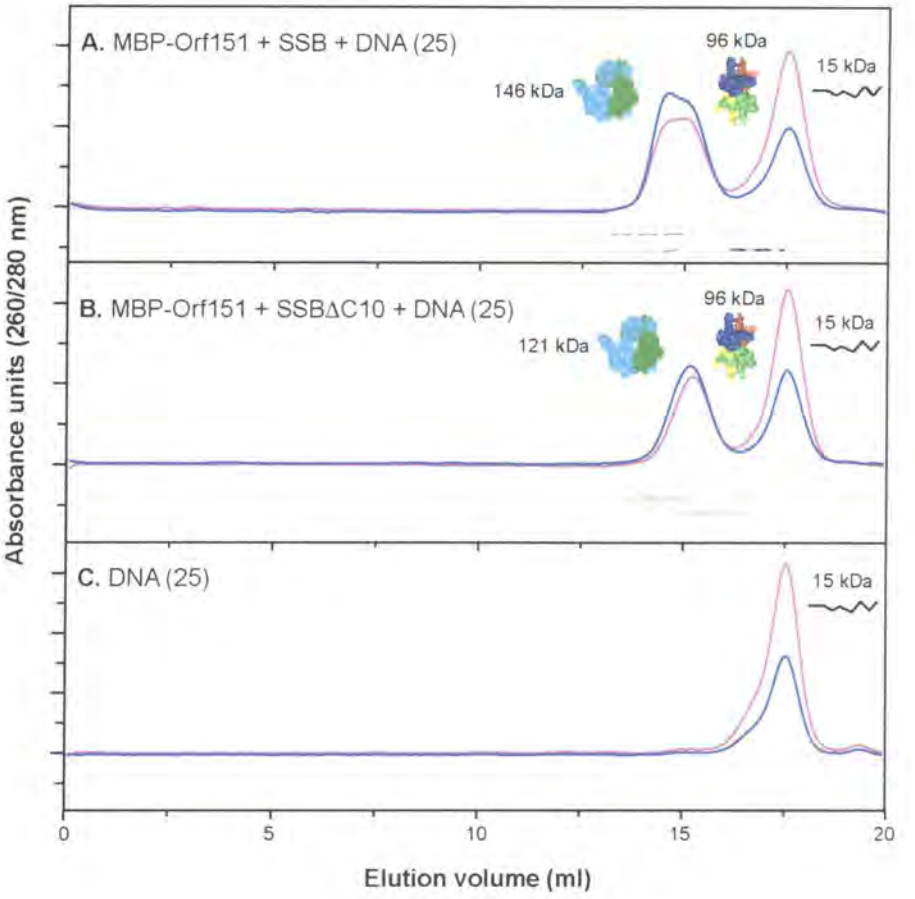

Figure 7.7. Size exclusion chromatography of MBP-Orf151, SSB and SSBΔC10 in the presence of a 25nt ssDNA. Proteins (1 mg/ml) and ssDNA (0.1 mg/ml) were mixed in 250 mM KCl Buffer A and incubated on ice. Reactions were analysed by loading on a 24 ml Superose-6HR column and eluting in 250 mM KCl Buffer A at a flow rate of 0.3 ml/min. Proteins and DNA were collected by absorption at 260 nm (pink) and 280 nm (blue), and were visualised by SDS-PAGE and agarose gel electrophoresis analysis respectively. The predicted molecular weight of each protein monomer is; MBP-Orf151 (60 kDa), SSB (19 kDa) and SSBΔC10 (17.7 kDa). **(A)** MBP-Orf151 and SSB did not form complexes in the presence of the 25nt ssDNA oligonucleotide, the trailing shoulder of the peak indicates the overlap of two separate elution profiles, and the proteins did not co-elute. The higher UV absorbance trace at 260 nm than at 280 nm indicated the presence of ssDNA in the lower molecular weight profile, and DNA was not detected in fractions containing proteins **(B)** MBP-Orf151 and SSBΔC10 did not interact with ssDNA, the proteins did not co-elute and ssDNA was absent from peak fractions containing the proteins **(C)** the 25nt ssDNA alone eluted at a molecular weight of 15 kDa.

7.5 Discussion

We have investigated the ability of Orf and Orf151 to interact with Exo, β and SSB by far-western blotting, affinity chromatography and gel filtration. We failed to demonstrate a direct interaction between either Orf, or Orf151 with Exo or β in far western and gel filtration analyses. It is possible that these proteins do not directly interact, or the presence of both Exo and β is required for complex formation. Exo and β are known to interact (Muniyappa and Radding, 1986; Passy *et al.*, 1999) and it is feasible that Orf, and possibly Orf151, requires the presence of both recombinases to function appropriately in λ recombination.

Gel filtration analysis revealed that neither MBP-Orf nor MBP-Orf151 associated with ssDNA or SSB under the conditions employed. However, Orf-SSB and Orf151-SSB complexes were formed in the presence of ssDNA. This could be due to Orf and Orf151 recognising a conformational change in SSB induced by ssDNA binding. How can this be reconciled with the far-western assays showing an Orf/Orf151 interactions with SSB in the absence of ssDNA (this study, unpublished data and (Maxwell et al., 2005). It is possible that transfer of SSB to the hydrophobic PVDF membrane under denaturing conditions induces the exposure of epitopes on SSB, normally hidden when in solution, facilitating the binding of Orf or Orf151. These epitopes may normally be revealed when SSB encounters ssDNA and serve as an assembly point for Orf and Orf151. Given the evidence presented here, it may be that Orf and Orf151 associate with the SSB-DNA interface to mediate the loading of a recombinase onto SSB-coated ssDNA. Further studies on the interactions between Orf/Orf151 and SSB in the presence and absence of DNA are necessary to uncover precisely how these proteins assemble on recombination intermediates. The preferred

genetic exchange partner for Orf, either RecA or β, has yet to be identified. Distinguishing between the two is a crucial feature, as it will determine whether strand annealing or invasion pathways are favoured in λ recombination.

Chapter 8

Complementation of *E. coli ruv* mutant UV sensitivity by heterologous RecU and RuvC proteins

8.1 Introduction

In Chapter 1 we described homologous recombination as a fundamental cellular process, its key role being the salvage of stalled or broken replication forks (Haber, 1999; Kuzminov, 1999; Cox *et al.*, 2000; Michel *et al.*, 2004). Homologous recombination ensures accurate repair of a variety of chromosomal lesions and generates the exchanges and rearrangements fundamental to the diversification of bacterial genomes. Recombination is initiated at the site of DNA damage by the action of presynaptic enzymes, exposing sufficient ssDNA for assembly of a RecA filament. RecA then mediates strand-invasion and strand-exchange (synapsis) with a homologous (undamaged) duplex. The resulting intermediate, the four-stranded (Holliday) junction, can also be generated by regression of a stalled replication fork (Seigneur *et al.*, 1998). Once formed, the Holliday junction can be moved by branch migration and is usually eliminated by paired cleavages at the cross-over. Resealing of the nicked duplex products by DNA ligase reconstitutes intact chromosomes (Seigneur *et al.*, 1998; Zerbib *et al.*, 1998). The importance of Holliday junction-resolving enzymes in homologous recombination and DNA repair is highlighted by their widespread occurrence in nature (Lilley and White, 2001; Sharples, 2001; Aravind *et al.*, 2000).

In *E. coli* and many other Gram-negative bacteria, Holliday junctions can be branch migrated and/or resolved by a number of enzymes, including the RusA resolvase (Chan *et al.*, 1998), the RecG helicase (Lloyd and Sharples, 1993; Whitby *et al.*, 1994) and the RuvABC complex (Sharples *et al.*, 1999b). The RuvC protein is a homodimeric endonuclease that resolves Holliday intermediates by cleavage at specific sequences located symmetrically at the crossover (Hagan *et al.*, 1998; Dunderdale *et al.*, 1991; Iwasaki *et al.*, 1991; Yamada *et al.*, 2004)(Figure 8.1A). The RuvA protein forms a four-fold symmetric tetramer that binds specifically to the Holliday structure, with each of the four arms of the structure being bound within grooves on surface of the tetramer (Rafferty *et al.*, 1996). Binding by RuvA forces the junction into an approximately square planar configuration, optimal for branch migration (Hargreaves *et al.*, 1998). Assembly of the RuvA tetramer onto the junction promotes loading of two hexameric RuvB helicase-motor proteins onto opposing duplexes, flanking the RuvA-junction complex, with the DNA fed through the central cavity of each RuvB ring (Stasiak *et al.*, 1994). Branch migration of the Holliday junction is achieved by the helicase motor of RuvB drawing DNA through the complex, each duplex arm rotating within the surface channels of RuvA (Rafferty *et al.*, 1996; Parsons *et al.*, 1995). The RuvC endonuclease component can bind the accessible face of the junction and interacts with the RuvAB branch migration complex to form the RuvABC resolvasome (van Gool *et al.*, 1998; van Gool *et al.*, 1999; Davies and West, 1998)(Figure 8.1A, B and C). RuvA and RuvC target the X-structure directly, while interactions between these two proteins and RuvB serve to stabilise complex formation (Sharples *et al.*, 1999b; Yamada *et al.*, 2004). RuvC specifically cleaves between the third and fourth positions in the consensus sequence

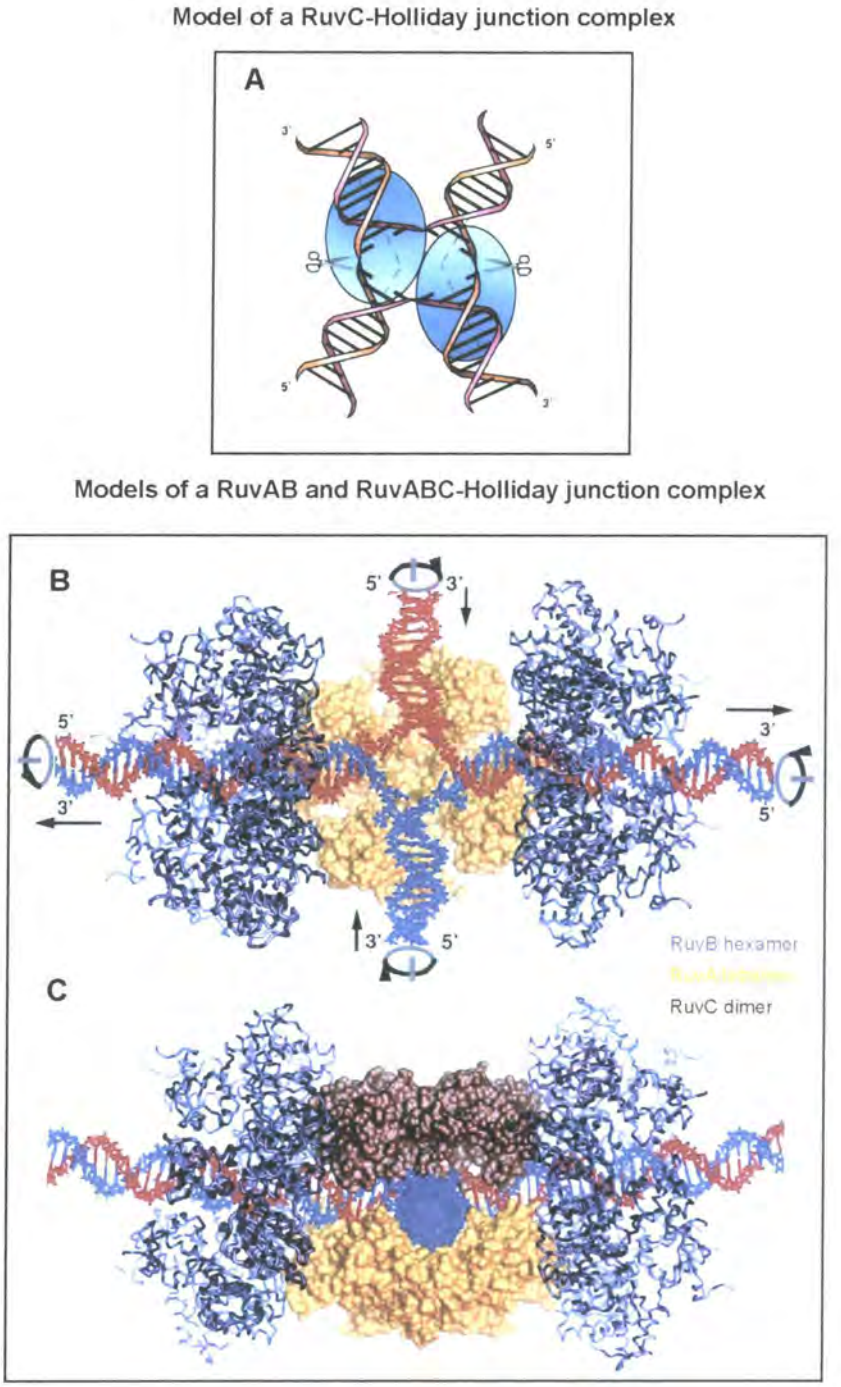

Figure 8.1. Models of RuvC, RuvAB and RuvABC bound to the Holliday junction. **(A)** A model for the RuvC Holliday junction resolvase cleaving the X structure in an unfolded conformation imposed by RuvC binding. **(B)** A model of the RuvAB-Holliday junction complex, showing RuvB hexamers (dark blue and purple) on opposite sides of a single RuvA tetramer (yellow). The direction of DNA migration and rotation during unwinding by RuvAB is indicated by arrows. **(C)** A model of the RuvABC-Holliday junction (resolvasome), rotated 45° relative to (B). The RuvC dimer (brown) is accommodated opposite RuvA with its active site positioned for dual strand cleavage. Adapted from Putnam et al 2001 and Hagan et al 1998.

5'-$^{A}/_{T}$TT$^{G}/_{C}$-3' (Shah *et al.*, 1994). Thus the RuvABC resolvasome functions to 'scan' the DNA as it is drawn through the complex until target sequences are found, allowing for the twin RuvC active sites to be positioned appropriately for dual strand cleavage.

Representatives of the RuvA and RuvB proteins are widespread in bacteria, with orthologs present in diverse species, such as *Neisseria sp*, *Yersinia pestis*, *Streptococcus sp* and *Helicobacter pylori* (Sharples *et al.*, 1999b). The possibility that RuvAB can restore regressed replication forks by branch migration in the absence of RuvC (Seigneur *et al.*, 1998) may explain their ubiquity in bacteria. Homologs of RuvC are less prevalent and Gram-positive bacteria lack an identifiable orthologue altogether, utilising an unrelated resolvase, RecU (Ayora *et al.*, 2004). Distantly-related RuvC-like proteins have however been identified in phages and prophages from Lactococci and Streptococci (Bidnenko *et al.*, 1998). Phage λ also encodes a branch-specific endonuclease, Rap, which can eliminate branched structures arising from λ Red-mediated recombination and can function as a Holliday junction resolvase (Hollifield *et al.*, 1987; Tarkowski *et al.*, 2002; Sharples *et al.*, 1998; Sharples *et al.*, 1999a; Sharples *et al.*, 2004). Rap is thought to be a functional equivalent of the *E. coli* RusA Holliday junction resolvase (Sharples *et al.*, 2002; Mahdi *et al.*, 1996) and can substitute for *E. coli* RuvC in Red-mediated recombination reactions (Poteete *et al.*, 2002).

Although closely related homologs of *E. coli* RuvC are lacking in many bacterial species, functional equivalents have been isolated from numerous bacteria and phages (as outlined above). To explore the importance and functionality of different Holliday junction resolvases from phage and bacterial systems, we decided

to investigate their impact on the survival of *E. coli ruv* mutants after exposure to UV light. The availability of different of *ruvA*, *ruvB* and *ruvC* mutant combinations allowed us to probe the requirements for both branch migration and resolution *in vivo*. Studies were conducted with *Bacillus subtilis* RuvA, RuvB and RecU, *Lactococcus lactis* phage bIL67 RuvC (67RuvC) and *Helicobacter pylori* RuvC. The results described in this chapter, in conjunction with data obtained from *in vitro* experiments, have provided significant new insights into the workings of these enzymes in the removal of branched DNA intermediates.

8.2 *B. subtilis* RecU suppresses the UV sensitivity of *E. coli ruvC* mutants

Bacillus subtilis, a model Gram-positive bacterium, shares most of the recombination genes found in *E. coli* and many have been given the same gene designation. These include *recA*, *recF*, *recG*, *recJ*, *recN*, *recO*, *recQ*, *recR*, *ruvA* and *ruvB*. However, certain key enzymatic components of the recombinational repair apparatus differ significantly from those found in *E. coli*. One example is *addAB*, encoding functional analogs of the *E. coli recBCD* genes (Fernandez *et al.*, 2000). Similarly RecU appears to function as a structurally-unrelated equivalent of RuvC. RecU is highly conserved among Gram-positive bacteria and involved in homologous recombination, DNA repair and chromosome segregation (Ayora *et al.*, 2004; Fernandez *et al.*, 2000). The protein binds preferentially to three- and four-stranded junctions, although it can also associate with ssDNA and dsDNA with much reduced affinity. In the presence of Mg^{2+}, RecU cleaves mobile Holliday junctions, introducing symmetrically-related nicks with a certain sequence specificity (Ayora *et*

al., 2004). RecU presumably functions to resolve Holliday junctions following branch migration by the *B. subtilis* RuvAB helicase. However, at the outset of this study it was unclear if RuvAB and RecU acted in concert in a manner analogous to the *E. coli* RuvABC complex (Ayora *et al.*, 2004).

Holliday junction processing in Gram-positive bacteria was investigated by examining the effect of the cloned *B. subtilis ruvA, ruvB, ruvAB* and *recU* genes (provided by JC Alonso, University of Madrid) on the UV sensitivity of *E. coli ruvA, ruvB, ruvAB* and *ruvC* mutants.

8.2.1 Bacterial strains and plasmids

The *E. coli* K12 mutant strains used in this study were: SR2210 (*ruvA200*), N1057 (*ruvB4*), N2057 (*ruvAB60*::Tn*10*), GS1481 (Δ*ruvC*::*kan*), GS2053 (*ruvC64*::kan) and CS85 (*ruvC53 eda*::Tn*10*) and all are derivatives of the *ruv*$^+$ wild-type strain AB1157 (Sargentini and Smith, 1989; Sharples *et al.*, 1990; Mandal *et al.*, 1993). Plasmid constructs were as described (Sanchez *et al.*, 2005; Sharples *et al.*, 1990; Sharples and Lloyd, 1991) and genes from *E. coli* and *B. subtilis* designated as *Ec* or *Bs*, respectively. The *ruv* mutant and wild-type strains were transformed with the appropriate plasmids and their vector controls. Successful transformants were selected on LB agar containing ampicillin for pUC18 vectors or chloramphenicol for pHP13 or pACYC184 derivatives. The UV sensitivity of transformants was analysed as described in Chapter 2 (section 2.3.5).

8.2.2 Complementation analysis

The plasmid constructs carrying *B. subtilis* RuvA, RuvB or RuvAB failed to improve the UV sensitivity of the relevant *E. coli ruv* mutants (Figure 8.2), suggesting that they are incapable of replacing *E. coli* RuvAB function. The *Bs*RuvA and *Bs*RuvAB clones conferred a significant negative effect on wild-type *E. coli* cells following exposure to UV light (Figure 8.2). This effect is similar to that observed with plasmids expressing high levels of *Ec*RuvA (Sharples *et al.*, 1990). In both cases, this negative effect on the survival of UV-damaged cells is probably a consequence of RuvA binding the Holliday junction DNA and preventing the access other junction processing enzymes (Sharples *et al.*, 1990). The results suggest that *Bs*RuvA and *Bs*RuvB either cannot function as a complex with *Ec*RuvC or the detection of any complementation is masked by the negative effect of *Bs*RuvA.

In contrast, plasmids carrying *B. subtilis* RecU fully restored the UV resistance of *ruvC* mutant strains (Figure 8.3), although it failed to improve the UV sensitivity of *E. coli ruvA*, *ruvB* or *ruvAB* mutants. An additional RecU clone with improved gene expression did partially improve the UV sensitivity of *E. coli ruvA*, *ruvB* and *ruvAB* mutants (data not shown). The results show that RecU can act as a Holliday junction resolvase *in vivo* by replacing *E. coli* RuvC function. *E. coli* RuvC is dependent upon the RuvAB branch migration complex in order to promote efficient Holliday junction resolution *in vivo* (Figure 8.3). The same appears to be true of RecU as it only managed to restore complete UV resistance to the *ruvC* mutant strain. We conclude that: i) RecU requires the branch migration activity of RuvAB for activity and can function with the *E. coli* RuvAB system. ii) Any important contacts that assist in stabilising a RuvABC or RuvAB-RecU complex must be

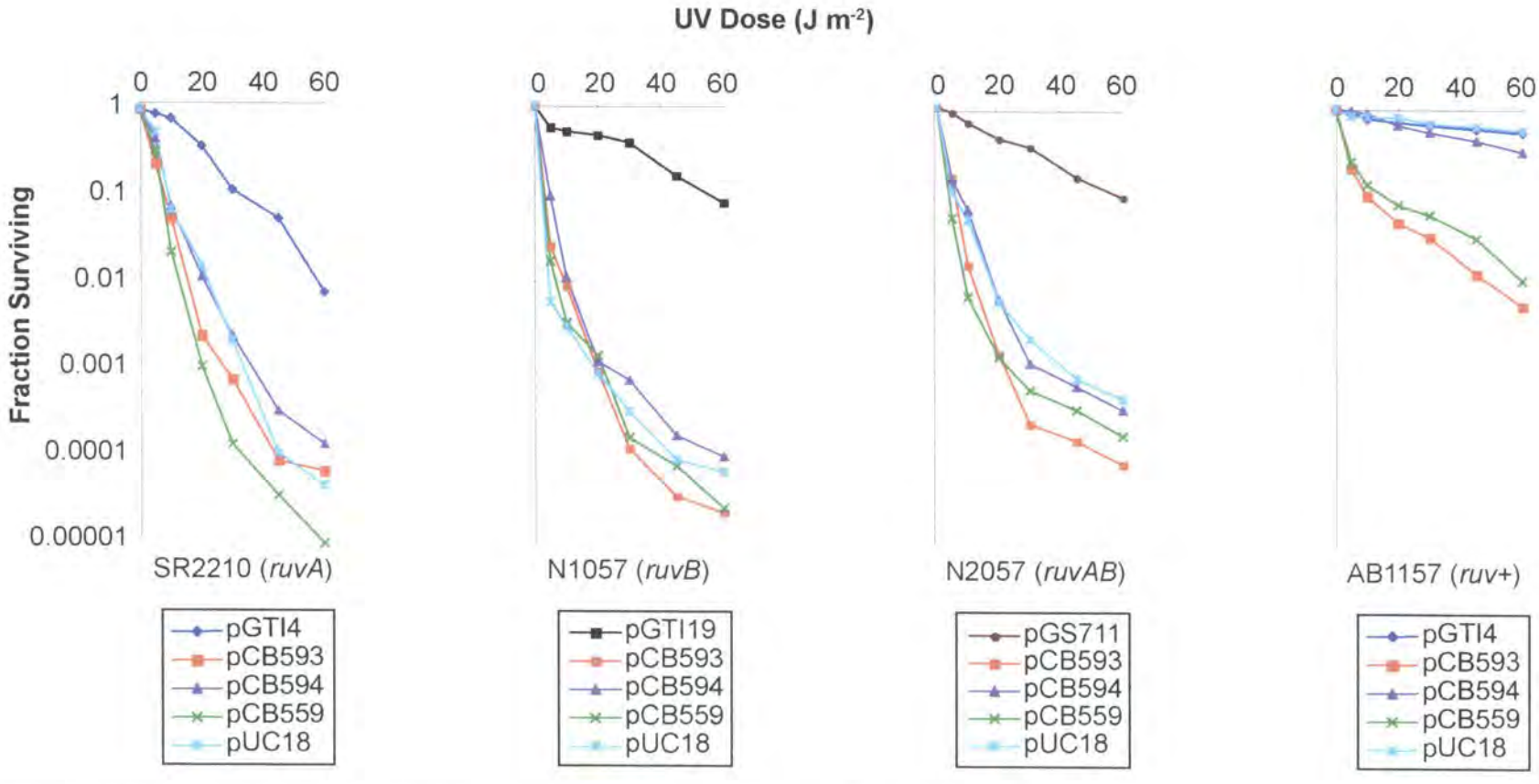

Figure 8.2. Survival of UV-irradiated *E. coli ruv* mutants carrying *Bacillus subtilis ruvAB gene* constructs.
Strain, genotype and a key to plasmid symbols used are shown below each panel. Plasmids used were pGTI4 (*Ec*RuvA in pUC18), pGTI19 (*Ec*RuvB in pUC18), pGS711 (*Ec*RuvAB in pUC18), pCB593 (*Bs*RuvA in pUC18), pCB594 (*Bs*RuvB in pUC18), and pCB559 (*Bs*RuvAB in pUC18).

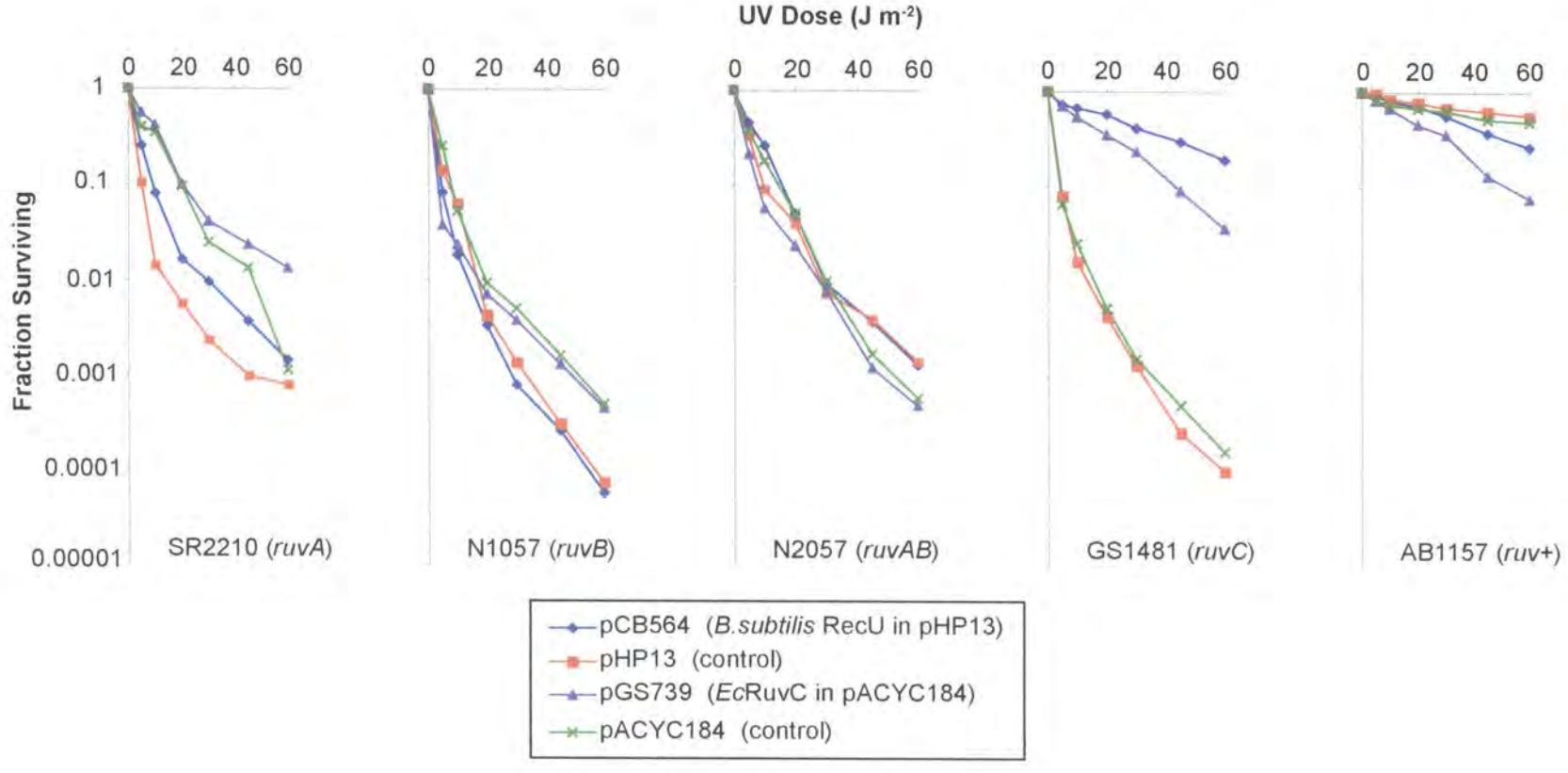

Figure 8.3. Survival of UV-irradiated *E. coli ruv* mutants carrying the *B.subtilis recU* gene construct.
Strain and genotype are indicated below each panel. Plasmids and symbols used are shown below the figure.

conserved between the two species. iii) RecU does not function effectively as a stand-alone resolvase, implying that it depends on RuvAB for junction targeting or branch migration to its preferred cleavage sites.

Several RecU mutants have been generated in the putative endonuclease active site, producing proteins that retain the ability to bind but not cleave Holliday junction DNA (McGregor *et al.*, 2005). Two of the resolution-defective mutants (E36A and H119A) were examined for their ability to complement the UV sensitivity of *E. coli ruvC* mutants. The RecU mutants failed to suppress the UV sensitive phenotype of the *ruvC* mutants (Figure 8.4), confirming the functional importance of these residues.

8.3 *Lactococcus lactis* phage bIL67 RuvC restores UV resistance to *E. coli ruvA, ruvAB* and *ruvABC* mutant strains

Lactococcus lactis is a Gram-positive facultative anaerobe used extensively in the dairy industry for the production of buttermilk and cheese. The use of *L. lactis* in dairy factories is not without issues, as they are subject to strong selective pressure from specific bacteriophages present in the dairy environment. Infection by phages results in significant economic losses each year by preventing the bacteria from fully metabolizing the milk substrate (Coffey and Ross, 2002). *L. lactis* has evolved a defensive strategy, known as abortive phage infection (Abi), to deal with the threat of phage attack. In Abi, normal phage infection is followed by an interruption to phage development, leading to the release of a small number of progeny particles and death of the bacterial cell (Snyder, 1995). Studies on *L. lactis* phage bIL66, in an effort to define the molecular targets of one of these Abi pathways, uncovered an

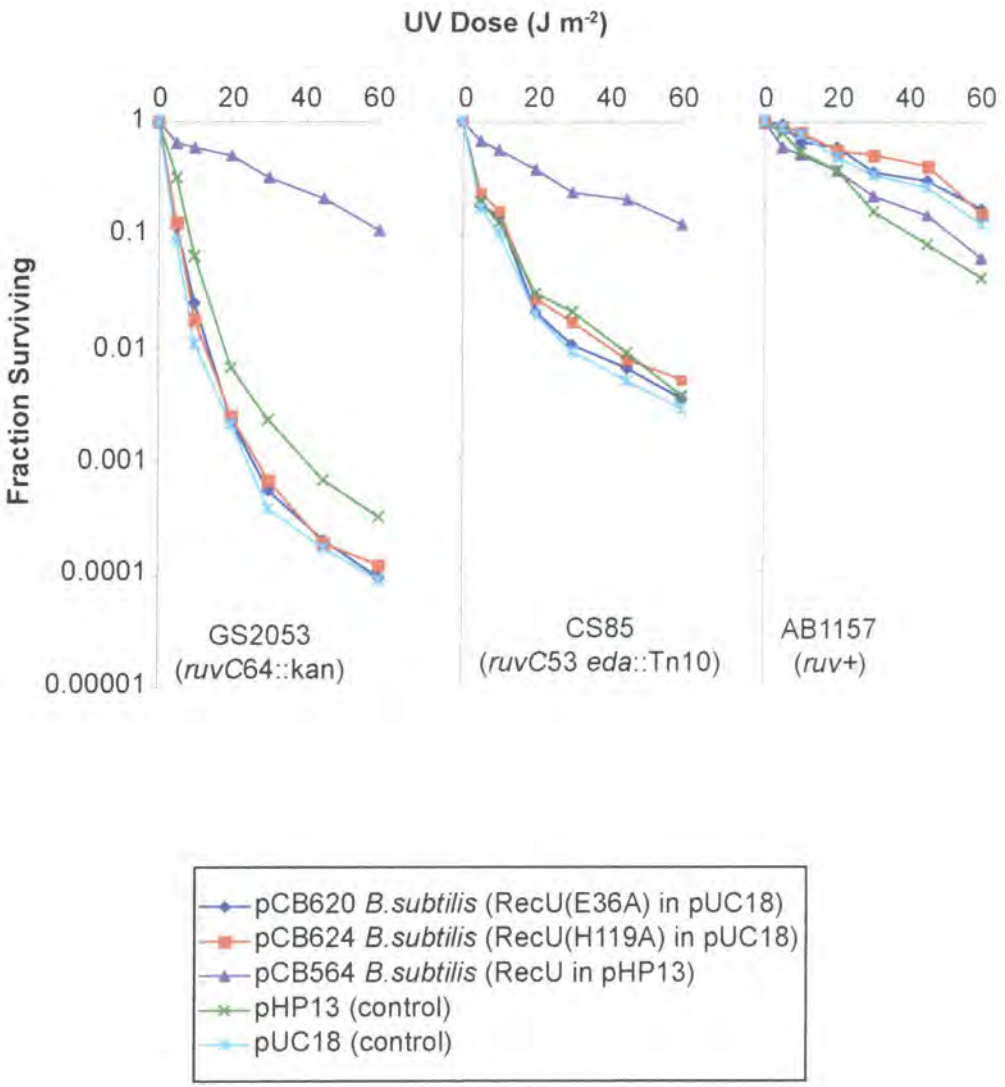

Figure 8.4. Survival of UV-irradiated *E. coli ruv* mutants carrying *B.subtilis recU* wild-type and mutant gene constructs. Strain and genotype are indicated below each panel. Plasmids and symbols used are shown below the figure.

operon encoding an endonuclease with homology to *E. coli* RuvC (Bidnenko *et al.*, 1998). They found that the phage bIL66 RuvC-like protein shared 57% sequence similarity to *E. coli* RuvC, and 43% identity to its equivalent from the lactococcal phage bIL67 (Schouler *et al.*, 1994). Both phage RuvC-like proteins have three of the four conserved acidic residues (Asp-7, Glu-66, Asp-138 and Asp-141) known to comprise the active site of *E. coli* RuvC (Saito *et al.*, 1995).

In an attempt to characterise the RuvC-like protein from *L. lactis* phage bIL67, the relevant gene (67*ruvC*) was cloned in the pT7-7 expression vector to give pGS905 (Curtis, 2004). As demonstrated with bIL66 constructs (Bidnenko *et al.*, 1998), *recA* mutant strains could not be transformed with the construct, suggesting both strains share a similar capacity to induce double-strand breaks. A plasmid construct (pFC105) with improved stability was obtained by transferring the pGS905 insert into pET-24a. As with studies on *B. subtilis* RecU, these clones were transferred into *E. coli ruv* mutant strains to see whether they could improve cell survival following UV irradiation.

The experiments revealed that phage bIL67 RuvC can suppress the UV sensitive phenotype of *ruvA* and *ruvABC* mutants (Figure 8.5). Further analysis, carried out by GJ Sharples, confirmed that pFC105 could restore UV resistance in a *ruvAB* mutant background (Curtis, 2004). Thus 67RuvC can function as a genuine Holliday junction resolvase *in vivo*, although unlike its bacterial counterpart it can do so independently of the RuvAB branch migration complex. 67RuvC failed to improve the UV repair defect in *ruvB* (Figure 8.5) or *ruvC* (Curtis, 2004) mutant strains, suggesting that the presence of RuvA protein may block access of 67RuvC to the junction. A similar, though less extreme, phenotype was detected with the *E. coli*

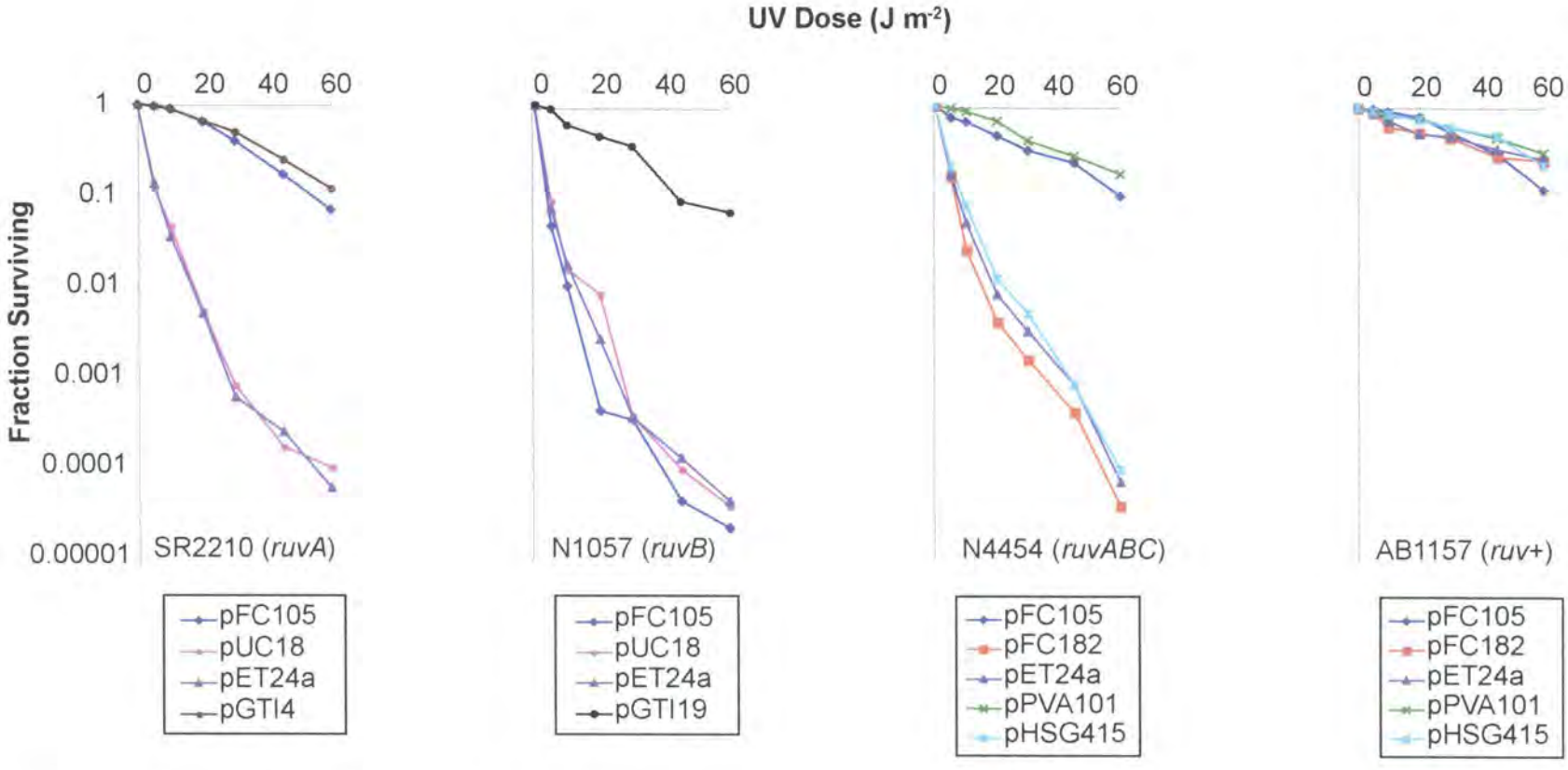

Figure 8.5. Survival of UV irradiated *E. coli ruv* mutants carrying *L. lactis* phage bIL67 *ruvC* gene constructs. Strain, genotype and a key to the plasmid symbols used are shown below each panel. Plasmids used were pFC105 (67RuvC in pET-24a), pFC182 (67RuvC D8N in pET-24a), pGTI4 (*Ec*RuvA in pUC18), pGTI19 (*Ec*RuvB in pUC18), and pPVA101 (*Ec*RuvABC in pHSG415).

RusA Holliday junction resolvase in *ruvB* and *ruvC* mutant strains (Mandal *et al.*, 1993; Mahdi *et al.*, 1996). Holliday junction cleavage activity by 67RuvC was also demonstrated *in vitro* using the purified protein (Curtis, 2004).

To help determine whether *E. coli* RuvC and 67RuvC share a similar arrangement within the catalytic site, one of the four conserved residues within 67RuvC, Asp-8 (corresponding to Asp-7 in *E. coli* RuvC) was targeted by mutagenesis (Curtis, 2004). The resulting 67RuvC D8N mutant failed to improve the UV sensitive phenotype of an *E. coli ruvABC* mutant. This defect, coupled with the inability of the purified mutant protein to cleave branched DNA substrates (Curtis, 2004), revealed that the bacterial and phage RuvC proteins are likely to share a common catalytic core.

8.4 *Helicobacter pylori* RuvC is unable to suppress the UV sensitivity of an *E. coli ruvC* mutant

Helicobacter pylori is an important human pathogen responsible for chronic gastritis, peptic ulcer disease, gastric cancer and MALT (mucosa-associated lymphoid tissue) lymphoma (Suerbaum and Michetti, 2002) and infects approximately half of the worlds population (Dunn *et al.*, 1997). It is well adapted for colonisation and survival within gastric mucosa and is usually acquired in early childhood, persisting for years or even decades (Dunn *et al.*, 1997; Feldman, 2001).

H. pylori isolates are extremely diverse in their genetic structure, primarily though its ability to acquire DNA by natural transformation (Nedenskov-Sorensen *et al.*, 1990) and an unusually high rate of homologous recombination (Achtman *et al.*, 1999; Suerbaum *et al.*, 1998). These features are thought to contribute to rapid

switching of surface antigens and the ability to adapt to environmental changes within the stomach. Mutations in the DNA repair systems of other bacterial pathogens, such as *Salmonella sp* and *Listeria monocytogenes* are known to attenuate virulence (Cano *et al.*, 2002; Merino *et al.*, 2002). A potentially similar defect was observed in *H. pylori ruvC* (*Hp*RuvC) mutants which were unable to sustain infection in the mouse gut (Loughlin *et al.*, 2003). This raises the possibility that homologous recombination may be important for chronic infections by generating the diversity necessary to evade an immune response.

Mutations of the *ruvA*, *ruvB* or *ruvC* genes in *E. coli* results in a slight reduction in recombination frequency but a significant defect in DNA repair (Lloyd, 1991), whereas mutations in *H. pylori ruvC* result in significant defects in both DNA repair and recombination (Loughlin *et al.*, 2003). The *H. pylori* ortholog of RuvC possesses all the key resides known to be required for binding and resolution by *E. coli* RuvC. However, it lacks 16 amino acids from the C-terminus that are likely to contact RuvB and stabilise assembly of the RuvABC resolvasome complex. It may be that the RuvAB branch migration and RuvC cleavage steps are uncoupled in *H. pylori*. To investigate this possibility, we examined the effect of plasmids carrying *Hp*RuvC on the survival of *E. coli ruvC* mutants after exposure to UV light.

Plasmid constructs (pFC120 and pMP101) carrying the *H. pylori ruvC* gene were provided by FC Curtis and introduced into *E. coli* K12 strains GS2053 (*ruvC64::kan*), GS1922 (*ruvC::cat*) and wild-type AB1157 (Sharples *et al.*, 1990; Lloyd, 1991). The *H. pylori* RuvC endonuclease was unable to restore UV resistance to either of the *E. coli ruvC* mutants (Figure 8.6). There are several possible explanations for these results. It is may be that *Hp*RuvC cannot interact with the

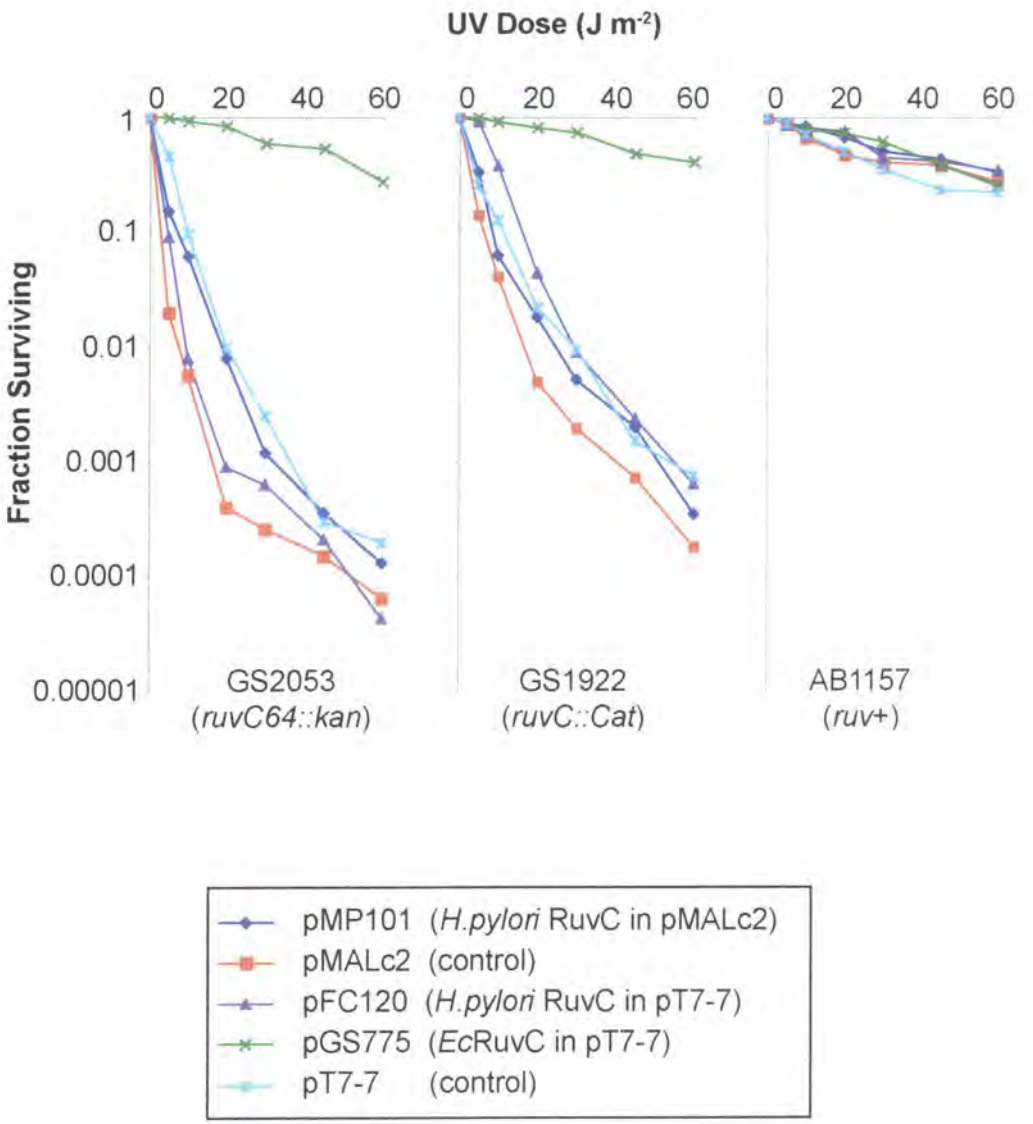

Figure 8.6. Survival of UV-irradiated *E. coli ruv* mutants carrying *H. pylori ruvC* gene constructs. Strain and genotype are indicated below each panel. Plasmids and symbols used are shown below the figure.

RuvAB branch migration complex of *E. coli*, or it cannot assemble on Holliday junction DNA in their presence. Alternatively, it may require other *H. pylori* encoded recombinases to function effectively *in vivo*. Perhaps the most likely explanation is that there were simply problems with gene expression or protein solubility in *E. coli* cells. This latter possibility fits with the significant difficulties encountered during attempts to purify *Hp*RuvC for *in vitro* analysis (FA Curtis, personal communication). These negative results serve to highlight the limitations of these *in vivo* experiments and are not too surprising given the study of foreign enzymes out of their proper context.

8.5 Discussion

To further our understanding of Holliday junction processing, we initiated an *in vivo* study of three resolving enzymes from a Gram-positive bacterium, a Gram-positive phage and an important Gram-negative pathogen. Plasmids carrying each gene were tested for their ability to restore UV resistance in *E. coli* strains lacking different components of the well-characterised RuvABC recombination system.

B. subtilis RecU, which cleaves Holliday junctions *in vitro* (Ayora *et al.*, 2004), conferred UV resistance upon an *E. coli ruvC* mutant, but was unable to complement *ruvA*, *ruvB* or *ruvAB* strains. The results suggest that *B. subtilis* RuvAB branch migration proteins are required to target the RecU resolvase, much as in the *E. coli* RuvABC system. It seems likely that the combined activity of all three proteins at Holliday junctions is necessary for recombinational repair in *B. subtilis*. Support for such a RuvAB-RecU resolvasome model has been obtained recently. Mutations in *B. subtilis recU* and *ruvAB* genes render cells extremely sensitive to

DNA-damaging agents, as with *E. coli ruv* mutants (Sanchez *et al.*, 2005). In addition, the recruitment of RecU to dsDNA breaks in damaged cells is dependent upon the presence of RuvAB (Sanchez *et al.*, 2005).

Phage bIL67 RuvC resolves model Holliday junction and fork DNA substrates with almost equal efficiency, unlike *E. coli* RuvC, which preferentially cleaves Holliday junctions (Curtis, 2004). The sequence specificity of 67RuvC also differs from *E. coli* RuvC, cleaving preferentially at the consensus 5'-T$\downarrow^{A}/_{G}$ rather than 5'-TT$\downarrow$, although both proteins do prefer to cleave 3'of a thymidine residue (Curtis, 2004). In this chapter, we described the ability of plasmids carrying 67RuvC to suppress the UV sensitive phenotype of *E. coli ruv* mutants. 67RuvC was able to restore UV resistance to *ruvA*, *ruvAB*, and *ruvABC* mutant strains, confirming that the phage protein mediates Holliday junction resolution *in vivo*. The combination of *in vitro* and *in vivo* analyses gave us a greater understanding of how the structure-specificity of a bacterial Holliday junction resolvase might evolve to suit the requirements of phage recombination.

Homologous recombination may play a crucial role in the diversity and adaptability of the human pathogen *H. pylori* and mutation of the *ruvC* gene resulted in an inability to maintain chronic infections in mice (Loughlin *et al.*, 2003). However, our attempts to demonstrate that *H. pylori* RuvC could function as a Holliday junction resolvase, by suppressing the UV sensitivity of an *E. coli ruvC* mutant, were unsuccessful. Further studies are therefore required to unravel the mechanics of junction resolution in *H. pylori*

Although this particular study forms only a small part of the ongoing studies into enzymes involved in Holliday junction resolution, it serves to highlight the

importance of these proteins in a variety of organisms. It is vital that recombination intermediates formed during replication fork regression or by homologous recombination are eliminated prior to cell division. The mechanism of branch resolution has been dissected in considerable detail using *E. coli* as a model organism. However the data presented here reveals the importance of having more than one model system to evaluate the mechanics of DNA repair/replication and recombination.

Chapter 9

Final discussion

Bacteriophages contribute to the evolution of new infectious diseases by mediating the dissemination of virulence factors within and between bacterial populations. Phage recombinases are able to induce genomic rearrangements at sites of limited sequence homology, leading to the acquisition of new genetic material. If the appropriation of novel properties confers a selective advantage to the host, the genes may be retained and transmitted to subsequent generations. The molecular mechanisms of phage recombination have received surprisingly little attention, despite the fact that an understanding of illegitimate exchange would help explain how phage facilitate the transmission of virulence genes and contribute to the emergence of new pathogenic bacteria. Of the known temperate phages, λ is the best characterised and its relatives frequently carry important virulence determinants. It therefore presents an ideal model system in which to study the mechanisms that promote bacteriophage genome rearrangements.

The work presented in this thesis describes the purification and biochemical analysis of the Orf recombination protein from phage λ and two homologs, *E. coli* DLP12 Orf151 and *S. aureus* phage ϕETA Orf20 (ETA20). The purpose of these studies was to elucidate how the 17 kDa Orf protein can substitute for a complex of three much larger *E. coli* proteins (RecF, RecO and RecR), and to further our understanding of how it functions in the initial phases of genetic exchange. In this chapter we summarise the major findings of this study and suggest experiments to address some of the unanswered questions concerning the function of Orf.

9.1 A family of phage Orf proteins

The primary sequence of λ Orf was used to search databases for homologous proteins. Multiple matches sharing significant sequence identity were identified and all were associated with phages, especially among the lambdoid group. More divergent Orf-like proteins were also found and two of these, *E. coli* cryptic prophage DLP12 Orf151 and *S. aureus* phage φETA Orf20 (ETA20), were shown to share significant secondary structure similarities to λ Orf. The genomic organisation of phages containing Orf homologs was investigated and revealed that the gene order was preserved in many instances. The *orf* genes were located between replication and lysis functions and upstream of alternative Holliday junction resolvases supplied by either Rap or RusA. This data provided additional support for the similarities identified by sequence homology and suggests the Orf protein is functionally important in phage biology.

9.2 Orf belongs to a group of recombination/replication mediators

To study the activities of λ Orf and the two potential orthologs, *E. coli* DLP12 Orf151 and *S. aureus* φETA Orf20 (ETA20) were purified to apparent homogeneity. λ Orf was successfully purified in its native form, as a histidine-tagged protein and as an MBP-fusion protein, and we demonstrated it exists as a dimer in solution. The Orf homologs, Orf151 and ETA20, were also successfully purified in milligram quantities from histidine-tagged and MBP-fusion expression constructs. Orf151, like Orf, exists as a dimer in solution. In contrast, ETA20 forms multimeric complexes, indicating its divergence from Orf.

Biochemical studies revealed functional similarities between Orf and Orf151; both preferentially bind ssDNA and interact with SSB in the presence of ssDNA. ETA20 failed to show any DNA binding activity; instead it displayed a non-specific nuclease function absent from both Orf and Orf151. Although further studies are required, this nuclease activity is consistent with the presence of a C-terminal HNH domain extension in ETA20. While ETA20 may possess an Orf-like DNA binding domain and participate in the early stages of phage genetic exchange, it appears not to be a true otholog of λ Orf.

The Orf protein has previously been shown to substitute for the *E. coli* RecFOR complex in λ-by-λ recombination but not during host conjugational exchange (Sawitzke and Stahl, 1992; Sawitzke and Stahl, 1994). However, Orf can partially substitute for RecFOR in *E. coli* recombination when Exo and Bet are present and partially suppresses the UV sensitive phenotype of *ruv*AB and *ruv*C mutants (Poteete, 2004). This latter activity could be due to Orf encouraging a repair pathway that does not result in Holliday junction formation. The RecFOR proteins belong to a structurally diverse group of recombination/replication mediator proteins (RMP's) that includes phage T4 UvsY and eukaryotic Rad52 proteins (Beernink and Morrical, 1999). All of these proteins function by helping strand exchange enzymes overcome the physical barrier presented by ssDNA binding proteins such as SSB and RPA. They mediate assembly of the required recombinase onto ssDNA in a species-specific manner enabling presynaptic filament formation and thus allowing recombination to proceed.

The prototype RMP is the UvsY accessory protein of phage T4. It facilitates the assembly of the UvsX recombinase onto ssDNA, overcoming the block presented

by the Gp32 ssDNA binding protein. UvsY binds preferentially to ssDNA and interacts with both the UvsX recombinase and Gp32. The Gp32 protein, like SSB, binds cooperatively and with high affinity to ssDNA, eliminating the inhibitory secondary structures. However, the competitive effects of its binding to the ssDNA lattice interfere with UvsX loading, in a similar manner to SSB inhibiting RecA binding (Griffith and Formosa, 1985; Kowalczykowski *et al.*, 1987). It is thought that the UvsY protein eliminates this competition by binding to the Gp32-ssDNA complex and weakening it, enabling UvsX to bind, allowing filament formation and displacement of Gp32. The protein has been shown to stabilise the UvsX-ssDNA interaction and destabilise Gp32-ssDNA interactions (Kodadek *et al.*, 1989; Sweezy and Morrical, 1999). The mechanism of UvsY presynaptic filament stabilisation is unknown; it may induce a conformational change in the ssDNA causing an increase in the affinity of UvsX, or it may simply hold the complex together by interacting with both ssDNA and UvsX (Sweezy and Morrical, 1997; Yassa *et al.*, 1997). The most recent studies indicate that UvsY interactions with ssDNA play a major role in its stabilisation of presynaptic filaments, consistent with the theory that it functions by organising the ssDNA lattice into a favourable conformation for interaction of UvsX with ssDNA (Liu *et al.*, 2006).

The RecFOR complex stimulates the assembly of RecA filaments onto gapped SSB-coated ssDNA, overcoming the rate limiting step of presynaptic filament formation, thereby accelerating DNA strand exchange (Morimatsu and Kowalczykowski, 2003). RecO binds preferentially to ssDNA and interacts with SSB, while the RecOR complex stabilises RecA-ssDNA complexes (Shan *et al.*, 1997). These proteins clearly function in an analogous manner to the UvsY, UvsX

and Gp32 components of bacteriophage T4, and are therefore classified as members of the RMP family (Beernink and Morrical, 1999).

Orf and Orf151 proteins shared remarkably similar biochemical properties to each other, displaying a preference for binding to substrates containing a ssDNA component rather than those consisting only of dsDNA. Both proteins also associated with *E. coli* SSB protein, however they required strict conditions for recognition, the presence of ssDNA being required for complex formation. This suggests the proteins require exposure of concealed SSB epitopes, available only when SSB is bound to ssDNA. Regardless of this requirement, the Orf/Orf151-SSB-ssDNA tripartite assemblies are analogous to those formed between UvsY, Gp32 and ssDNA in the phage T4 system and are consistent with the activities of the RecFOR complex. Orf and Orf151 should therefore be considered as members of the recombination/ replication mediator protein family.

Interactions between Orf and β protein could not be authenticated during this study and further experimentation is required to elucidate whether RecA or β protein is the preferred recombinase partner for Orf.

Formation of complexes between RMP proteins and their cognate SSB bound to ssDNA appears to be a universal strategy for facilitating the loading of enzymes that would otherwise be inhibited by SSB (Beernink and Morrical, 1999). A common role in alleviating the SSB/Gp32 barrier during phage recombination is indicated by the similarities between the action of Orf/Orf151 and T4 UvsY. The widespread conservation of Orf and the functional similarity displayed by one of its very distant relatives (Orf151) highlights the importance of overcoming the barrier presented by SSB proteins during lambdoid phage recombination.

9.3 The Orf DNA binding site

In Chapter 6 we described the DNA binding properties of Orf and Orf151. We discovered that both bind preferentially to ssDNA over dsDNA and form complexes with gapped substrates and oligonucleotides containing 3' or 5' ssDNA overhangs. The crystal structure of Orf indicated two possible DNA binding regions, a positively charged, tapered central channel, and a shallow U-shaped cleft that traverses the top of the central cavity and is lined with positively charged residues. The central channel ranges in diameter from ~20Å at the top to 8Å at the bottom, and it is thought that this hole could accommodate ssDNA throughout and dsDNA at the widest end (Maxwell *et al.*, 2005). The lack of preference for binding to gapped and tailed substrates by Orf suggests that the protein interacts with ssDNA regions via the surface cleft, as dsDNA cannot be accommodated throughout the central channel. However, until we can prove this theory, either by crystallising Orf bound to DNA, or by site-directed mutagenesis of residues in the proposed binding sites, we cannot be certain that DNA is not bound within the central hole.

If the central cavity is the DNA binding site of Orf we have to account for its ability to assemble on gapped duplex substrates, given that the diameter of DNA double helix should preclude aperture penetration. The simplest explanation is that the Orf dimer is a hinge protein, able to open in order to bind at the ssDNA-dsDNA junction. The electrostatic potential calculations show that the interior of the central channel is extremely positively charged (Figure 9.1A) due to several conserved residues (Lys3, Arg41, Lys73 and Lys81; Figure 9.1B; see also Figure 3.2A). This feature is supportive of a model where ssDNA threads through the central channel. The residues Trp50 and Asn46 project from opposite sides of each monomer into the

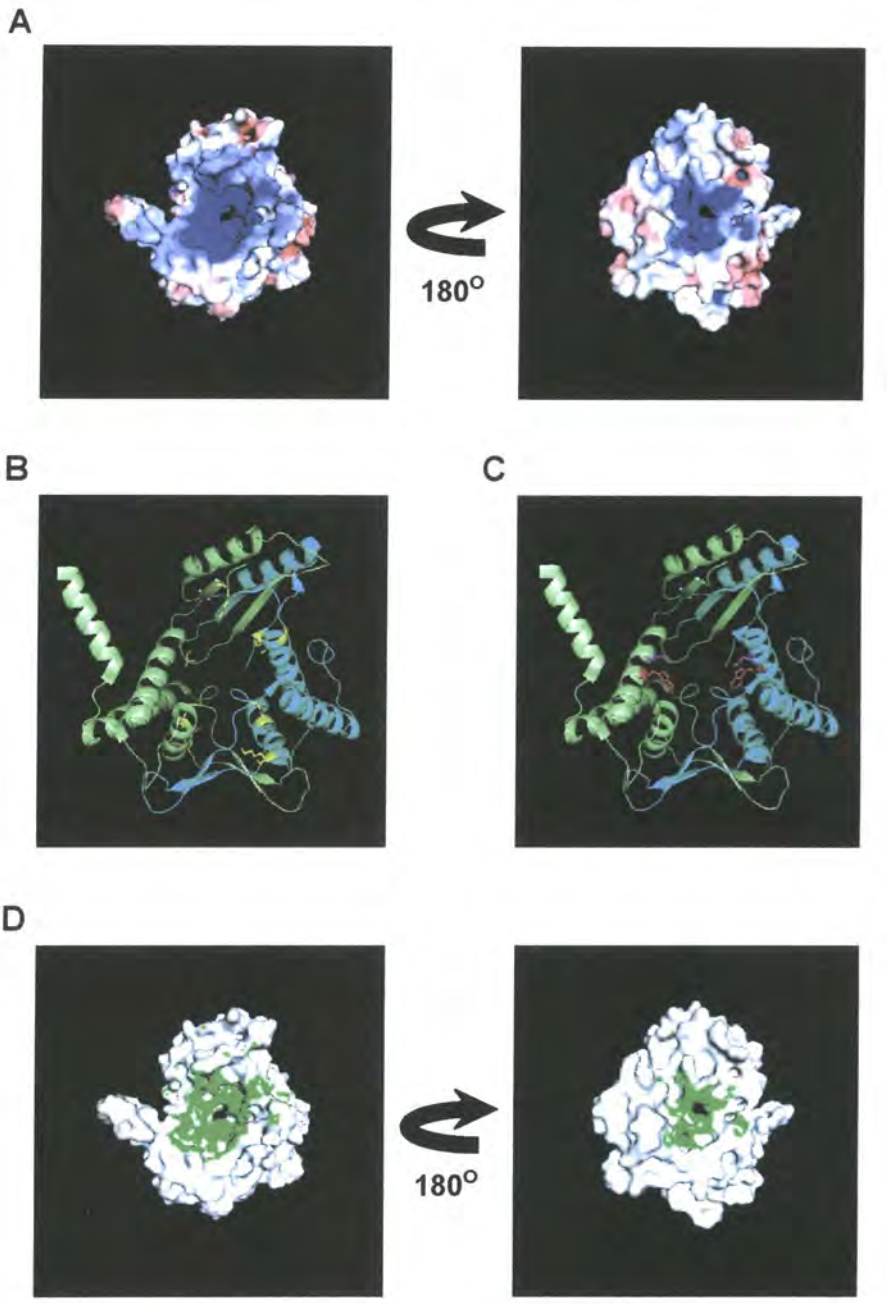

Figure 9.1. Potential Orf DNA-binding sites.
(A) Electrostatic surface potential of the Orf dimer. The surface is coloured according to electrostatic potential, -0.1 V (red) to 0.1 V (blue). A view of the potential DNA binding site looking through the central cavity from the top (left) and bottom (right) is shown. **(B)** Ribbon diagram of the Orf dimer, with monomer A shown in green and monomer B in blue. The conserved basic residues (Lys3, Arg41, Lys73 and Lys81) responsible for the positively charged central channel are shown in yellow. **(C)** Ribbon diagram of the Orf dimer with Tryptophan 50 (red) and Asparagine 46 (pink) residues highlighted. **(D)** Molecular surface representation of the Orf dimer with the predicted dsDNA-binding surface highlighted in green, viewed from the top (left), and bottom (right) of the central channel.

cavity and could interact with nucleotide bases to stabilise interactions with DNA (Figure 9.1C). Using a DNA binding site prediction server (University of Osaka; http://pre-s.protein.osaka-u.ac.jp/~preds/), which calculates potential dsDNA-binding surfaces, Orf was considered to be a probable dsDNA binding protein, interacting with dsDNA in the areas indicated in Figure 9.1D. The prediction method generates the molecular surface of the protein, calculates the electrostatic potential and uses 63 protein-dsDNA complexes determined by X-ray crystallography to predict the dsDNA-binding sites on protein surfaces (Tsuchiya *et al.*, 2004). Orf showed limited binding to the duplex substrates in band shift assays, however it did bind to gapped and tailed substrates containing duplex regions. It is possible that Orf threads ssDNA through the central cavity and interacts with the dsDNA region of ssDNA-dsDNA junctions via the widest part of the cavity opening (Figure 9.2).

However, this still leaves unanswered the question of how Orf loads onto a gapped duplex substrate. The N-terminal region of the Orf dimer consists of two α-helices (αA from each monomer) and four β-strands (β1 and β2 from each monomer) that form an intertwined β-sheet (Figure 9.3A). Calculation of the position of hydrogen bonds within the Orf dimer was performed using Swiss-Pdb viewer (GlaxoSmithKline; www.expasy.org/spdbv/). Hydrogen bonding in this region is extensive, within and between the monomers. In contrast the larger C-terminal region is composed of two three-helix bundles (αB, αC and αD) and two anti-parallel β-sheets (β3 and β4) connected by long loops (Figure 9.3A). Hydrogen bonding within the distal region of the C-terminal (αC, β3 and β4) is less extensive, and any bonding is between strands and helices within a single monomer. Asymmetry in the Orf dimer is due to a twist in the backbone at residues Asn40-

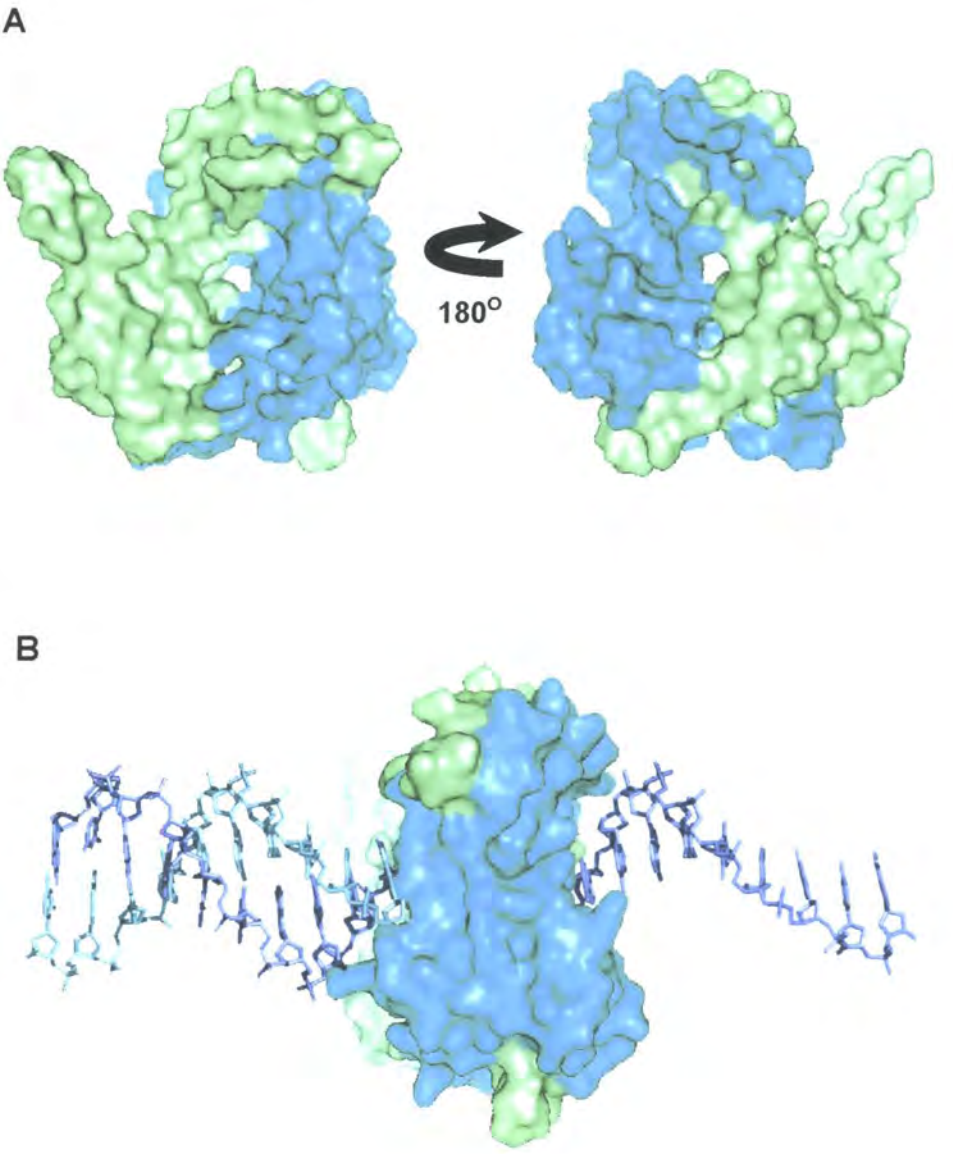

Figure 9.2. Model of Orf binding at the dsDNA-ssDNA junction.
(A) Surface representation of the ~20A wide end (left) and the ~8A narrow end (right) of the central cavity in the Orf dimer. Monomer A is shown in green and monomer B in blue. **(B)** Model for the interaction of Orf with DNA. The dsDNA region of the dsDNA-ssDNA junction is accomodated in the mouth of the central opening and ssDNA is threaded through the channel, protruding out of the other side.

Ser42, the hinge point involving residues Asn40 and Arg41 (Figure 9.3A). We propose that the stable packing and strong hydrogen bonding within the N-terminal asymmetric region of the Orf dimer allows it to act as a hinge, allowing for opening of the dimer in the C-terminal region between the antiparallel β-sheets (Figure 9.3B). In this model the dimer bends in the region between β2 and αB, possibly at residues Asn40 and Arg41, allowing dissociation of the C-terminal β3 and β4 from each monomer, opening the central channel (Figure 9.3B). The extended conformation of the αE helix in monomer A, and the disorder observed at the C-terminus of monomer B suggests that the last 20 residues are flexible and could adopt different conformations. It is possible that these helices stack against the N-terminal region of the dimer, offering stabilisation of the open conformation.

This model could account for Orf binding to gapped substrates and those with ssDNA overhangs with equally affinity and help explain the observed weak/unstable binding to dsDNA. In the 'open' conformation, Orf could associate with duplex DNA *via* the central cavity, but the interaction would be unstable due to the lack of a ssDNA substrate to close around. The highly positively charged nature of the central cavity certainly suggests that this region is involved in DNA binding and fits well with this model. Our proposed 'hinge protein' model is also in agreement with the predicted dsDNA binding regions of Orf. While this model can offer certain explanations we cannot be sure whether Orf binds DNA via the central channel or within the surface cleft without further experimental evidence.

Many DNA binding proteins interact with the negatively charged phosphodiester backbone of DNA via positively charged residues. Orf contains 19 lysine or arginine residues per subunit, most of which lie in the putative binding cleft

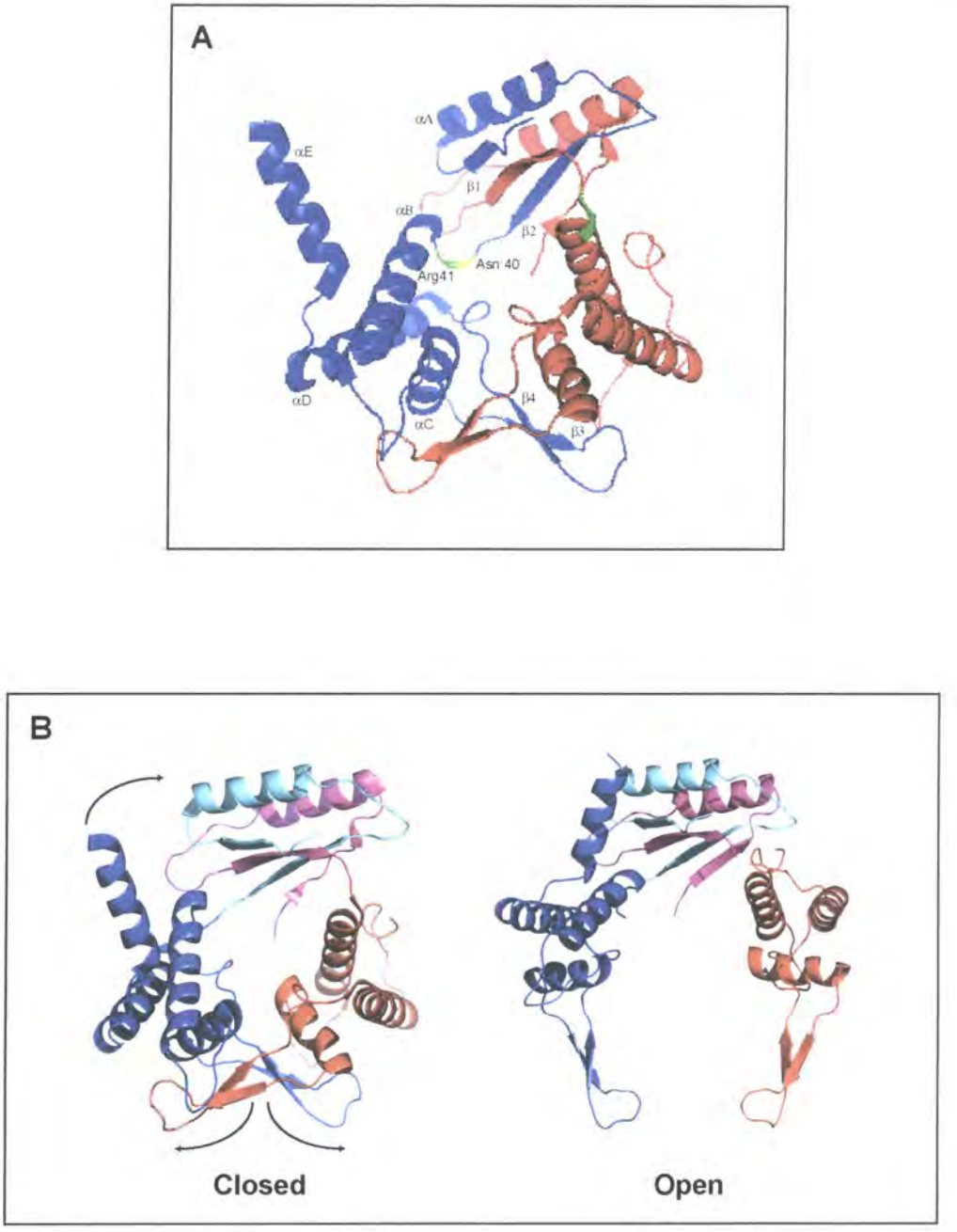

Figure 9.3. Model for opening of the Orf dimer.
(A) Ribbon diagram of the Orf dimer. Monomer A is coloured blue and monomer B is red. The secondary structure elements are labelled for monomer A and the proposed 'hinge point' residues asparagine 40 (yellow) and arginine 41 (green) are highlighted. **(B)** Ribbon representation of the 'closed' (left) and 'open' (right) conformations of the Orf dimer. Monomers are coloured as for A, the N-terminal 'hinge' regions of each monomer are shaded; monomer A - cyan, monomer B - pink. Arrows indicate the proposed movement of the C-terminal regions. Figure created using MacPyMol.

or line the walls of the central cavity. Seven of these residues are conserved in alignments of closely related lambdoid phages, five of which are in the proximity of the putative binding domains. Residues 48 and 132 (Lys48, Arg132) are positioned within the cleft, Lys73 and Arg41 project into the central cavity and Lys3 is in a position that indicates it could interact with DNA in either position. Mutation of these residues could provide valuable information about the location of the DNA binding site. The residues Trp50 and Asn46 protrude into the central cavity and could stabilise interactions with DNA. If proteins mutated in these residues displayed an impaired ability to bind DNA, this would provide support for the hinged protein model with ssDNA passing through the channel.

9.4 A model for the role of Orf in phage recombination

The λ Orf protein can provide a function analogous to that of RecFOR in λ recombination. RecO interacts with ssDNA and SSB and RecFOR act in a concerted manner at ssDNA-dsDNA junctions to facilitate the loading of RecA onto SSB-coated DNA. We have shown that Orf binds ssDNA and gapped duplex substrates and associates with SSB bound to ssDNA. However, we do not know the precise role of Orf in genetic recombination and can only provide a basic model for its function in the early stages of the Red-recombination pathways (Figure 9.4).

When a dsDNA break occurs in the λ genome, the dsDNA end is processed by Exo to generate a 3' ssDNA tail. The highly processive nature of Exo and the small size (48.5 kb) of the phage genome suggests that the activity of Exo is modulated, possibly by the β protein (Matsuura *et al.*, 2001; Cassuto *et al.*, 1971; Cassuto and Radding, 1971). The dsDNA end is probably converted to a β protein coated 3'-

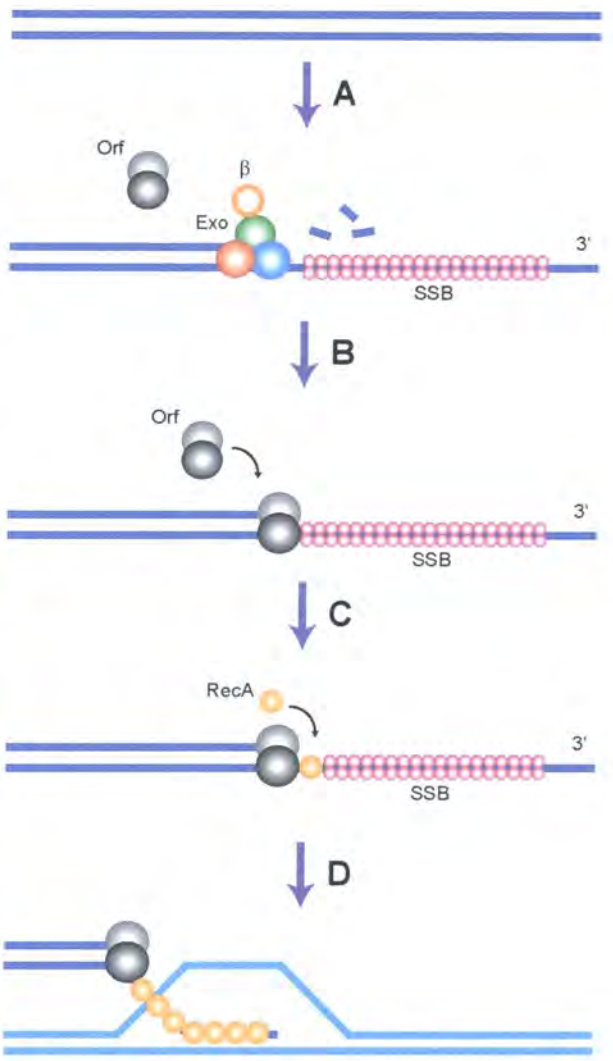

Figure 9.4. Model for the role of Orf in the Red-mediated strand invasion pathway. **(A)** Lambda exonuclease recognises the free dsDNA end and begins digestion of a single strand in the 5'-3' direction, producing 3'-tailed ssDNA which is quickly bound by SSB. The β protein is thought to interact with Exo and modulate its activities. **(B)** Orf recognises and binds the SSB-ssDNA complex at the dsDNA-ssDNA junction. **(C)** Orf facilitates the loading and subsequent polymerisation of RecA on the ssDNA, displacing SSB. **(D)** The RecA-coated ssDNA invades an intact homologous duplex, followed by annealing of the complementary strands. *E. coli* RuvABC or λ structure-specific endonuclease, Rap, action is required to process the branched intermediates generated.

ended ssDNA by the concerted action of Exo and β, consistent with their known interaction and DNA binding properties of β protein (Radding, 1971; Muniyappa and Radding, 1986; Kmiec and Holloman, 1981). It is possible that a monomer of β is associated with Exo and β protein immediately loads on the free ssDNA exposed by Exo. Assembly of additional β subunits to form a multimeric ring may halt Exo progression. This could explain the ability of recombination in λ to occur at short tracts of homology and to occur close to the initiating dsDNA break. If a complementary ssDNA is available, recombination will occur via the strand-annealing pathway; β anneals the homologous strands to form a splice recombinant molecule. It is possible that Orf function is not required for this pathway, perhaps explaining why it is not essential for Red pathway recombination. It has been shown that λ phage lacking Orf can still recombine via the Red pathway, though recombinational exchanges occur further away from the dsDNA break (Tarkowski *et al.*, 2002). However, Orf may function in some way to assist binding of β onto the ssDNA, perhaps interacting at the SSB-ssDNA interface to serve as a nucleation site for β. Orf may modulate the Exo/β complex ensuring minimal degradation of the dsDNA by Exo, promoting efficient binding of β onto the ssDNA and thereby acting to focus recombination events close to the initiating dsDNA break. In the absence of a homologous ssDNA, λ recombination will proceed via the RecA-dependent strand-invasion pathway. It is not known whether β protein participates in this pathway; if it is loaded onto the ssDNA as in the annealing pathway, this would suggest that RecA has to displace β for the presynaptic filament to form. Although β protein can promote strand exchange under certain circumstances, it has not been shown to promote the duplex strand invasion reactions seen with RecA (Li *et al.*, 1998). The

requirement for RecA in the strand invasion pathway of λ recombination fits with β being restricted to the annealing pathway (Stahl *et al.*, 1997). It is possible therefore that Orf is required to facilitate the polymerisation of RecA onto β-coated ssDNA. If β protein is not bound to the ssDNA and instead it is coated with SSB, Orf may bind to the SSB-ssDNA interface and weaken the interaction, in a way analogous to the action of UvsY in the phage T4 system, allowing for binding and subsequent polymerisation of RecA. The Orf protein could also modulate the activity of SSB, helping RecA displace SSB or compete with β protein for ssDNA binding.

Many questions remain unanswered concerning how the activities of phage and host enzymes are coordinated and how they function in the molecular processes leading to genomic rearrangements in lambda and lambdoid phages. The current model proposed here cannot incorporate all possible proteins involved, or indicate their activities with any great certainty. The new model (Figure 9.4) can only be slightly modified from that shown in the introduction (Figure 1.8), by using the experimental evidence gathered during this study. It depicts Orf functioning in the Red-mediated strand invasion pathway by recognising the ssDNA-dsDNA junction, created from a dsDNA end by Exo nuclease action and its association with the SSB-ssDNA complex. Assuming Orf supplies a function analogous to that of the *E. coli* RecFOR proteins, it will facilitate loading of RecA onto the SSB-coated ssDNA. Whether β protein is present only to modulate Exo activity, or is bound to the ssDNA and thus requires removal prior to loading of RecA is unknown. To simplify the diagram, β is indicated solely as a regulatory factor for Exo (Figure 9.4).

9.5 Future directions

We have performed a basic biochemical analysis of the Orf protein and have uncovered properties that clearly fit with a role in the initial steps of genetic exchange. Further work is needed to establish whether bacterial or phage recombination pathways are stimulated by Orf, and to elucidate its precise action. The order of assembly of λ and host recombination proteins onto ssDNA and the precise role of these proteins in the strand-annealing and strand-invasion pathways is not yet fully understood. Many questions remain unanswered and further investigation is essential for the unravelling of the mechanisms of phage genetic exchange in the context of the bacterial cell. The future directions envisioned for this work and the reasoning behind these proposals are outlined below.

(i) Defining the DNA binding site

Site-directed mutagenesis of the residues predicted to be involved in Orf:DNA interactions could provide evidence for binding within the central cavity or along the surface cleft. Site-directed cross linking of the C-terminal region of Orf (between β3 and β4 of different monomers) to prevent opening of the proposed 'hinge' region could also help prove this model. Binding of Orf to gapped substrates followed by circularisation of the DNA could help indicate whether Orf opens to bind around the ssDNA, threads the ssDNA through the central hole like a bead on a string, or binds DNA within the surface cleft. Complexes could be visualised by electron microscopy. The crystal structure of Orf bound to a short oligonucleotide or gapped DNA substrate would be extremely useful in furthering our understanding of how Orf interacts with DNA.

(ii) DNA binding specificity of Orf

Recombination junctions in the genomes of many lambdoid phages tend to occur at purine-rich ribosome binding sites (Juhala *et al.*, 2000). Investigations, using synthetic oligonucleotides, into the preference of Orf for binding to purine or pyrimidine-rich sequences could be investigated using fluorescence quench assays. If a preference is detected, methods such as "SELEX' could be employed to identify the specific nucleotide sequence elements recognised (Pollock and Treisman, 1990).

(iii) Protein:protein interactions between Orf and other λ or host proteins

Far western blotting can be employed to investigate the interaction between Orf with RecA, Exo, β, and other lambda encoded proteins such as Ea10, a putative λ SSB analogue, and Ea47, a DNA-binding protein from lambda, shown to interact weakly with Orf in yeast-two hybrid studies. Inclusion of a guanidinium hydrochloride denaturation-renaturation cycle may improve the folding of blotted proteins and potentially enhance interactions with Orf.

Affinity chromatography could be utilised to identify interacting partners, binding His- or MBP-tagged Orf to nickel agarose or amylose matrices as described in Chapter 7. Untagged proteins such as Exo, β and RecA could be passed across the pre-bound matrix, and co-elution would indicate a positive interaction. Performing these experiments, and repeating those attempted with SSB (Chapter 7), in the presence of DNA could help us determine whether joint assembly on DNA enhances protein:protein association.

Size exclusion chromatography could be used as described in Chapter 7 for investigating other possible partnerships involving Orf. Experimentation with the salt concentration and incubation conditions may be required to find the ideal conditions

for the formation of protein:protein complexes. DNA can be incorporated into the protein mix as mentioned previously to establish co-assembly of interacting protein partners upon substrates resembling those likely to be encountered during the initial stages of genetic exchange.

(iv) Orf-SSB interaction

To gain further insights into the domains involved in the Orf:SSB interactions, the OrfΔC6 and OrfΔC19 C-terminal deletion mutant proteins should be investigated for their ability to bind DNA and SSB utilising the approaches outlined above. The SSB C-terminal deletion mutants SSBΔC10 and SSB113 can be used to examine the loss of interaction in a reciprocal manner. The C-terminus of SSB has previously been shown to be important for association with other proteins involved in DNA replication and repair (Curth *et al.*, 1996; Meyer and Laine, 1990).

9.6 Concluding remarks

This study has provided us with new information about a phage protein involved in the early stages of genetic exchange. Precisely how this protein functions remains unclear, in part due to the limited knowledge about its interacting partners and the molecular mechanisms involved in phage recombination. However we have purified a novel protein, discovered its multimeric status and uncovered some interesting biochemical activities in keeping with its predicted role in recombination. The knowledge gained will hopefully provide a solid foundation for the design of further experiments, aimed at increasing our understanding of how phage enzymes collaborate with the systems of the host to drive the genome rearrangements that ultimately spawn new pathogenic bacteria.

References

Achtman, M., Azuma, T., Berg, D.E., Ito, Y., Morelli, G., Pan, Z.J., Suerbaum, S., Thompson, S.A., van der Ende, A. and van Doorn, L.J. (1999) Recombination and clonal groupings within *Helicobacter pylori* from different geographical regions. *Mol Microbiol*, **32**, 459-70.

Alonso, J.C., Steige, A.C., Dobrinski, B. & Lurz, R. (1993) Purification and properties of the RecR protein from *Bacillus subtilis* 168. *J Biol Chem*, **268**, 1424-1429.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol*, **215**, 403-10.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*, **25**, 3389-402.

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

Aravind, L., Makarova, K.S. and Koonin, E.V. (2000) SURVEY AND SUMMARY: holliday junction resolvases and related nucleases: identification of new families, phyletic distribution and evolutionary trajectories. *Nucleic Acids Res*, **28**, 3417-32.

Ayora, S., Carrasco, B., Doncel, E., Lurz, R. and Alonso, J.C. (2004) *Bacillus subtilis* RecU protein cleaves Holliday junctions and anneals single-stranded DNA. *Proc Natl Acad Sci U S A*, **101**, 452-7.

Bachmann, B.J. (1972) Pedigrees of some mutant strains of *Escherichia coli* K-12. *Bacteriol Rev*, **36**, 525-57.

Baranska, S., Gabig, M., Wegrzyn, A., Konopa, G., Herman-Antosiewicz, A., Hernandez, P., Schvartzman, J.B., Helinski, D.R. and Wegrzyn, G. (2001) Regulation of the switch from early to late bacteriophage lambda DNA replication. *Microbiology*, **147**, 535-47.

Barondess, J.J. and Beckwith, J. (1990) A bacterial virulence determinant encoded by lysogenic coliphage lambda. *Nature*, **346**, 871-4.

Barondess, J.J. and Beckwith, J. (1995) *bor* gene of phage lambda, involved in serum resistance, encodes a widely conserved outer membrane lipoprotein. *J Bacteriol*, **177**, 1247-53.

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

Bidnenko, E., Ehrlich, S.D. and Chopin, M.C. (1998) *Lactococcus lactis* phage operon coding for an endonuclease homologous to RuvC. *Mol Microbiol*, **28**, 823-34.

Bishai, W.R., Rappuoli, R. and Murphy, J.R. (1987) High-level expression of a proteolytically sensitive diphtheria toxin fragment in *Escherichia coli*. *J Bacteriol*, **169**, 5140-51.

Blackwood, E.M. and Eisenman, R.N. (1991) Max: a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with Myc. *Science*, **251**, 1211-7.

Blanar, M.A., Sandler, S.J., Armengod, M.E., Ream, L.W. and Clark, A.J. (1984) Molecular analysis of the *recF* gene of *Escherichia coli*. *Proc Natl Acad Sci U S A*, **81**, 4622-6.

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

Boehmer, P.E. and Emmerson, P.T. (1991) *Escherichia coli* RecBCD enzyme: inducible overproduction and reconstitution of the ATP-dependent deoxyribonuclease from purified subunits. *Gene*, **102**, 1-6.

Bork, J.M., Cox, M.M. and Inman, R.B. (2001) The RecOR proteins modulate RecA protein function at 5' ends of single-stranded DNA. *Embo J*, **20**, 7313-22.

Boyd, E.F. and Brussow, H. (2002) Common themes among bacteriophage-encoded virulence factors and diversity among the bacteriophages involved. *Trends Microbiol*, **10**, 521-9.

Braun, P., Hu, Y., Shen, B., Halleck, A., Koundinya, M., Harlow, E. and LaBaer, J. (2002) Proteome-scale purification of human proteins from bacteria. *Proc Natl Acad Sci U S A*, **99**, 2654-9.

Brooks, K. and Clark, A.J. (1967) Behavior of lambda bacteriophage in a recombination deficienct strain of *Escherichia coli*. *J Virol*, **1**, 283-93.

Buck, G.A., Cross, R.E., Wong, T.P., Loera, J. and Groman, N. (1985) DNA relationships among some *tox*-bearing corynebacteriophages. *Infect Immun*, **49**, 679-84.

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

Casjens, S., Winn-Stapley, D.A., Gilcrease, E.B., Morona, R., Kuhlewein, C., Chua, J.E., Manning, P.A., Inwood, W. and Clark, A.J. (2004) The chromosome of *Shigella flexneri* bacteriophage Sf6: complete nucleotide sequence, genetic mosaicism, and DNA packaging. *J Mol Biol*, **339**, 379-94.

Casjens, S.R., Gilcrease, E.B., Winn-Stapley, D.A., Schicklmaier, P., Schmieger, H., Pedulla, M.L., Ford, M.E., Houtz, J.M., Hatfull, G.F. and Hendrix, R.W. (2005) The generalized transducing Salmonella bacteriophage ES18: complete genome sequence and DNA packaging strategy. *J Bacteriol*, **187**, 1091-104.

Campbell, A. (1994) Comparative molecular biology of lambdoid phages. *Annu Rev Microbiol*, **48**, 193-222.

Cano, D.A., Pucciarelli, M.G., Garcia-del Portillo, F. and Casadesus, J. (2002) Role of the RecBCD recombination pathway in Salmonella virulence. *J Bacteriol*, **184**, 592-5.

Carr, K.M. and Kaguni, J.M. (2002) *Escherichia coli* DnaA protein loads a single DnaB helicase at a DnaA box hairpin. *J Biol Chem*, **277**, 39815-22.

Carter, D.M. and Radding, C.M. (1971) The role of exonuclease and beta protein of phage lambda in genetic recombination. II. Substrate specificity and the mode of action of lambda exonuclease. *J Biol Chem*, **246**, 2502-12.

Cassuto, E., Lash, T., Sriprakash, K.S. and Radding, C.M. (1971) Role of exonuclease and protein of phage lambda in genetic recombination. V.

Recombination of lambda DNA in vitro. *Proc Natl Acad Sci U S A*, **68**, 1639-43.

Cassuto, E. and Radding, C.M. (1971) Mechanism for the action of lambda exonuclease in genetic recombination. *Nat New Biol*, **229**, 13-6.

Chalker, A.F., Leach, D.R. and Lloyd, R.G. (1988) *Escherichia coli sbcC* mutants permit stable propagation of DNA replicons containing a long palindrome. *Gene*, **71**, 201-5.

Chan, S.N., Vincent, S.D. and Lloyd, R.G. (1998) Recognition and manipulation of branched DNA by the RusA Holliday junction resolvase of *Escherichia coli*. *Nucleic Acids Res*, **26**, 1560-6.

Chen, S.L., Hung, C.S., Xu, J., Reigstad, C.S., Magrini, V., Sabo, A., Blasiar, D., Bieri, T., Meyer, R.R., Ozersky, P., Armstrong, J.R., Fulton, R.S., Latreille, J.P., Spieth, J., Hooton, T.M., Mardis, E.R., Hultgren, S.J. and Gordon, J.I. (2006) Identification of genes subject to positive selection in uropathogenic strains of *Escherichia coli*: a comparative genomics approach. *Proc Natl Acad Sci U S A*, **103**, 5977-82.

Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T.J., Higgins, D.G. and Thompson, J.D. (2003) Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res*, **31**, 3497-500.

Chow, K.H. and Courcelle, J. (2004) RecO acts with RecF and RecR to protect and maintain replication forks blocked by UV-induced DNA damage in Escherichia coli. *J Biol Chem*, **279**, 3492-6.

Clark, A.J. and Margulies, A.D. (1965) Isolation and Characterization of Recombination-Deficient Mutants of *Escherichia Coli* K12. *Proc Natl Acad Sci U S A*, **53**, 451-9.

Clark, A.J. (1973) Recombination deficient mutants of *E. coli* and other bacteria. *Annu Rev Genet*, **7**, 67-86.

Clark, A.J. and Sandler, S.J. (1994) Homologous genetic recombination: the pieces begin to fall into place. *Crit Rev Microbiol*, **20**, 125-42.

Clark, A.J., Inwood, W., Cloutier, T. and Dhillon, T.S. (2001) Nucleotide sequence of coliphage HK620 and the evolution of lambdoid phages. *J Mol Biol*, **311**, 657-79.

Coffey, A. and Ross, R.P. (2002) Bacteriophage-resistance systems in dairy starter strains: molecular analysis to application. *Antonie Van Leeuwenhoek*, **82**, 303-21.

Courcelle, J., Crowley, D.J. and Hanawalt, P.C. (1999) Recovery of DNA replication in UV-irradiated *Escherichia coli* requires both excision repair and RecF protein function. *J Bacteriol*, **181**, 916-22.

Cox, M.M. and Lehman, I.R. (1987) Enzymes of general recombination. *Annu Rev Biochem*, **56**, 229-62.

Cox, M.M. (1998) A broadening view of recombinational DNA repair in bacteria. *Genes Cells*, **3**, 65-78.

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

Cox, M.M. (2001) Recombinational DNA repair of damaged replication forks in *Escherichia coli*: questions. *Annu Rev Genet*, **35**, 53-82.

Curth, U., Genschel, J., Urbanke, C. and Greipel, J. (1996) *In vitro* and *in vivo* function of the C-terminus of *Escherichia coli* single-stranded DNA binding protein. *Nucleic Acids Res*, **24**, 2706-11.

Curtis, F.A., Reed, P., Sharples, G.J. (2004) Evolution of a phage RuvC endonuclease for resolution of both Holliday and branched DNA junctions. *Molecular Microbiology*, **in p66**, 1322-1345.

Daniels, D.L., Schroeder, J.L., Szybalski, F.W., Sanger, F, Coulson, A.R., Hong, G.B., Hill, D., Petersen, G., and Blattner, F.R., in *Lambda II* R.W. Hendrix, J.W.R., F.W Stahl, and R.A. Weisberg, Ed. (Cold Spring Harbour, NY, New York, 1983), vol. II, pp. 519-676.

Davies, A.A. and West, S.C. (1998) Formation of RuvABC-Holliday junction complexes *in vitro*. *Curr Biol*, **8**, 725-7.

Deng, W., Burland, V., Plunkett, G., 3rd, Boutin, A., Mayhew, G.F., Liss, P., Perna, N.T., Rose, D.J., Mau, B., Zhou, S., Schwartz, D.C., Fetherston, J.D., Lindler, L.E., Brubaker, R.R., Plano, G.V., Straley, S.C., McDonough, K.A., Nilles, M.L., Matson, J.S., Blattner, F.R. and Perry, R.D. (2002) Genome sequence of Yersinia pestis KIM. *J Bacteriol*, **184**, 4601-11.

Dillingham, M.S., Spies, M. and Kowalczykowski, S.C. (2003) RecBCD enzyme is a bipolar DNA helicase. *Nature*, **423**, 893-7.

Dodd, I.B., Shearwin, K.E. and Egan, J.B. (2005) Revisited gene regulation in bacteriophage lambda. *Curr Opin Genet Dev*, **15**, 145-52.

Dunderdale, H.J., Benson, F.E., Parsons, C.A., Sharples, G.J., Lloyd, R.G. and West, S.C. (1991) Formation and resolution of recombination intermediates by *E. coli* RecA and RuvC proteins. *Nature*, **354**, 506-10.

Dunn, B.E., Cohen, H. and Blaser, M.J. (1997) *Helicobacter pylori. Clin Microbiol Rev*, **10**, 720-41.

Echolas, H. and Gingery, R. (1968) Mutants of bacteriophage lambda defective in vegetative genetic recombination. *J Mol Biol*, **34**, 239-49.

Echols, H., in *Lambda II* R.W. Hendrix, J.W.R., F.W Stahl, and R.A. Weisberg, Ed. (Cold Spring Harbour, New York, 1983) pp. 75-93.

Errami, M., Geourjon, C. and Deleage, G. (2003) Detection of unrelated proteins in sequences multiple alignments by using predicted secondary structures. *Bioinformatics*, **19**, 506-12.

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

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

Feldman, R.A., in *Helicobacter pylori: Molecular and Cellular Biology* Achtman, M., Suerbaum, S., Ed. (Horizon Scientific Press, Wymondham, 2001) pp. 29-51.

Fernandez, S., Ayora, S. and Alonso, J.C. (2000) *Bacillus subtilis* homologous recombination: genes and products. *Res Microbiol*, **151**, 481-6.

Franklin, N.C. (1967) Extraordinary recombinational events in *Escherichia coli.* Their independence of the rec+ function. *Genetics*, **55**, 699-707.

Friedman, D., and Gottesman, M., in *Lambda II* R.W. Hendrix, J.W.R., F.W Stahl, and R.A. Weisberg, Ed. (Cold Spring Harbour, New York, 1983) pp. 21-53.

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

Geourjon, C., Combet, C., Blanchet, C. and Deleage, G. (2001) Identification of related proteins with weak sequence identity using secondary structure information. *Protein Sci*, **10**, 788-97.

Gottesman, M. (1999) Bacteriophage lambda: the untold story. *J Mol Biol*, **293**, 177-80.

Griffin, T.J.t. and Kolodner, R.D. (1990) Purification and preliminary characterization of the *Escherichia coli* K-12 *recF* protein. *J Bacteriol*, **172**, 6291-9.

Griffith, J. and Formosa, T. (1985) The uvsX protein of bacteriophage T4 arranges single-stranded and double-stranded DNA into similar helical nucleoprotein filaments. *J Biol Chem*, **260**, 4484-91.

Habeeb, A.J. and Hiramoto, R. (1968) Reaction of proteins with glutaraldehyde. *Arch Biochem Biophys*, **126**, 16-26.

Haber, J.E. (1999) DNA recombination: the replication connection. *Trends Biochem Sci*, **24**, 271-5.

Hagan, N.F., Vincent, S.D., Ingleston, S.M., Sharples, G.J., Bennett, R.J., West, S.C. and Lloyd, R.G. (1998) Sequence-specificity of Holliday junction resolution:

identification of RuvC mutants defective in metal binding and target site recognition. *J Mol Biol*, **281**, 17-29.

Hanahan, D. (1983) Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol*, **166**, 557-80.

Handa, N., Ohashi, S., Kusano, K. and Kobayashi, I. (1997) Chi-star, a chi-related 11-mer sequence partially active in an *E. coli recC*1004 strain. *Genes Cells*, **2**, 525-36.

Handa, P., Acharya, N. and Varshney, U. (2001) Chimeras between single-stranded DNA-binding proteins from *Escherichia coli* and *Mycobacterium tuberculosis* reveal that their C-terminal domains interact with uracil DNA glycosylases. *J Biol Chem*, **276**, 16992-7.

Hargreaves, D., Rice, D.W., Sedelnikova, S.E., Artymiuk, P.J., Lloyd, R.G. and Rafferty, J.B. (1998) Crystal structure of *E.coli* RuvA with bound DNA Holliday junction at 6 A resolution. *Nat Struct Biol*, **5**, 441-6.

Hegde, S.P., Qin, M.H., Li, X.H., Atkinson, M.A., Clark, A.J., Rajagopalan, M. and Madiraju, M.V. (1996a) Interactions of RecF protein with RecO, RecR, and single-stranded DNA binding proteins reveal roles for the RecF-RecO-RecR complex in DNA repair and recombination. *Proc Natl Acad Sci U S A*, **93**, 14468-73.

Hegde, S.P., Rajagopalan, M. and Madiraju, M.V. (1996b) Preferential binding of *Escherichia coli* RecF protein to gapped DNA in the presence of adenosine (gamma-thio) triphosphate. *J Bacteriol*, **178**, 184-90.

Hendrix, R.W., Smith, M.C., Burns, R.N., Ford, M.E. and Hatfull, G.F. (1999) Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage. *Proc Natl Acad Sci U S A*, **96**, 2192-7.

Hendrix, R.W. (2002) Bacteriophages: evolution of the majority. *Theor Popul Biol*, **61**, 471-80.

Hendrix, R.W., Roberts, J.W., Stahl, F.W., and Weisberg, R.A., *Lambda II* (Cold Spring Harbour, NY; Cold Spring Harbour Lab. Press, 1983).

Hermann, R., Rubolph, R. and Jaenicke, R. (1979) Kinetics of *in vitro* reconstitution of oligomeric enzymes by cross-linking. *Nature*, **277**, 243-5.

Herskowitz, I. and Hagen, D. (1980) The lysis-lysogeny decision of phage lambda: explicit programming and responsiveness. *Annu Rev Genet*, **14**, 399-445.

Highton, P.J., Chang, Y. and Myers, R.J. (1990) Evidence for the exchange of segments between genomes during the evolution of lambdoid bacteriophages. *Mol Microbiol*, **4**, 1329-40.

Hill, S.A., Stahl, M.M. and Stahl, F.W. (1997) Single-strand DNA intermediates in phage lambda's Red recombination pathway. *Proc Natl Acad Sci U S A*, **94**, 2951-6.

Hollifield, W.C., Kaplan, E.N. and Huang, H.V. (1987) Efficient RecABC-dependent, homologous recombination between coliphage lambda and plasmids requires a phage *ninR* region gene. *Mol Gen Genet*, **210**, 248-55.

Honda, M., Inoue, J., Yoshimasu, M., Ito, Y., Shibata, T. and Mikawa, T. (2006) Identification of the RecR toprim domain as the binding site for both RecF and RecO: A role of RecR in RecFOR assembly at dsDNA-ssDNA junctions. *J Biol Chem*.

Horii, Z. and Clark, A.J. (1973) Genetic analysis of the *recF* pathway to genetic recombination in *Escherichia coli* K12: isolation and characterization of mutants. *J Mol Biol*, **80**, 327-44.

Iwasaki, H., Takahagi, M., Shiba, T., Nakata, A. and Shinagawa, H. (1991) *Escherichia coli* RuvC protein is an endonuclease that resolves the Holliday structure. *Embo J*, **10**, 4381-9.

Iyer, L.M., Koonin, E.V. and Aravind, L. (2002) Classification and evolutionary history of the single-strand annealing proteins, RecT, Redbeta, ERF and RAD52. *BMC Genomics*, **3**, 8.

Juhala, R.J., Ford, M.E., Duda, R.L., Youlton, A., Hatfull, G.F. and Hendrix, R.W. (2000) Genomic sequences of bacteriophages HK97 and HK022: pervasive genetic mosaicism in the lambdoid bacteriophages. *J Mol Biol*, **299**, 27-51.

Kaelin, W.G., Jr., Krek, W., Sellers, W.R., DeCaprio, J.A., Ajchenbaum, F., Fuchs, C.S., Chittenden, T., Li, Y., Farnham, P.J., Blanar, M.A. and et al. (1992) Expression cloning of a cDNA encoding a retinoblastoma-binding protein with E2F-like properties. *Cell*, **70**, 351-64.

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

Kapust, R.B. and Waugh, D.S. (1999) *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci*, **8**, 1668-74.

Karakousis, G., Ye, N., Li, Z., Chiu, S.K., Reddy, G. and Radding, C.M. (1998) The beta protein of phage lambda binds preferentially to an intermediate in DNA renaturation. *J Mol Biol*, **276**, 721-31.

Karu, A.E., Sakaki, Y., Echols, H. and Linn, S. (1975) The gamma protein specified by bacteriophage gamma. Structure and inhibitory activity for the *recBC* enzyme of *Escherichia coli*. *J Biol Chem*, **250**, 7377-87.

Kelley, L.A., MacCallum, R.M. and Sternberg, M.J. (2000) Enhanced genome annotation using structural profiles in the program 3D-PSSM. *J Mol Biol*, **299**, 499-520.

Kelman, Z., Yuzhakov, A., Andjelkovic, J. and O'Donnell, M. (1998) Devoted to the lagging strand-the subunit of DNA polymerase III holoenzyme contacts SSB to promote processive elongation and sliding clamp assembly. *Embo J*, **17**, 2436-49.

Keyer, K., Gort, A.S. and Imlay, J.A. (1995) Superoxide and the production of oxidative DNA damage. *J Bacteriol*, **177**, 6782-90.

Kmiec, E. and Holloman, W.K. (1981) Beta protein of bacteriophage lambda promotes renaturation of DNA. *J Biol Chem*, **256**, 12636-9.

Kodadek, T., Gan, D.C. and Stemke-Hale, K. (1989) The phage T4 uvs Y recombination protein stabilizes presynaptic filaments. *J Biol Chem*, **264**, 16451-7.

Kogoma, T. (1996) Recombination by replication. *Cell*, **85**, 625-7.

Kolodner, R., Hall, S.D. and Luisi-DeLuca, C. (1994) Homologous pairing proteins encoded by the *Escherichia coli recE* and *recT* genes. *Mol Microbiol*, **11**, 23-30.

Kovall, R. and Matthews, B.W. (1997) Toroidal structure of lambda-exonuclease. *Science*, **277**, 1824-7.

Kovall, R.A. and Matthews, B.W. (1998) Structural, functional, and evolutionary relationships between lambda-exonuclease and the type II restriction endonucleases. *Proc Natl Acad Sci U S A*, **95**, 7893-7.

Kowalczykowski, S.C., Clow, J., Somani, R. and Varghese, A. (1987) Effects of the *Escherichia coli* SSB protein on the binding of *Escherichia coli* RecA protein to single-stranded DNA. Demonstration of competitive binding and the lack of a specific protein-protein interaction. *J Mol Biol*, **193**, 81-95.

Kowalczykowski, S.C. and Krupp, R.A. (1987) Effects of *Escherichia coli* SSB protein on the single-stranded DNA-dependent ATPase activity of *Escherichia coli* RecA protein. Evidence that SSB protein facilitates the binding of RecA protein to regions of secondary structure within single-stranded DNA. *J Mol Biol*, **193**, 97-113.

Kowalczykowski, S.C., Dixon, D.A., Eggleston, A.K., Lauder, S.D. and Rehrauer, W.M. (1994) Biochemistry of homologous recombination in *Escherichia coli*. *Microbiol Rev*, **58**, 401-65.

Kowalczykowski, S.C. (2000) Initiation of genetic recombination and recombination-dependent replication. *Trends Biochem Sci*, **25**, 156-165.

Kreuzer, K.N. (2005) Interplay between DNA replication and recombination in prokaryotes. *Annu Rev Microbiol*, **59**, 43-67.

Kroger, M. and Hobom, G. (1982) A chain of interlinked genes in the *ninR* region of bacteriophage lambda. *Gene*, **20**, 25-38.

Kulkarni, S.K. and Stahl, F.W. (1989) Interaction between the *sbcC* gene of *Escherichia coli* and the *gam* gene of phage lambda. *Genetics*, **123**, 249-53.

Kushner, S.R., Nagaishi, H., Templin, A. and Clark, A.J. (1971) Genetic recombination in *Escherichia coli*: the role of exonuclease I. *Proc Natl Acad Sci U S A*, **68**, 824-7.

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

Lederberg, E.M. (1951) Lysogenicity in *E. coli* K12. *Genetics*, **36**, 560.

Lee, B.I., Kim, K.H., Park, S.J., Eom, S.H., Song, H.K. and Suh, S.W. (2004) Ring-shaped architecture of RecR: implications for its role in homologous recombinational DNA repair. *Embo J*, **23**, 2029-38.

Li, Z., Karakousis, G., Chiu, S.K., Reddy, G. and Radding, C.M. (1998) The beta protein of phage lambda promotes strand exchange. *J Mol Biol*, **276**, 733-44.

Lilley, D.M. and White, M.F. (2001) The junction-resolving enzymes. *Nat Rev Mol Cell Biol*, **2**, 433-43.

Little, J.W. (1967) An exonuclease induced by bacteriophage lambda. II. Nature of the enzymatic reaction. *J Biol Chem*, **242**, 679-86.

Little, J.W., Lehman, I.R. and Kaiser, A.D. (1967) An exonuclease induced by bacteriophage lambda. I. Preparation of the crystalline enzyme. *J Biol Chem*, **242**, 672-8.

Liu, J., Bond, J.P. and Morrical, S.W. (2006) Mechanism of presynaptic filament stabilization by the bacteriophage T4 UvsY recombination mediator protein. *Biochemistry*, **45**, 5493-502.

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

Lloyd, R.G. (1991) Conjugational recombination in resolvase-deficient *ruvC* mutants of *Escherichia coli* K-12 depends on recG. *J Bacteriol*, **173**, 5414-8.

Lloyd, R.G. and Sharples, G.J. (1993) Dissociation of synthetic Holliday junctions by *E. coli* RecG protein. *Embo J*, **12**, 17-22.

Lloyd, R.G., and Brooks Low, K., in *Escherichia coli and Salmonella Cellular and Molecular Biology* Neidhardt, F.C., Ed. (ASM Press, 1996) pp. 2236-2255.

Lohman, T.M. and Overman, L.B. (1985) Two binding modes in *Escherichia coli* single strand binding protein-single stranded DNA complexes. Modulation by NaCl concentration. *J Biol Chem*, **260**, 3594-603.

Loughlin, M.F., Barnard, F.M., Jenkins, D., Sharples, G.J. and Jenks, P.J. (2003) *Helicobacter pylori* mutants defective in RuvC Holliday junction resolvase display reduced macrophage survival and spontaneous clearance from the murine gastric mucosa. *Infect Immun*, **71**, 2022-31.

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

Luisi-DeLuca, C. (1995) Homologous pairing of single-stranded DNA and superhelical double-stranded DNA catalyzed by RecO protein from *Escherichia coli*. *J Bacteriol*, **177**, 566-72.

Lusetti, S.L. and Cox, M.M. (2002) The bacterial RecA protein and the recombinational DNA repair of stalled replication forks. *Annu Rev Biochem*, **71**, 71-100.

Madiraju, M.V. and Clark, A.J. (1991) Effect of RecF protein on reactions catalyzed by RecA protein. *Nucleic Acids Res*, **19**, 6295-300.

Madiraju, M.V. and Clark, A.J. (1992) Evidence for ATP binding and double-stranded DNA binding by *Escherichia coli* RecF protein. *J Bacteriol*, **174**, 7705-10.

Mahdi, A.A. and Lloyd, R.G. (1989) The *recR* locus of *Escherichia coli* K-12: molecular cloning, DNA sequencing and identification of the gene product. *Nucleic Acids Res*, **17**, 6781-94.

Mahdi, A.A., Sharples, G.J., Mandal, T.N. and Lloyd, R.G. (1996) Holliday junction resolvases encoded by homologous *rusA* genes in *Escherichia coli* K-12 and phage 82. *J Mol Biol*, **257**, 561-73.

Maisnier-Patin, S., Nordstrom, K. and Dasgupta, S. (2001) Replication arrests during a single round of replication of the *Escherichia coli* chromosome in the absence of DnaC activity. *Mol Microbiol*, **42**, 1371-82.

Mandal, T.N., Mahdi, A.A., Sharples, G.J. and Lloyd, R.G. (1993) Resolution of Holliday intermediates in recombination and DNA repair: indirect suppression of *ruvA*, *ruvB*, and *ruvC* mutations. *J Bacteriol*, **175**, 4325-34.

Marians, K.J. (1999) PriA: at the crossroads of DNA replication and recombination. *Prog Nucleic Acid Res Mol Biol*, **63**, 39-67.

Matsuura, S., Komatsu, J., Hirano, K., Yasuda, H., Takashima, K., Katsura, S. and Mizuno, A. (2001) Real-time observation of a single DNA digestion by

lambda exonuclease under a fluorescence microscope field. *Nucleic Acids Res*, **29**, E79.

Maxwell, K.L., Reed, P., Zhang, R.G., Beasley, S., Walmsley, A.R., Curtis, F.A., Joachimiak, A., Edwards, A.M. and Sharples, G.J. (2005) Functional similarities between phage lambda Orf and *Escherichia coli* RecFOR in initiation of genetic exchange. *Proc Natl Acad Sci U S A*, **102**, 11260-5.

McGlynn, P. and Lloyd, R.G. (2002) Recombinational repair and restart of damaged replication forks. *Nat Rev Mol Cell Biol*, **3**, 859-70.

McGregor, N., Ayora, S., Sedelnikova, S., Carrasco, B., Alonso, J.C., Thaw, P. and Rafferty, J. (2005) The structure of *Bacillus subtilis* RecU Holliday junction resolvase and its role in substrate selection and sequence-specific cleavage. *Structure*, **13**, 1341-51.

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

Mehta, P., Katta, K. and Krishnaswamy, S. (2004) HNH family subclassification leads to identification of commonality in the His-Me endonuclease superfamily. *Protein Sci*, **13**, 295-300.

Menetski, J.P. and Kowalczykowski, S.C. (1985) Interaction of RecA protein with single-stranded DNA. Quantitative aspects of binding affinity modulation by nucleotide cofactors. *J Mol Biol*, **181**, 281-95.

Merino, D., Reglier-Poupet, H., Berche, P. and Charbit, A. (2002) A hypermutator phenotype attenuates the virulence of *Listeria monocytogenes* in a mouse model. *Mol Microbiol*, **44**, 877-87.

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

Meyer, R.R. and Laine, P.S. (1990) The single-stranded DNA-binding protein of *Escherichia coli*. *Microbiol Rev*, **54**, 342-80.

Miao, E.A. and Miller, S.I. (1999) Bacteriophages in the evolution of pathogen-host interactions. *Proc Natl Acad Sci U S A*, **96**, 9452-4.

Michel, B., Grompone, G., Flores, M.J. and Bidnenko, V. (2004) Multiple pathways process stalled replication forks. *Proc Natl Acad Sci U S A*, **101**, 12783-8.

Miyamoto, H., Nakai, W., Yajima, N., Fujibayashi, A., Higuchi, T., Sato, K. and Matsushiro, A. (1999) Sequence analysis of Stx2-converting phage VT2-Sa shows a great divergence in early regulation and replication regions. *DNA Res*, **6**, 235-40.

Mmolawa, P.T., Schmieger, H., Tucker, C.P. and Heuzenroeder, M.W. (2003) Genomic structure of the *Salmonella enterica serovar Typhimurium* DT 64 bacteriophage ST64T: evidence for modular genetic architecture. *J Bacteriol*, **185**, 3473-5.

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

Morimatsu, K. and Kowalczykowski, S.C. (2003) RecFOR proteins load RecA protein onto gapped DNA to accelerate DNA strand exchange: a universal step of recombinational repair. *Mol Cell*, **11**, 1337-47.

Morrison, P.T., Lovett, S.T., Gilson, L.E. and Kolodner, R. (1989) Molecular analysis of the *Escherichia coli recO* gene. *J Bacteriol*, **171**, 3641-9.

Muniyappa, K., Shaner, S.L., Tsang, S.S. and Radding, C.M. (1984) Mechanism of the concerted action of RecA protein and helix-destabilizing proteins in homologous recombination. *Proc Natl Acad Sci U S A*, **81**, 2757-61.

Muniyappa, K. and Radding, C.M. (1986) The homologous recombination system of phage lambda. Pairing activities of beta protein. *J Biol Chem*, **261**, 7472-8.

Murphy, K.C. (1991) Lambda Gam protein inhibits the helicase and chi-stimulated recombination activities of *Escherichia coli* RecBCD enzyme. *J Bacteriol*, **173**, 5808-21.

Murphy, K.C. (1998) Use of bacteriophage lambda recombination functions to promote gene replacement in *Escherichia coli*. *J Bacteriol*, **180**, 2063-71.

Muyrers, J.P., Zhang, Y., Buchholz, F. and Stewart, A.F. (2000) RecE/RecT and Redalpha/Redbeta initiate double-stranded break repair by specifically interacting with their respective partners. *Genes Dev*, **14**, 1971-82.

Myers, R.S. and Stahl, F.W. (1994) Chi and the RecBC D enzyme of *Escherichia coli*. *Annu Rev Genet*, **28**, 49-70.

Nedenskov-Sorensen, P., Bukholm, G. and Bovre, K. (1990) Natural competence for genetic transformation in *Campylobacter pylori*. *J Infect Dis*, **161**, 365-6.

Neely, M.N. and Friedman, D.I. (1998) Functional and genetic analysis of regulatory regions of coliphage H-19B: location of shiga-like toxin and lysis genes suggest a role for phage functions in toxin release. *Mol Microbiol*, **28**, 1255-67.

Oppenheim, A.B., Kobiler, O., Stavans, J., Court, D.L. and Adhya, S. (2005) Switches in bacteriophage lambda development. *Annu Rev Genet*, **39**, 409-29.

Parkhill, J., Sebaihia, M., Preston, A., Murphy, L.D., Thomson, N., Harris, D.E., Holden, M.T., Churcher, C.M., Bentley, S.D., Mungall, K.L., Cerdeno-Tarraga, A.M., Temple, L., James, K., Harris, B., Quail, M.A., Achtman, M., Atkin, R., Baker, S., Basham, D., Bason, N., Cherevach, I., Chillingworth, T., Collins, M., Cronin, A., Davis, P., Doggett, J., Feltwell, T., Goble, A., Hamlin, N., Hauser, H., Holroyd, S., Jagels, K., Leather, S., Moule, S., Norberczak, H., O'Neil, S., Ormond, D., Price, C., Rabbinowitsch, E., Rutter, S., Sanders, M., Saunders, D., Seeger, K., Sharp, S., Simmonds, M., Skelton, J., Squares, R., Squares, S., Stevens, K., Unwin, L., Whitehead, S., Barrell, B.G. and Maskell, D.J. (2003) Comparative analysis of the genome sequences of *Bordetella pertussis*, *Bordetella parapertussis* and *Bordetella bronchiseptica*. *Nat Genet*, **35**, 32-40.

Parsons, C.A., Stasiak, A., Bennett, R.J. and West, S.C. (1995) Structure of a multisubunit complex that promotes DNA branch migration. *Nature*, **374**, 375-8.

Passy, S.I., Yu, X., Li, Z., Radding, C.M. and Egelman, E.H. (1999) Rings and filaments of beta protein from bacteriophage lambda suggest a superfamily of recombination proteins. *Proc Natl Acad Sci U S A*, **96**, 4279-84.

Perna, N.T., Plunkett, G., 3rd, Burland, V., Mau, B., Glasner, J.D., Rose, D.J., Mayhew, G.F., Evans, P.S., Gregor, J., Kirkpatrick, H.A., Posfai, G., Hackett, J., Klink, S., Boutin, A., Shao, Y., Miller, L., Grotbeck, E.J., Davis, N.W., Lim, A., Dimalanta, E.T., Potamousis, K.D., Apodaca, J., Anantharaman,

T.S., Lin, J., Yen, G., Schwartz, D.C., Welch, R.A. and Blattner, F.R. (2001) Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature*, **409**, 529-33.

Petit, M.A.E., D. (2002) Essential bacterial helicases that counteract the toxicity of recombination proteins. *Embo J*, **21**, 3137-3147.

Plunkett, G., 3rd, Rose, D.J., Durfee, T.J. and Blattner, F.R. (1999) Sequence of Shiga toxin 2 phage 933W from *Escherichia coli* O157:H7: Shiga toxin as a phage late-gene product. *J Bacteriol*, **181**, 1767-78.

Pollock, R. and Treisman, R. (1990) A sensitive method for the determination of protein-DNA binding specificities. *Nucleic Acids Res*, **18**, 6197-204.

Poteete, A.R. (2001) What makes the bacteriophage lambda Red system useful for genetic engineering: molecular mechanism and biological function. *FEMS Microbiol Lett*, **201**, 9-14.

Poteete, A.R., Fenton, A.C. and Wang, H.R. (2002) Recombination-promoting activity of the bacteriophage lambda Rap protein in *Escherichia coli* K-12. *J Bacteriol*, **184**, 4626-9.

Poteete, A.R. (2004) Modulation of DNA repair and recombination by the bacteriophage lambda Orf function in *Escherichia coli* K-12. *J Bacteriol*, **186**, 2699-707.

Prescott, D.M. and Kuempel, P.L. (1972) Bidirectional replication of the chromosome in *Escherichia coli*. *Proc Natl Acad Sci U S A*, **69**, 2842-5.

Ptashne, M., *Genetic Switch: Phage Lambda Revisited* (Cold Spring Harbour, NY: Cold Spring Harbour Lab. Press, New York, 2004).

Radding, C.M., Rosensweig, J., Richards, F., and Cassuto, E. (1971) Separation and characterisation of exonuclease, β protein and a complex of both. *J Biol Chem*, **246**, 2510-2512.

Rafferty, J.B., Sedelnikova, S.E., Hargreaves, D., Artymiuk, P.J., Baker, P.J., Sharples, G.J., Mahdi, A.A., Lloyd, R.G. and Rice, D.W. (1996) Crystal structure of DNA recombination protein RuvA and a model for its binding to the Holliday junction. *Science*, **274**, 415-21.

Resch, G., Kulik, E.M., Dietrich, F.S. and Meyer, J. (2004) Complete genomic nucleotide sequence of the temperate bacteriophage Aa Phi 23 of *Actinobacillus actinomycetemcomitans*. *J Bacteriol*, **186**, 5523-8.

Roca, A.I. and Cox, M.M. (1997) RecA protein: structure, function, and role in recombinational DNA repair. *Prog Nucleic Acid Res Mol Biol*, **56**, 129-223.

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

Rothstein, R., Michel, B. and Gangloff, S. (2000) Replication fork pausing and recombination or "gimme a break". *Genes Dev*, **14**, 1-10.

Rybalchenko, N., Golub, E.I., Bi, B. and Radding, C.M. (2004) Strand invasion promoted by recombination protein beta of coliphage lambda. *Proc Natl Acad Sci U S A*, **101**, 17056-60.

Saito, A., Iwasaki, H., Ariyoshi, M., Morikawa, K. and Shinagawa, H. (1995) Identification of four acidic amino acids that constitute the catalytic center of the RuvC Holliday junction resolvase. *Proc Natl Acad Sci U S A*, **92**, 7470-4.

Sambrook, J., and Russell, D.W., *Molecular Cloning: A laboratory manual* (Cold Spring Harbour Press, Cold Spring Harbour, NY, 2nd Edition, 2001).

Sanchez, H., Kidane, D., Reed, P., Curtis, F.A., Cozar, M.C., Graumann, P.L., Sharples, G.J. and Alonso, J.C. (2005) The RuvAB branch migration translocase and RecU Holliday junction resolvase are required for double-stranded DNA break repair in *Bacillus subtilis*. *Genetics*, **171**, 873-83.

Sandhya, S., Kishore, S., Sowdhamini, R. and Srinivasan, N. (2003) Effective detection of remote homologues by searching in sequence dataset of a protein domain fold. *FEBS Lett*, **552**, 225-30.

Sandigursky, M. and Franklin, W.A. (1994) *Escherichia coli* single-stranded DNA binding protein stimulates the DNA deoxyribophosphodiesterase activity of exonuclease I. *Nucleic Acids Res*, **22**, 247-50.

Sargentini, N.J. and Smith, K.C. (1989) Role of r*uvAB* genes in UV- and gamma-radiation and chemical mutagenesis in *Escherichia coli*. *Mutat Res*, **215**, 115-29.

Savvides, S.N., Raghunathan, S., Futterer, K., Kozlov, A.G., Lohman, T.M. and Waksman, G. (2004) The C-terminal domain of full-length *E. coli* SSB is disordered even when bound to DNA. *Protein Sci*, **13**, 1942-7.

Sawitzke, J.A. and Stahl, F.W. (1992) Phage lambda has an analog of *Escherichia coli recO*, *recR* and *recF* genes. *Genetics*, **130**, 7-16.

Sawitzke, J.A. and Stahl, F.W. (1994) The phage lambda orf gene encodes a trans-acting factor that suppresses *Escherichia coli recO*, *recR*, and *recF* mutations for recombination of lambda but not of *E. coli*. *J Bacteriol*, **176**, 6730-7.

Sawitzke, J.A. and Stahl, F.W. (1997) Roles for lambda Orf and *Escherichia coli* RecO, RecR and RecF in lambda recombination. *Genetics*, **147**, 357-69.

Schouler, C., Ehrlich, S.D. and Chopin, M.C. (1994) Sequence and organization of the lactococcal prolate-headed bIL67 phage genome. *Microbiology*, **140 (Pt 11)**, 3061-9.

Seigneur, M., Bidnenko, V., Ehrlich, S.D. and Michel, B. (1998) RuvAB acts at arrested replication forks. *Cell*, **95**, 419-30.

Shah, R., Bennett, R.J. and West, S.C. (1994) Genetic recombination in *E. coli*: RuvC protein cleaves Holliday junctions at resolution hotspots *in vitro*. *Cell*, **79**, 853-64.

Shan, Q., Bork, J.M., Webb, B.L., Inman, R.B. and Cox, M.M. (1997) RecA protein filaments: end-dependent dissociation from ssDNA and stabilization by RecO and RecR proteins. *J Mol Biol*, **265**, 519-40.

Sharff, A.J., Rodseth, L.E., Spurlino, J.C. and Quiocho, F.A. (1992) Crystallographic evidence of a large ligand-induced hinge-twist motion between the two domains of the maltodextrin binding protein involved in active transport and chemotaxis. *Biochemistry*, **31**, 10657-63.

Sharples, G.J., Benson, F.E., Illing, G.T. and Lloyd, R.G. (1990) Molecular and functional analysis of the ruv region of *Escherichia coli* K-12 reveals three genes involved in DNA repair and recombination. *Mol Gen Genet*, **221**, 219-26.

Sharples, G.J. and Lloyd, R.G. (1991) Resolution of Holliday junctions in *Escherichia coli*: identification of the *ruvC* gene product as a 19-kilodalton protein. *J Bacteriol*, **173**, 7711-5.

Sharples, G.J., Corbett, L.M. and Graham, I.R. (1998) lambda Rap protein is a structure-specific endonuclease involved in phage recombination. *Proc Natl Acad Sci U S A*, **95**, 13507-12.

Sharples, G.J., Corbett, L.M. and McGlynn, P. (1999a) DNA structure specificity of Rap endonuclease. *Nucleic Acids Res*, **27**, 4121-7.

Sharples, G.J., Ingleston, S.M. and Lloyd, R.G. (1999b) Holliday junction processing in bacteria: insights from the evolutionary conservation of RuvABC, RecG, and RusA. *J Bacteriol*, **181**, 5543-50.

Sharples, G.J. (2001) The X philes: structure-specific endonucleases that resolve Holliday junctions. *Mol Microbiol*, **39**, 823-34.

Sharples, G.J., Bolt, E.L. and Lloyd, R.G. (2002) RusA proteins from the extreme thermophile *Aquifex aeolicus* and lactococcal phage r1t resolve Holliday junctions. *Mol Microbiol*, **44**, 549-59.

Sharples, G.J., Curtis, F.A., McGlynn, P. and Bolt, E.L. (2004) Holliday junction binding and resolution by the Rap structure-specific endonuclease of phage lambda. *J Mol Biol*, **340**, 739-51.

Sherratt, D.J. (2003) Bacterial chromosome dynamics. *Science*, **301**, 780-5.

Shurvinton, C.E., Lloyd, R.G., Benson, F.E. and Attfield, P.V. (1984) Genetic analysis and molecular cloning of the *Escherichia coli ruv* gene. *Mol Gen Genet*, **194**, 322-9.

Silva, C.J.S.M., Sousa, F., Gubitz, G., and Cavaco-Paulo, A. (2004) Chemical Modifications on Proteins Using Gluteraldehyde. *Food Technology and Biotechnology*, **42**, 51-56.

Simmons, L.A., Felczak, M. and Kaguni, J.M. (2003) DnaA Protein of Escherichia coli: oligomerization at the *E. coli* chromosomal origin is required for initiation and involves specific N-terminal amino acids. *Mol Microbiol*, **49**, 849-58.

Singleton, M.R., Wentzell, L.M., Liu, Y., West, S.C. and Wigley, D.B. (2002) Structure of the single-strand annealing domain of human RAD52 protein. *Proc Natl Acad Sci U S A*, **99**, 13492-7.

Singleton, M.R., Dillingham, M.S., Gaudier, M., Kowalczykowski, S.C. and Wigley, D.B. (2004) Crystal structure of RecBCD enzyme reveals a machine for processing DNA breaks. *Nature*, **432**, 187-93.

Smith, T.F. and Waterman, M.S. (1981) Identification of common molecular subsequences. *J Mol Biol*, **147**, 195-7.

Snyder, L. (1995) Phage-exclusion enzymes: a bonanza of biochemical and cell biology reagents? *Mol Microbiol*, **15**, 415-20.

Spurlino, J.C., Lu, G.Y. and Quiocho, F.A. (1991) The 2.3-A resolution structure of the maltose- or maltodextrin-binding protein, a primary receptor of bacterial active transport and chemotaxis. *J Biol Chem*, **266**, 5202-19.

Stahl, F. (1996) Meiotic recombination in yeast: coronation of the double-strand-break repair model. *Cell*, **87**, 965-8.

Stahl, F.W., Kobayashi, I. and Stahl, M.M. (1985) In phage lambda, *cos* is a recombinator in the red pathway. *J Mol Biol*, **181**, 199-209.

Stahl, M.M., Thomason, L., Poteete, A.R., Tarkowski, T., Kuzminov, A. and Stahl, F.W. (1997) Annealing vs. invasion in phage lambda recombination. *Genetics*, **147**, 961-77.

Stasiak, A., Tsaneva, I.R., West, S.C., Benson, C.J., Yu, X. and Egelman, E.H. (1994) The *Escherichia coli* RuvB branch migration protein forms double hexameric rings around DNA. *Proc Natl Acad Sci U S A*, **91**, 7618-22.

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

Subramanian, K., Rutvisuttinunt, W., Scott, W. and Myers, R.S. (2003) The enzymatic basis of processivity in lambda exonuclease. *Nucleic Acids Res*, **31**, 1585-96.

Suerbaum, S., Smith, J.M., Bapumia, K., Morelli, G., Smith, N.H., Kunstmann, E., Dyrek, I. and Achtman, M. (1998) Free recombination within *Helicobacter pylori*. *Proc Natl Acad Sci U S A*, **95**, 12619-24.

Suerbaum, S. and Michetti, P. (2002) *Helicobacter pylori* infection. *N Engl J Med*, **347**, 1175-86.

Sweezy, M.A. and Morrical, S.W. (1997) Single-stranded DNA binding properties of the uvsY recombination protein of bacteriophage T4. *J Mol Biol*, **266**, 927-38.

Sweezy, M.A. and Morrical, S.W. (1999) Biochemical interactions within a ternary complex of the bacteriophage T4 recombination proteins uvsY and gp32 bound to single-stranded DNA. *Biochemistry*, **38**, 936-44.

Szostak, J.W., Orr-Weaver, T.L., Rothstein, R.J. and Stahl, F.W. (1983) The double-strand-break repair model for recombination. *Cell*, **33**, 25-35.

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

Takahashi, N. and Kobayashi, I. (1990) Evidence for the double-strand break repair model of bacteriophage lambda recombination. *Proc Natl Acad Sci U S A*, **87**, 2790-4.

Takano, T. (1966) Behaviour of some episomal elements in recombination-deficient mutant of *Escherichia coli*. *Jpn. J. Microbiol*, **10**, 201-210.

Tanaka, K., Nishimori, K., Makino, S., Nishimori, T., Kanno, T., Ishihara, R., Sameshima, T., Akiba, M., Nakazawa, M., Yokomizo, Y. and Uchida, I. (2004) Molecular characterization of a prophage of *Salmonella enterica serotype Typhimurium* DT104. *J Clin Microbiol*, **42**, 1807-12.

Tarkowski, T.A., Mooney, D., Thomason, L.C. and Stahl, F.W. (2002) Gene products encoded in the *ninR* region of phage lambda participate in Red-mediated recombination. *Genes Cells*, **7**, 351-63.

Taylor, A.F. and Smith, G.R. (2003) RecBCD enzyme is a DNA helicase with fast and slow motors of opposite polarity. *Nature*, **423**, 889-93.

Taylor, K. and Wegrzyn, G. (1995) Replication of coliphage lambda DNA. *FEMS Microbiol Rev*, **17**, 109-19.

Taylor, K., and Wegrzyn, G., in *Molecular Microbiology* Busby, S.J.W., Thomas, C.M., and Brown, N.L., Ed. (Berlin and Heidelberg: Springer, 1998) pp. 81-97.

Thresher, R.J., Christiansen, G. and Griffith, J.D. (1988) Assembly of presynaptic filaments. Factors affecting the assembly of RecA protein onto single-stranded DNA. *J Mol Biol*, **201**, 101-13.

Tsang, S.S., Muniyappa, K., Azhderian, E., Gonda, D.K., Radding, C.M., Flory, J. and Chase, J.W. (1985) Intermediates in homologous pairing promoted by recA protein. Isolation and characterization of active presynaptic complexes. *J Mol Biol*, **185**, 295-309.

Tsuchiya, Y., Kinoshita, K. and Nakamura, H. (2004) Structure-based prediction of DNA-binding sites on proteins using the empirical preference of electrostatic potential and the shape of molecular surfaces. *Proteins*, **55**, 885-94.

Umezu, K., Chi, N.W. and Kolodner, R.D. (1993) Biochemical interaction of the *Escherichia coli* RecF, RecO, and RecR proteins with RecA protein and single-stranded DNA binding protein. *Proc Natl Acad Sci U S A*, **90**, 3875-9.

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

van de Putte, P., Zwenk, H. and Rorsch, A. (1966) Properties of four mutants of *Escherichia coli* defective in genetic recombination. *Mutat Res*, **3**, 381-92.

van Gool, A.J., Shah, R., Mezard, C. and West, S.C. (1998) Functional interactions between the holliday junction resolvase and the branch migration motor of *Escherichia coli*. *Embo J*, **17**, 1838-45.

van Gool, A.J., Hajibagheri, N.M., Stasiak, A. and West, S.C. (1999) Assembly of the *Escherichia coli* RuvABC resolvasome directs the orientation of holliday junction resolution. *Genes Dev*, **13**, 1861-70.

Vander Byl, C. and Kropinski, A.M. (2000) Sequence of the genome of Salmonella bacteriophage P22. *J Bacteriol*, **182**, 6472-81.

Wagner, P.L. and Waldor, M.K. (2002) Bacteriophage control of bacterial virulence. *Infect Immun*, **70**, 3985-93.

Waldor, M.K. and Mekalanos, J.J. (1996) Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science*, **272**, 1910-4.

Waldor, M.K. (1998) Bacteriophage biology and bacterial virulence. *Trends Microbiol*, **6**, 295-7.

Wang, T.V., Chang, H., & Hung, J. (1993) Cosuppression of *recF*, *recR*, and *recO* mutations by mutant *recA* alleles in *Escherichia coli* cells. *Mutat Res*, **294**, 157-166.

Watson, J.D. and Crick, F.H. (1953) Genetical implications of the structure of deoxyribonucleic acid. *Nature*, **171**, 964-7.

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

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

Webb, B.L., Cox, M.M. and Inman, R.B. (1999) ATP hydrolysis and DNA binding by the *Escherichia coli* RecF protein. *J Biol Chem*, **274**, 15367-74.

Weiner, M.P., Anderson, C., Jerpseth, B., Wells, S., Johnson-Browne, B. *et al.* (1994) Strategies Newsletter (Stratagene). *Strategies*, **7**, 41-43.

Weisberg, R., and Landy, A., in *Lambda II* R.W. Hendrix, J.W.R., F.W Stahl, and R.A. Weisberg, Ed. (Cold Spring Harbour, New York, 1983) pp. 211-251.

Whitby, M.C., Vincent, S.D. and Lloyd, R.G. (1994) Branch migration of Holliday junctions: identification of RecG protein as a junction specific DNA helicase. *Embo J*, **13**, 5220-8.

Wu, Z., Xing, X., Bohl, C.E., Wisler, J.W., Dalton, J.T. and Bell, C.E. (2006) Domain structure and DNA binding regions of beta protein from bacteriophage lambda. *J Biol Chem*, **281**, 25205-14.

Yamada, K., Ariyoshi, M. and Morikawa, K. (2004) Three-dimensional structural views of branch migration and resolution in DNA homologous recombination. *Curr Opin Struct Biol*, **14**, 130-7.

Yamaguchi, T., Hayashi, T., Takami, H., Nakasone, K., Ohnishi, M., Nakayama, K., Yamada, S., Komatsuzawa, H. and Sugai, M. (2000) Phage conversion of exfoliative toxin A production in *Staphylococcus aureus*. *Mol Microbiol*, **38**, 694-705.

Yassa, D.S., Chou, K.M. and Morrical, S.W. (1997) Characterization of an amino-terminal fragment of the bacteriophage T4 uvsY recombination protein. *Biochimie*, **79**, 275-85.

Yeung, T., Mullin, D.A., Chen, K.S., Craig, E.A., Bardwell, J.C. and Walker, J.R. (1990) Sequence and expression of the *Escherichia coli recR* locus. *J Bacteriol*, **172**, 6042-7.

Yu, D., Ellis, H.M., Lee, E.C., Jenkins, N.A., Copeland, N.G. and Court, D.L. (2000) An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc Natl Acad Sci U S A*, **97**, 5978-83.

Yu, H.B., Zhang, Y.L., Lau, Y.L., Yao, F., Vilches, S., Merino, S., Tomas, J.M., Howard, S.P. and Leung, K.Y. (2005) Identification and characterization of putative virulence genes and gene clusters in *Aeromonas hydrophila* PPD134/91. *Appl Environ Microbiol*, **71**, 4469-77.

Yu, M., Souaya, J. and Julin, D.A. (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 U S A*, **95**, 981-6.

Zerbib, D., Mezard, C., George, H. and West, S.C. (1998) Coordinated actions of RuvABC in Holliday junction processing. *J Mol Biol*, **281**, 621-30.

Appendix A

Publications

Molecular Microbiology (2005) 55(5), 1332–1345 doi:10.1111/j.1365-2958.2004.04476.x

Evolution of a phage RuvC endonuclease for resolution of both Holliday and branched DNA junctions

Fiona A. Curtis, Patricia Reed and Gary J. Sharples*
Centre for Infectious Diseases, Wolfson Research Institute, University of Durham, Queen's Campus, Stockton-on-Tees TS17 6BH, UK.

Summary

Resolution of Holliday junction recombination intermediates in most Gram-negative bacteria is accomplished by the RuvC endonuclease acting in concert with the RuvAB branch migration machinery. Gram-positive species, however, lack RuvC, with the exception of distantly related orthologues from bacteriophages infecting Lactococci and Streptococci. We have purified one of these proteins, 67RuvC, from *Lactococcus lactis* phage bIL67 and demonstrated that it functions as a Holliday structure resolvase. Differences in the sequence selectivity of resolution between 67RuvC and *Escherichia coli* RuvC were noted, although both enzymes prefer to cleave 3′ of thymidine residues. However, unlike its cellular counterpart, 67RuvC readily binds and cleaves a variety of branched DNA substrates in addition to Holliday junctions. Plasmids expressing 67RuvC induce chromosomal breaks, probably as a consequence of replication fork cleavage, and cannot be recovered from recombination-defective *E. coli* strains. Despite these deleterious effects, 67RuvC constructs suppress the UV light sensitivity of *ruvA*, *ruvAB* and *ruvABC* mutant strains confirming that the phage protein mediates Holliday junction resolution *in vivo*. The characterization of 67RuvC offers a unique insight into how a Holliday junction-specific resolvase can evolve into a debranching endonuclease tailored to the requirements of phage recombination.

Introduction

Genetic recombination salvages stalled replication forks, ensures accurate chromosomal repair and generates the exchanges and rearrangements that fuel evolution. The central intermediate in these reactions is the four-way (Holliday) junction, derived from recombinational exchange or regression of replication fork structures (Seigneur *et al.*, 1998). In typical recombination reactions, the Holliday junction is fashioned when RecA polymerizes on single-stranded DNA (ssDNA) and mediates the exchange of strands between homologous partners (West, 1992). Once formed, the junction is free to branch migrate along the linked duplexes, generating tracts of heteroduplex DNA, and is resolved by paired cleavages at the cross-over. The nicked duplex products of resolution are repaired by DNA ligase to reconstitute intact chromosomes. Holliday junction-resolving enzymes are widespread in nature, underlining their importance in genome maintenance and transmission (Lilley and White, 2001; Sharples, 2001).

In *Escherichia coli*, and most other Gram-negative bacteria, the RuvC-resolving enzyme operates in conjunction with the RuvAB complex, thus facilitating the coupling of branch migration and resolution (reviewed in Sharples *et al.*, 1999). Each of the four duplex arms are engaged within channels on the surface of a RuvA tetramer ensuring that the junction adopts an approximately square planar configuration optimal for branch migration (Hargreaves *et al.*, 1998). Two hexamers of the RuvB helicase motor join RuvA on a pair of opposing duplexes with the DNA fed through a central cavity in each RuvB ring (Stasiak *et al.*, 1994). The accessible face of the junction accommodates the dimeric endonuclease component, RuvC, which directs the paired incisions crucial for Holliday junction resolution (Ariyoshi *et al.*, 1994). Protein assembly at the cross-over is stabilized by RuvB interacting with both RuvA and RuvC, while RuvA and RuvC target the Holliday structure directly (reviewed in Sharples *et al.*, 1999). The tripartite assembly accomplishes branch migration by each RuvB helicase motor drawing the DNA through the static complex in a reaction fuelled by ATP hydrolysis. RuvC exhibits a distinct sequence specificity, preferentially cleaving between the third and fourth positions of a tetranucleotide consensus, 5′-$^{A}/_{T}$TT$^{G}/_{C}$-3′ (Shah *et al.*, 1994a). Hence, DNA drawn through the RuvABC complex can be 'scanned' until appropriate target sequences are encountered, allowing RuvC to bring its pair of active sites into position for dual strand scission.

The model for co-ordinated junction processing by RuvABC is widely applicable as judged by the presence of cognate representatives in many bacterial species

Accepted 10 November, 2004. *For correspondence. E-mail gary.sharples@durham.ac.uk; Tel. (+44) 191 334 0466; Fax (+44) 191 334 0468.

(Sharples *et al.*, 1999). However, the Gram-positive group lack an identifiable orthologue of RuvC despite the presence of homologues of the RuvAB branch migration machinery (Sharples *et al.*, 1999). They utilize an unrelated junction resolution enzyme, RecU, that also seems to function as part of a complex with RuvAB (Ayora *et al.*, 2004; G.J. Sharples and J. Alonso, unpublished results). RuvC-like proteins have, however, been found in phages and prophages from Lactococci and Streptococci (Bidnenko *et al.*, 1998). Phage RuvC proteins share only 19–21% identity with *E. coli* RuvC (*Ec*RuvC; see Fig. 9A; data not shown), although they do possess equivalents of the four catalytic carboxylates necessary for RuvC catalysis (Saito *et al.*, 1995; Bidnenko *et al.*, 1998). Genetic analysis of *Lactococcus lactis* phage bIL66 RuvC showed that its expression is detrimental to cell growth and lethal in *E. coli* mutants lacking the *recA* gene. These effects can be explained by incision at replication forks to generate double-strand breaks; indeed, random chromosomal breaks were detected in cells expressing bIL66 RuvC, while nicks generated in theta-replicating plasmids were located close to the origin of replication (Bidnenko *et al.*, 1998). This toxicity mirrors that observed with the bacteriophage T4 endonuclease VII and T7 endonuclease I Holliday junction resolvases when overexpressed in *E. coli* (de Massy *et al.*, 1987; Kosak and Kemper, 1990).

In this study, we purified the RuvC-like protein from lactococcal phage bIL67 (67RuvC) which shares 43% identity with its equivalent from bIL66 (Bidnenko *et al.*, 1998; see Fig. 9A). Experiments *in vitro* confirmed that 67RuvC resolves synthetic Holliday junctions to generate nicked duplex products. However, compared with *Ec*RuvC, it shows a relaxed DNA structure and sequence selectivity. Like *Ec*RuvC, 67RuvC resolves Holliday structures on the 3′ side of thymidine residues; however, resolution occurs preferentially at the consensus 5′-T↓$^{A}/_{G}$, rather than the 5′-TT↓ core favoured by *Ec*RuvC. Unlike *Ec*RuvC, it cleaves DNA substrates resembling replication forks almost as well as it resolves Holliday junctions; fork resolution is largely independent of the nucleotide sequences at the branch point. Both fork and Holliday cleavage activities were also evaluated *in vivo* using *E. coli recA* and *ruv* mutant strains. The analysis of 67RuvC provides a powerful tool to explore the architectural differences in RuvC proteins responsible for DNA branch recognition and resolution.

Results

Overexpression and purification of phage bIL67 RuvC

To characterize the RuvC-like proteins from *L. lactis*, we attempted to clone the relevant genes from phages bIL66 (66*ruvC*) and bIL67 (67*ruvC*) in the expression vector pT7-7. Only one construct, carrying the 67*ruvC* gene (pGS905), was obtained and this was deleterious to the growth of *E. coli* strains. As shown previously with bIL66 RuvC constructs (Bidnenko *et al.*, 1998), *recA* mutant strains could not be transformed with pGS905, suggesting that both phage proteins share a similar capacity to induce double-strand breaks. The construct could not be introduced into standard BL21 strains either, although it was recovered from BL21-SI and BL21-AI derivatives which offer much tighter regulation of target proteins (data not shown). Plasmid constructs with improved stability were obtained by transferring the pGS905 insert into pET24a and pET14b to generate pFC105 and pFC128 respectively. We used pFC105 to overproduce 67RuvC and purified the protein to near homogeneity following protocols similar to those employed in the purification of *Ec*RuvC (Dunderdale *et al.*, 1994). Recovery of large quantities of wild-type 67RuvC was hampered by low levels of protein overexpression (see Fig. 8C, lane b).

DNA structure selectivity of binding and cleavage

The *Ec*RuvC endonuclease is highly selective for Holliday junctions, although it can bind weakly to other branched DNA substrates and cleave three-way junctions with considerably reduced efficiency (Benson and West, 1994; Takahagi *et al.*, 1994; Fogg *et al.*, 1999). Experiments on bIL66 RuvC, however, indicate that phage RuvC proteins display a dramatically altered DNA junction specificity (Bidnenko *et al.*, 1998). Because of its structural similarity with bIL66 RuvC (see Fig. 9A), we anticipated that 67RuvC would behave in a similar manner and so assessed its ability to bind and cleave a range of branched DNA molecules. A 50 bp Holliday junction (J11), containing the *Ec*RuvC consensus sequence for resolution within an 11 bp core of homology, was employed alongside branched substrates missing one (flap) or two (fork) strands and a linear duplex (Figs 1 and 2). All substrates were labelled with ^{32}P in one of the constituent oligonucleotides. In gel retardation assays, 67RuvC bound to all four of the substrates tested, including the duplex, although it did show a preference for those substrates containing a branched element (Fig. 1A). Multiple complexes were formed with the branched substrates: at least three with X and flap, two with fork and only a single with duplex DNA. A similar pattern of complex formation was noted with the RusA Holliday junction resolvase from *E. coli* prophage DLP12 (Chan *et al.*, 1998). In both cases, complexes are probably formed by initial protein loading at the branch point followed by subsequent assembly on adjacent duplex sections. Binding to the Holliday junction and fork structures was compared over a greater range of protein concentration (Fig. 1C). The isotherms indicated a slight preference for 67RuvC binding to the X junction over

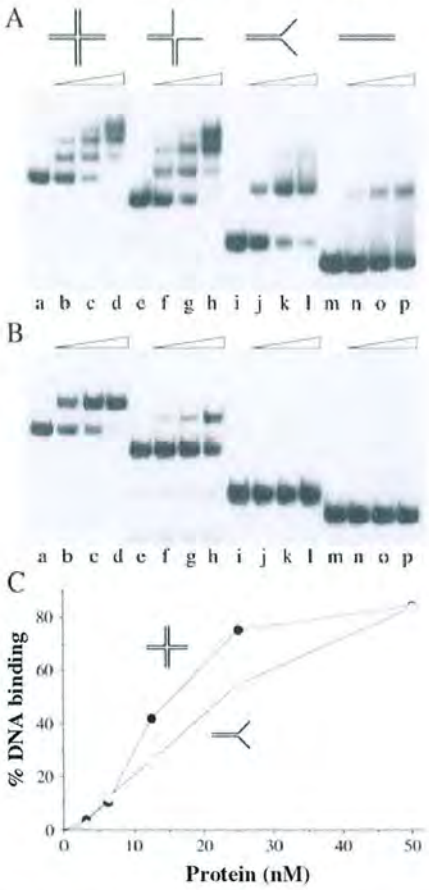

Fig. 1. DNA-branched structure binding by 67RuvC.
A and B. Binding reactions contained 0.3 nM ^{32}P-labelled Holliday (J11), flap, fork and duplex DNA as indicated and either 67RuvC (A) or *Ec*RuvC (B) at 12.5, 25 and 50 nM.
C. Comparison of 67RuvC binding to Holliday junction and fork structures. The binding reactions contained 0.3 nM ^{32}P-labelled J11 or F11 and 3.125, 6.25, 12.5, 25 and 50 nM 67RuvC protein. Data are the mean of two independent experiments and were quantified by phosphorimaging.

the fork; however, the availability of additional binding sites on the arms of the Holliday junction could account for this difference. We also compared binding of *Ec*RuvC to these substrates under identical conditions (Fig. 1B). *Ec*RuvC formed a single complex with Holliday junction and flap DNA (Fig. 1B, lanes b–d and f–h), the latter with significantly reduced affinity. No *Ec*RuvC–DNA complexes were detected on either the fork or duplex substrates (Fig. 1B, lanes j–l and n–p), in accord with previously published data (Benson and West, 1994; Takahagi *et al.*, 1994).

To assess DNA cleavage by 67RuvC, experiments were repeated at 37°C in the presence of 10 mM magnesium chloride, conditions optimal for *Ec*RuvC resolution (Dunderdale *et al.*, 1994; Shah *et al.*, 1994b). The results show that 67RuvC cleaves all three structures containing a branched component, with no degradative activity detected on duplex DNA (Fig. 2A). The products of Holliday junction cleavage migrated close to the position of duplex DNA (Fig. 2A, lanes b–d), potentially indicating formation of nicked duplex DNA. However, the cleavage

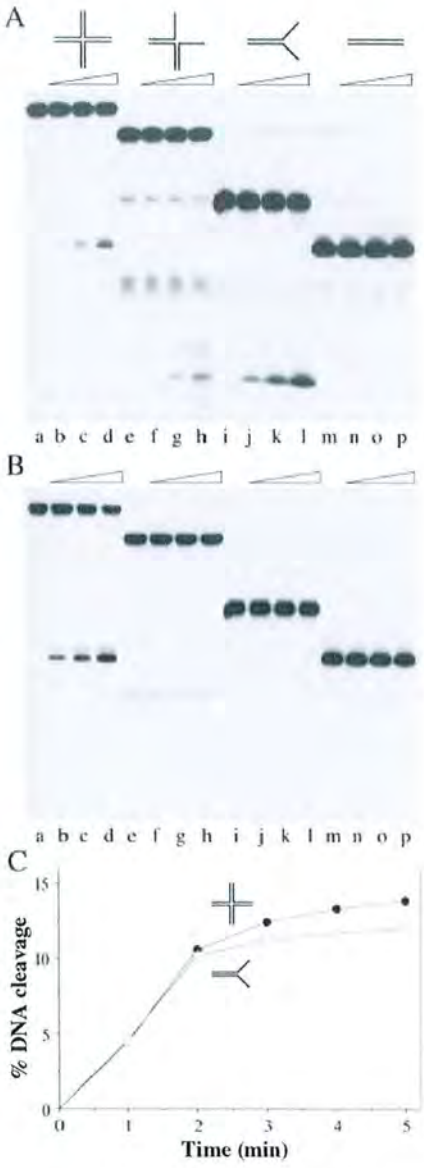

Fig. 2. DNA-branched structure cleavage by 67RuvC.
A and B. Cleavage reactions contained 0.3 nM ^{32}P-labelled Holliday (J11), flap, fork and duplex DNA as indicated and either 67RuvC (A) or *Ec*RuvC (B) at 12.5, 25 and 50 nM.
C. Time-course of Holliday junction and fork cleavage by 67RuvC. Bulk reactions contained 0.3 nM ^{32}P-labelled J11 junction or F11 fork and 67RuvC at 2.5 nM. Samples were removed and the reaction terminated at 0, 1, 2, 3, 4 and 5 min intervals. Data are the mean of two independent experiments and were quantified by phosphorimaging.

of flap and fork substrates suggests that 67RuvC has a significantly diminished specificity for four-way junctions (Fig. 2A, lanes h and l). The cleavage activity on junction and fork were examined in time-course experiments, with samples removed at intervals from bulk reactions (Fig. 2C). Cleavage of the two substrates was remarkably similar, especially in the initial rate of the reaction, with only a slight increase in the overall amount of product generated with the Holliday junction DNA over the fork. Similar results were obtained at higher protein concentrations (50 nM) with 51% of X junction cleaved after 2 min compared with 45% for the fork DNA. The reaction is rapid and tails off after this time point despite the availability of further substrate (Fig. 2C). This probably reflects protein instability under reaction conditions; both phage RusA- and Rap-resolving enzymes are similarly labile in junction cleavage assays (Chan *et al.*, 1997; Sharples *et al.*, 2004). Out of the four substrates tested, *Ec*RuvC only cleaved the Holliday junction (Fig. 2B) consistent with previous experiments showing bacterial RuvC is highly selective for X structures (Benson and West, 1994; Takahagi *et al.*, 1994).

Sequence specificity of Holliday junction resolution

To establish whether the products generated by 67RuvC cleavage on a Holliday junction were generated by bilateral strand scission, we mapped the location of incisions on two different four-way junction substrates (Fig. 3). The 50 bp junctions, containing 11 or 12 bp homologous cores constrained by heterologous sequences, are identical to those used previously to study *Ec*RuvC (Bennett *et al.*, 1993; Shah *et al.*, 1994a). Junction samples, each labelled in a different strand, were incubated with 67RuvC and the products of resolution separated on denaturing gels with a sequencing ladder to determine the position of cut sites. 67RuvC made a number of incisions on these junctions, positioned within or just outside the core of homology (Fig. 3B and D). The major sites located within the mobile core were symmetrically orientated, consistent

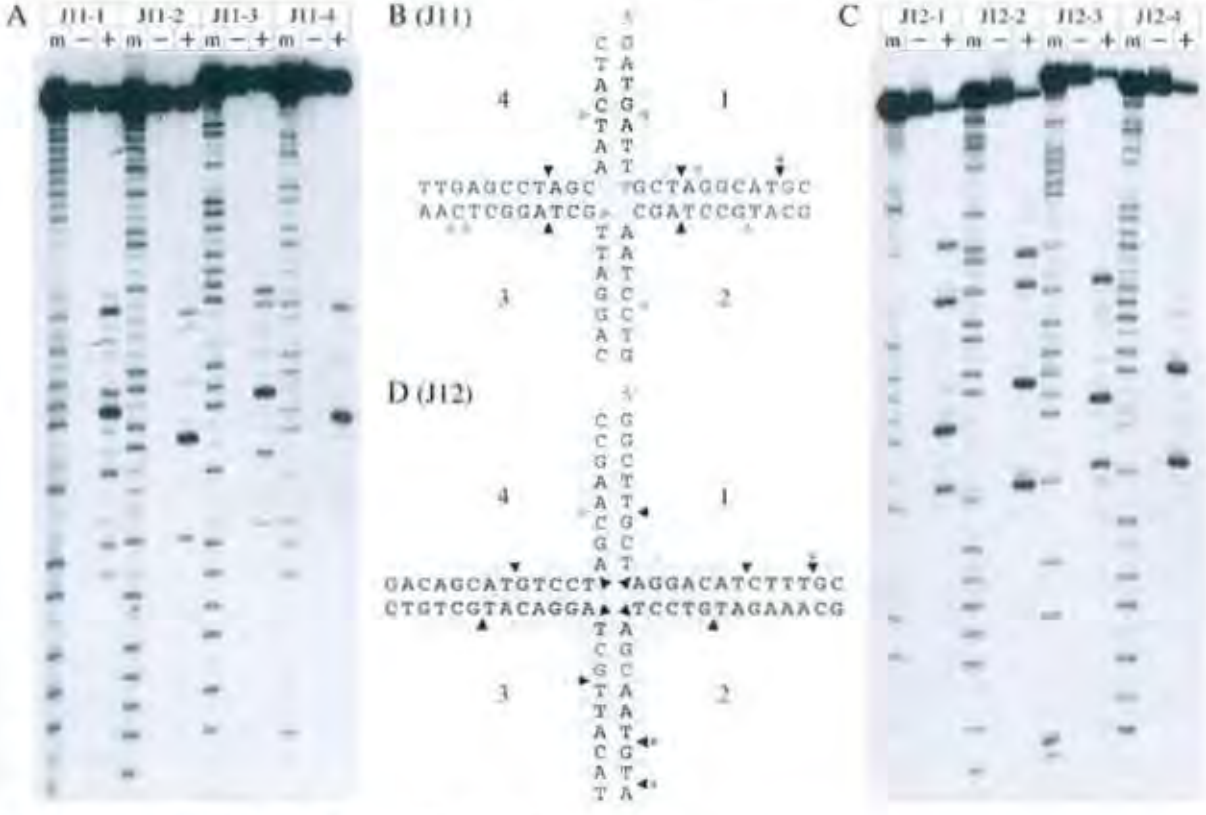

Fig. 3. Location of 67RuvC cleavage sites on Holliday junctions.
A and C. Denaturing gels showing the cleavage sites of 67RuvC on J11 and J12. Separate junction preparations of J11 or J12 (1.5 nM) each labelled in a different oligonucleotide (strands 1–4) were incubated with 50 nM 67RuvC. A Maxam-Gilbert A+G sequencing reaction on each labelled strand was used as a marker (m). Samples contained no protein (–) or 50 nM 67RuvC (+).
B and D. Location of cleavage sites at the cross-over in J11 and J12. Only part of the 50 bp junction sequence is shown with the homologous core region shaded. The position of cleavage is indicated by triangles at major (filled), intermediate (shaded) and minor (open) sites. Asterisks denote major sites that are asymmetrically orientated and located outside the core of sequence homology.

Table 1. Consensus for 67RuvC major cleavage sites on J11 and J12.

Number of sites	Major cut sites	Number of sites	Poorly cut or uncut sites 3′ of a T
8	CT↓AG	1	CT TG
4*	TT↓GC	2	GT CC
1	AT↓CT	3	AT TG
3*	AT↓GT	1	AT CC
1*	GT↓AA	1	AT CA
2*	AT↓GC	1	
Consensus	xT↓Rx		xT Yx

Cleavage sites (5′–3′) from the central core of homology in J11 and J12 were compiled from the data in Fig. 3. Four major cleavage sites lying outside this region are included in the list of sites (asterisks). R, purine; Y, pyrimidine.

with genuine Holliday junction resolution. Significantly, these cleavage sites differed from those found with *Ec*RuvC on these substrates; *Ec*RuvC preferentially cleaves at TT↓GC and 5′-GT↓CC in junctions J11 and J12 respectively (Connolly *et al.*, 1991; Bennett *et al.*, 1993; Shah *et al.*, 1994a). All of the major incisions made by 67RuvC occurred 3′ of a thymidine, with all but one being also 5′ of a purine residue (Table 1). This latter site (5′-AT↓CT) is located at the border of homology and heterology in J12 (Fig. 3D). Potential cleavage sites containing thymidine followed by a pyrimidine within the mobile core were either poorly cleaved or not cut at all, suggesting that cleavage is favoured at the consensus 5′-T↓$^{A}/_{G}$-3′ (Table 1; Fig. 3B and D). However, a number of major sites outside the region of homology and numerous other minor sites are not symmetrically related in these junctions (Fig. 3). The results suggest that 67RuvC is prone to cleave only one strand of a junction if a matching consensus site is unavailable for the second subunit to act. Such uncoupling of the dual strand cleavage reaction probably accounts for the products that migrate faster than a nicked duplex on neutral gels (see Fig. 7B, lane d).

Resolution symmetry

To confirm that bilateral resolution of Holliday junction was occurring at sites within the junction core, we probed for the formation of nicked duplex products using DNA ligase. J12 labelled in strand 2 was selected for these experiments as major symmetrical and asymmetrical cleavage sites are apparent (Fig. 3C and D). The experiment relies on the property of ligase to seal a nick in duplex DNA produced by symmetrical resolution, but not gaps or flaps resulting from cleavage asymmetry. Restoration of the cleaved DNA strand to its full length after ligation was monitored by denaturing PAGE. Two major cleavage products generated by 67RuvC located within the homologous core (sites 3 and 4) were significantly reduced after addition of T4 DNA ligase with a concomitant increase in the amount of the full-length strand (Fig. 4A, lanes b and c; Fig. 4B). These results fit with ligation of nicked duplex DNA and therefore symmetry-related junction cleavage. Similar results were obtained with *Ec*RuvC which cuts

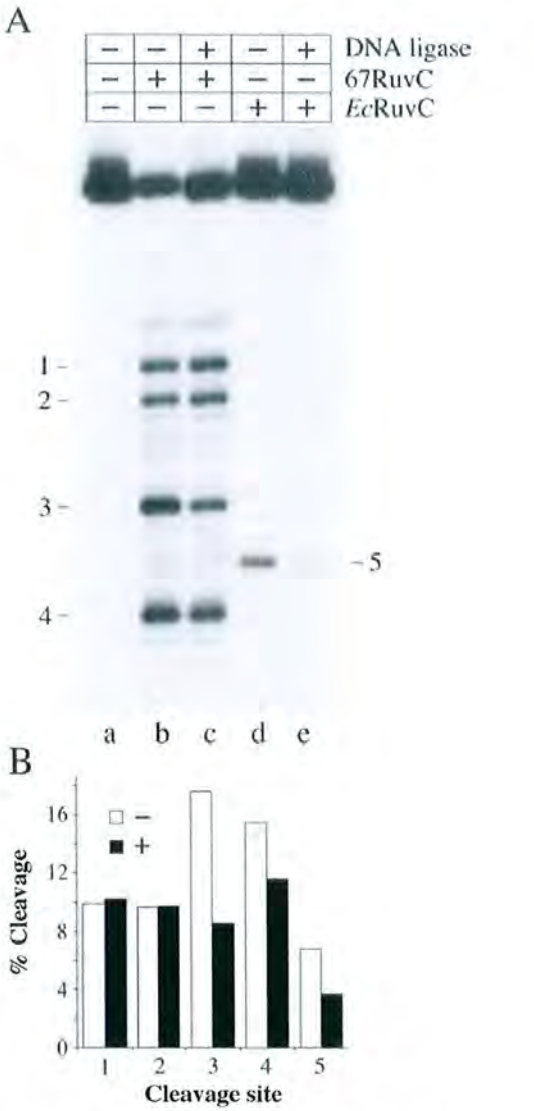

Fig. 4. Ligation of the products of 67RuvC Holliday junction resolution.
A. Reactions contained 1.5 nM J12 labelled in strand 2 and 50 nM of 67RuvC or *Ec*RuvC protein. Samples were incubated at 37°C for 10–15 min. One half of the reaction was terminated while 2.5 units of T4 DNA ligase was added to the other and incubation continued for a further 15 min. Samples were separated on a denaturing polyacrylamide gel.
B. Quantification of ligation at each cleavage site. Each 67RuvC (sites 1–4) or *Ec*RuvC (site 5) cleavage product, labelled in (A), was quantified by phosphorimaging. Values are expressed as a percentage of the total radioactivity units detected in each lane. Samples with (+) and without (–) DNA ligase are indicated.

poorly at a different location in J12 (Fig. 4; site 5: 5′-GT↓CC). In contrast, 67RuvC sites 1 and 2 located outside the 12 bp core of homology were not ligatable, indicating asymmetrical resolution to give products that cannot be sealed by DNA ligase (Fig. 4). We conclude that 67RuvC can resolve the Holliday structure to generate nicked duplexes, although it also readily cleaves at other positions where only one strand of the junction fits the preferred consensus. Single-strand nicking of junctions can also occur with *Ec*RuvC, especially on substrates lacking a homologous core.

Location of cleavage sites on fork substrates

The data from Figs 1 and 2 suggest that 67RuvC, unlike its bacterial counterpart, does not discriminate strongly between Holliday and fork substrates. We mapped the location of 67RuvC cleavage sites on two different fork substrates (F11 and F12), derived from the Holliday junctions J11 and J12, to determine the position of strand cleavage and evaluate any sequence preference in resolving these substrates (Fig. 5). As with the Holliday structure, cleavage of the forks occurred close to the branch point either in duplex or single-stranded sections or at the boundary between the two. In both cases, cleavage sites were extended slightly along the 3′ tail of the substrates (Fig. 5C and D). Major incisions were not restricted to 3′ of a thymidine, suggesting that recognition of the fork structure itself is more important than the particular nucleotide sequences at the branch point. The location of incisions fits with the neutral gel data (Fig. 5B) showing removal of both ssDNA arms with the products observed corresponding to which strand of the substrate was labelled with ^{32}P.

Holliday junction resolution in vivo

The observation that 67RuvC cleaves Holliday junctions

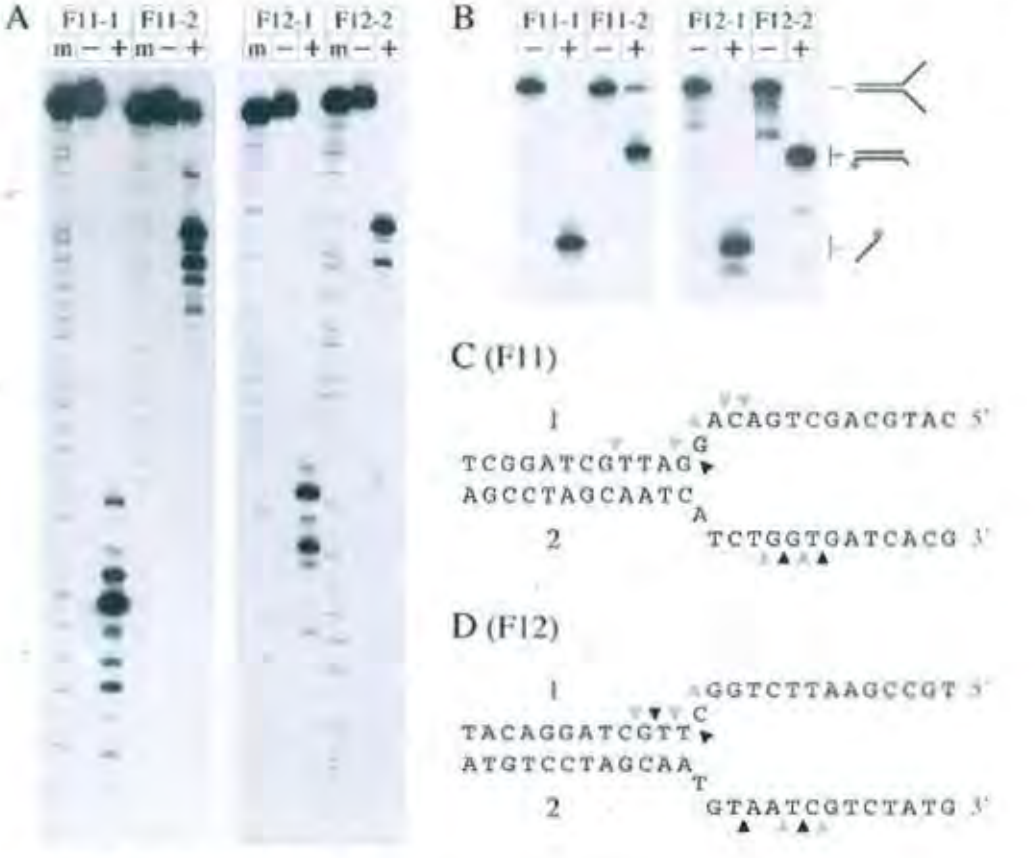

Fig. 5. Location of 67RuvC cleavage sites on fork DNA.
A. Denaturing gels showing the cleavage sites of 67RuvC on fork substrates derived from J11 and J12. Separate junction preparations of F11 or F12 (1.5 nM) each labelled in a different oligonucleotide (strands 1–2) were incubated with 50 nM 67RuvC. A Maxam-Gilbert A+G sequencing reaction on each labelled strand was used as a marker (m). Samples contained no protein (−) or 50 nM 67RuvC (+).
B. Neutral gels of the same reactions as in (A). The products of resolution are indicated on the right with an asterisk indicating the ^{32}P label.
C and D. Location of cleavage sites at the cross-over in F11 and F12. Only part of the 50 bp fork sequence is shown. The position of cleavage is indicated by triangles at major (filled), intermediate (shaded) and minor (open) sites.

Table 2. Effect of plasmids carrying 67RuvC on the survival of UV-irradiated *ruv* mutants.

	Fraction surviving (30 J m^{-2})					
Strain Genotype	SR2210 *ruvA*	N1057 *ruvB*	N2057 *ruvAB*	N4453 *ruvC*	N4454 *ruvABC*	AB1157 *ruv*$^+$
pFC105	0.28	0.00020	0.26	0.0066	0.17	0.53
pET24a	0.00048	0.00024	0.0016	0.00040	0.00078	0.68
pFC128	0.033	0.000078	0.086	0.00015	0.040	0.34
pET14b	0.00092	0.00032	0.00091	0.00027	0.00097	0.70
pRuv$^+$	0.36	0.19	0.42	0.52	0.22	
pUC/HSG	0.00075	0.00022	0.00096	0.00014	0.00082	

Plasmids used are indicated in the first column. pFC105 and pFC128 carry 67RuvC in pET24a and pET14b respectively. The Ruv$^+$ constructs (Sharples *et al.*, 1990; Sharples and Lloyd, 1991) used as controls were pGTI4 (RuvA), pGTI19 (RuvB), pGS711 (RuvAB), pGS762 (RuvC) and pPVA101 (RuvABC). Vector controls were either pUC18 or pHSG415. Cells carrying pPVA101 and pHSG415 were cultured at 30°C as these plasmids are unable to replicate at 37°C. Relevant strain genotypes are indicated.

in vitro prompted us to assay its ability to replace the function of *Ec*RuvC in *ruv* mutant strains. We utilized the UV light sensitivity of strains defective in different combinations of the RuvABC system to quantify the impact of clones carrying 67RuvC on Holliday junction resolution. The results, summarized in Fig. 6 and Table 2, reveal that plasmids carrying 67RuvC in wild-type (pFC105) or N-His_6-tagged (pFC128) forms are capable of suppressing the UV-sensitive repair phenotype of *ruvA*, *ruvAB* and *ruvABC* mutants. As resolution requires the participation of all three Ruv proteins in *E. coli*, these results confirm that 67RuvC can function as a Holliday junction resolvase *in vivo*. Furthermore, it does so independently of the RuvAB branch migration complex. The pFC128 (N-His_6-tagged 67RuvC) construct was less effective than pFC105 in improving the UV resistance of *ruv* mutants (Table 2). A greater negative effect on the survival of UV-irradiated *ruv*$^+$ strains probably accounts for this difference, suggesting that increased production of 67RuvC in pFC128 rather than addition of the histidine tag is responsible.

Surprisingly, both 67RuvC plasmids were unable to improve the UV repair defect in *ruvB* or *ruvC* mutant strains (Fig. 6; Table 2). The reason for this seems to be the presence of the *ruvA* gene which could block access of 67RuvC to the junction DNA. A similar, although less extreme, effect was seen with the RusA Holliday junction resolvase (Mandal *et al.*, 1993; Mahdi *et al.*, 1996). In fact, the presence of 67RuvC in strains carrying a chromosomal copy of *ruvA* had a severe negative effect on growth, with faster growing, UV-resistant revertants readily arising in the bacterial population. We suspected that these strains would carry mutations inactivating *ruvA* and so sequenced this gene from a fast-growing *ruvC* mutant carrying pFC128. We found a single nucleotide insertion of an adenine after position 187 in the *ruvA* gene which results in a frame shift immediately after Tyr63 in RuvA. The loss of both RuvA and RuvC function explains why 67RuvC plasmids confer UV resistance in this strain.

Both pFC105 and pFC128 slightly increased the UV light sensitivity of *ruv*$^+$ strains (Fig. 6; Table 2). This was dramatically enhanced when the original clone of 67RuvC (pGS905) was introduced, producing a survival profile similar to that of a *ruv* mutant derivative (Fig. 6). The extreme phenotype probably results from increased expression of 67RuvC creating double-stranded DNA breaks by multiple incision of replication forks which cannot be easily repaired in already UV-damaged cells. The inability to transform *recA* mutants with 67RuvC constructs fits with this conclusion. It is worth noting that the T7 promoter is inoperative in these strains and so 67RuvC expression must occur from fortuitous promoters located elsewhere in the vector. We attempted to assess the impact of RecG helicase activity on 67RuvC resolution as *ruv recG* strains exhibit an obvious synergism for recombination and DNA repair (Lloyd, 1991). However, as with the *recA* mutant, pFC128 could not be recovered in this resolution-deficient background.

Analysis of a resolution-defective mutant

To verify the junction and fork-cleaving activities of 67RuvC and its relationship with bacterial RuvC proteins, we constructed a mutant defective in one of four conserved residues known to comprise the catalytic site (see Fig. 9). The 67RuvC D8N mutant corresponds to *Ec*RuvC D7N characterized previously (Saito *et al.*, 1995; Hagan *et al.*, 1998). Relative to the wild-type protein, the mutant showed improved expression (see Fig. 8C, lane d) and larger quantities of purified protein were recovered. Its ability to bind and cleave branched DNA substrates was judged alongside wild-type 67RuvC. The D8N mutant showed slightly improved binding to DNA compared with the wild-type protein (Fig. 7A; data not shown). However, it was unable to cleave any of the four DNA substrates tested, including the Holliday junction and fork DNA (Fig. 7B, lanes e–g and l–n; data not

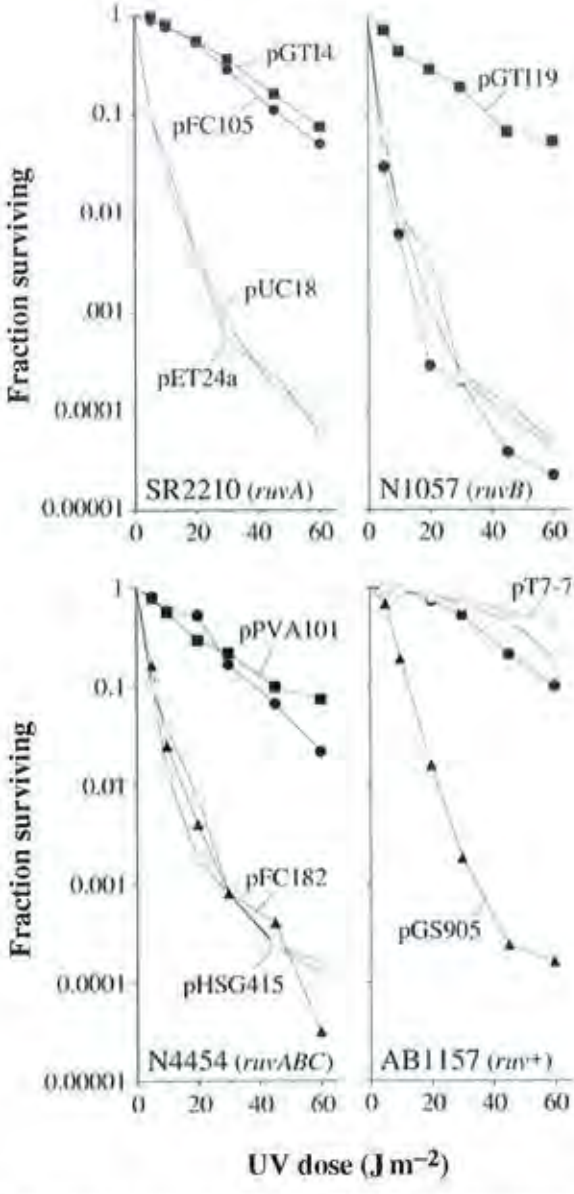

Fig. 6. Survival of UV-irradiated *E. coli ruv* mutants carrying plasmids encoding 67RuvC. Strain and genotype are indicated within each panel and symbols for pFC105 (67RuvC), pET24a and pUC18 labelled in the first panel only. Other plasmids used were: pGTI4 (*Ec*RuvA in pUC18), pGTI19 (*Ec*RuvB in pUC18), pPVA101 (*Ec*RuvABC in pHSG415), pFC182 (67RuvC D8N in pET24a) and pGS905 (67RuvC in pT7-7).

shown). Furthermore, the construct expressing 67RuvC D8N (pFC182) could be stably maintained in a *recA* mutant and failed to improve the UV sensitivity of a *ruvABC* mutant (Fig. 6). Neither did it confer a negative effect on survival of wild-type *E. coli* after exposure to UV light (data not shown). The results confirm that the D8N mutation inactivates all endonuclease activities of 67RuvC and that bacterial and phage RuvC proteins share a common catalytic core.

Formation of chromosomal breaks in vivo

The toxicity of bIL67 RuvC in *E. coli* matches that exhibited by the bIL66 *ruvC* gene (Bidnenko *et al.*, 1998). Clones expressing 67RuvC could not be recovered from recombination-defective *recA* or *ruv recG* mutant cells and pGS905 conferred extreme UV sensitivity on recombination-proficient strains. Taken together, these phenotypes suggest that 67RuvC also shares the capacity to introduce chromosomal breaks which are lethal in strains deficient in recombinational repair. To assay this directly, we investigated whether 67RuvC induces fragmentation of bacterial chromosomal DNA. Expression of 67RuvC in BL21-AI strains was induced and break formation was monitored in both chromosomal and plasmid DNA. 67RuvC D8N mutant and vector plasmids served as negative controls. Strains induced for expression of 67RuvC showed a much greater proportion of elongated and filamented cells consistent with damage to chromosome integrity (data not shown). Neither mutant nor vector plas-

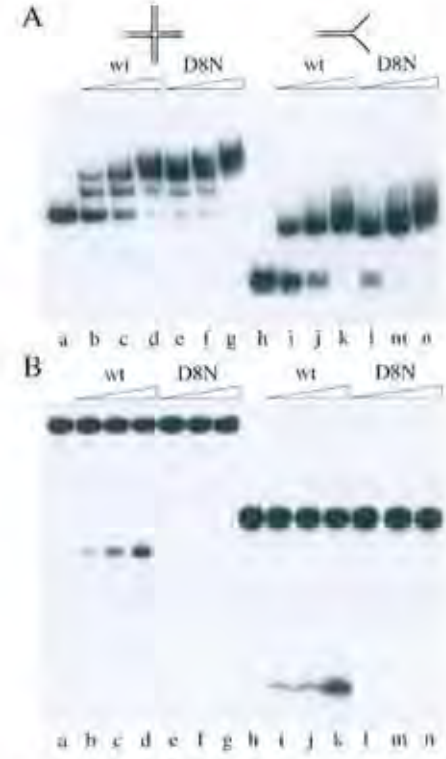

Fig. 7. DNA binding and cleavage activities of a 67RuvC D8N mutant protein. Binding (A) and cleavage (B) reactions contained 0.3 nM ^{32}P-labelled J11 junction or F11 fork DNA and either wild-type 67RuvC (lanes b–d and i–k) or 67RuvC D8N (lanes e–g and l–n) at 12.5, 25 and 50 nM in (A) or 25, 50 and 100 nM in (B).

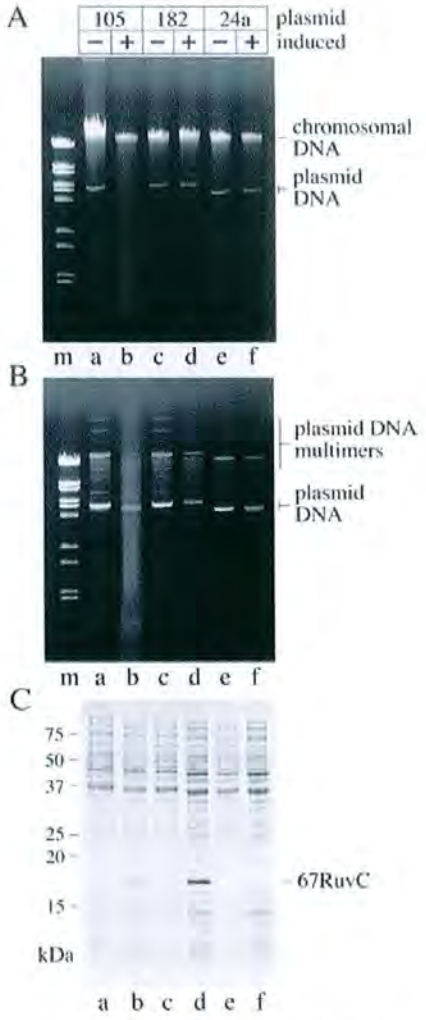

Fig. 8. Breakage of chromosomal DNA by 67RuvC. Agarose gel (0.8%) analysis of genomic DNA (A) or plasmid DNA (B) isolated from BL21-AI strains carrying pFC105 (67RuvC; lanes a and b), pFC182 (67RuvC D8N; lanes c and d) or pET24a (vector control; lanes e and f). The molecular weight marker (m) was λ Bst EII. Bacteria were grown to A_{650} of 0.6 and gene expression was induced in half of the culture by addition of IPTG and arabinose. Incubation was continued for 1 h before induced (+) and uninduced (–) cells were harvested by centrifugation. Total-cell proteins from these samples, ordered as in (A) and (B), were separated on 15% SDS-PAGE to monitor 67RuvC expression (C).

mids displayed this phenotype after addition of IPTG and arabinose. Genomic DNA harvested from these cells was separated on agarose gels. Generally similar quantities of genomic DNA were recovered from uninduced and induced BL21-AI carrying wild-type 67RuvC, D8N or the vector alone (Fig. 8A). A slight reduction in recovery and smearing of DNA was apparent in samples induced for 67RuvC expression (Fig. 8A, lane b). When plasmid DNA was analysed, a prominent smear of DNA fragments was observed in the induced 67RuvC sample (Fig. 8B, lane b). Because plasmid isolation protocols allow recovery of relatively small DNA molecules (less than 150 kb), this broad smear most probably constitutes an assortment of chromosome and plasmid fragments generated by nuclease action. Fragmentation is mediated by the endonucleolytic activity of 67RuvC as smearing was not detected from uninduced cells (Fig. 8B, lane a) or those induced for 67RuvC D8N protein (Fig. 8B, lane d). A similar set of results was obtained in strain BL21 pLysS (data not shown). Protein expression was also investigated by SDS-PAGE, revealing that small amounts of 67RuvC protein were produced from induced cells (Fig. 8C, lane b) with significantly more 67RuvC D8N formed under the same conditions (Fig. 8C, lane d). Destruction of chromosomal and/or plasmid DNA clearly has an adverse effect on wild-type protein expression. The results show that 67RuvC causes double-strand breaks in *E. coli*, most probably as a consequence of cleavage at replication forks distributed randomly across the chromosome.

Discussion

The RuvC protein from *E. coli* is a metal ion-dependent endonuclease with a conspicuous preference for resolving Holliday junction DNA by paired incisions across the branch point. In this study we demonstrated that the RuvC-like protein from *L. lactis* bIL67 can also function as a Holliday junction resolvase. As with its bacterial counterpart, 67RuvC resolved the X structure by introducing two symmetry-related incisions in opposing strands of the cross-over. Formation of nicked duplexes was validated by ligation of the products of resolution. 67RuvC substituted for *Ec*RuvC in resolution-deficient strains of *E. coli*, so long as the RuvA Holliday junction-binding protein was eliminated. *Ec*RuvC displays a predisposition for cleaving certain target sequences and although 67RuvC cleaved DNA at related sites, the complexity of this consensus was substantially reduced. While the precise nature of the 67RuvC consensus requires further investigation, the capacity to cleave at a 2 bp sequence, rather than 4 bp, means that extensive branch migration to locate favoured sites would be unnecessary for this phage protein. This feature may explain why 67RuvC acts as a resolving enzyme in cells lacking the RuvAB branch migration machinery. The RusA Holliday junction resolvase also functions in *E. coli* in the absence of RuvAB and similarly cleaves DNA at a 2 bp recognition sequence (Mandal *et al.*, 1993; Chan *et al.*, 1997; Giraud-Panis and Lilley, 1998). Reduced sequence selectivity may therefore be a prerequisite for those resolution endonucleases acting independently of specific branch migration enzymes.

While 67RuvC mediated Holliday junction resolution *in vitro* and *in vivo*, it also cleaved X structures outside the core of homology, frequently in an asymmetrical manner.

Moreover, it induced cleavage of a variety of other branched DNA junctions suggesting that phage RuvC has been adapted to suit a generalized role in branched DNA removal. Its toxicity in *E. coli* strains defective in recombinational repair appeared to result from this relaxed structure selectivity. Resolution of fork structures *in vitro* was not limited to the provisional consensus established on X junctions containing 11 and 12 bp mobile cores. 67RuvC, therefore, functionally resembles the resolving enzymes of T4, T7 and λ rather than the bacterial RuvC protein to which it is ancestrally related. In contrast with their cellular partners, phage resolvases tend to show a significant degree of flexibility in the substrates they accommodate, including forks, flaps, bulges, mismatches and even DNA adducts (reviewed in Lilley and White, 2001; Sharples, 2001). This versatility is undoubtedly beneficial for phage DNA metabolism, facilitating removal of any branches that could potentially impede genome packaging. For example, phage T4 and T7 mutants defective in their respective resolvases accumulate branched DNA which cannot be packaged (Kemper and Brown, 1976; Tsujimoto and Ogawa, 1978). The ability to cleave replication fork structures also contributes to breakdown of the host chromosome. T7 endonuclease I is known to function in this way, releasing nucleotides to supplement phage DNA replication (Center *et al.*, 1970). Phage RuvC proteins may serve in a similar capacity during lytic development in *L. lactis*, as suggested by the extensive degradation of *E. coli* genomic DNA after induction of 67RuvC expression.

How did a group of Gram-positive phages procure a Gram-negative RuvC-like resolving enzyme? Perhaps the ancestor of these phages arose in a Gram-negative host, acquired the *ruvC* gene by illegitimate recombination and passed it on to its descendants. In support of such a scenario, there is credible evidence for an evolutionary link between lactococcal phages and lambdoid coliphages based on similarities in their genomic content and organization (Brussow, 2001). In addition, phages and prophages with resident *ruvC* genes are not uncommon, indicating that uptake of host junction endonucleases is a relatively frequent event. Examples include the Rosebush phage from Mycobacteria, φAsp2 from *Actinoplanes* and φKZ from *Pseudomonas aeruginosa*. The RuvC endonucleases from these sources may be modified for general branch cleavage as are those from Lactococci and Streptococci, although they resemble more closely RuvC proteins from bacteria. Alternatively, RuvC could have been acquired by horizontal gene transfer rather than vertical transmission. This could have occurred during contact between their host and Gram-negatives in the mammalian gut. Donation of cellular genes from *L. lactis* to enterobacteria has been documented (Bolotin *et al.*, 2004). However, there are no known instances of Lactococcal or Streptococcal phages infecting Gram-negative organisms. Several Lactococcal phages of the P335 species (e.g. r1t, Tuc 2009) carry an alternative Holliday junction resolvase, RusA (Sharples *et al.*, 1999; 2002). One of these, from *L. lactis* phage r1t, resolves Holliday junctions *in vitro* and substitutes for *E. coli* RuvC *in vivo* (Sharples *et al.*, 2002). Like 67RuvC, r1t RusA shows a propensity to cleave various DNA branches in addition to the Holliday structures. This is significant, as *E. coli* RusA displays a Holliday junction cleavage specificity akin to that of *Ec*RuvC (Bolt and Lloyd, 2002). Lactococcal phages therefore acquired distinct resolution systems, both of which evolved into Holliday junction resolvases with degenerate branch specificity. Interestingly, none of these phages carry RecU, the alternative to RuvC in Gram-positive bacteria (Ayora *et al.*, 2004), even though it is present in the chromosomes of both Lactococci and Streptococci.

High resolution crystal structures of *Ec*RuvC (Ariyoshi *et al.*, 1994) and the distantly related *Sp*Ccr1/Ydc2-resolving enzyme from fission yeast are available (Ceschini *et al.*, 2001; Fig. 9). Thus, an analysis of the architectural features responsible for the remarkable disparity in branched DNA selectivity between phage and bacterial RuvC proteins can be undertaken. Unsurprisingly, given that both phage and bacterial RuvC proteins can resolve Holliday junctions, the structures appear to have much in common. Phage RuvC proteins retain equivalents of the four acidic residues (D7, E66, D138, D141) known to coordinate magnesium binding for phosphodiester bond hydrolysis in *Ec*RuvC (Saito *et al.*, 1995). One of these (D138) is replaced by asparagine in 67RuvC although this, or an adjacent aspartic acid, could potentially substitute as a metal ligand (Fig. 9A). We discovered that a 67RuvC D8N mutant protein, equivalent to D7N in *Ec*RuvC, was defective in DNA cleavage both *in vitro* and *in vivo*. Previous studies revealed that an E67Q mutant of bIL66 RuvC, corresponding to E66Q in *Ec*RuvC, similarly blocked nucleolytic activity. Taken together, these results suggest that the active site of phage RuvC proteins is structurally analogous to their relatives from Gram-negative bacteria.

Several additional features of *Ec*RuvC architecture are apparently conserved in the phage RuvC family (Fig. 9A), including the αB helix which forms the main interface between subunits in the homodimer (Ariyoshi *et al.*, 1994; Ichiyanagi *et al.*, 1998). Any alterations in the configuration of this helix could affect either subunit arrangement or integrity of dimerization and therefore influence targeting of particular DNA structures. Phage RuvC proteins lack an equivalent of F69, a residue implicated in disrupting base stacking interactions, possibly to position the scissile phosphate for hydrolysis (Yoshikawa *et al.*, 2001). Insertion of the aromatic ring of phenylalanine close to one of the thymidines at the TT dinucleotide target site favoured by *Ec*RuvC could influence sequence specificity

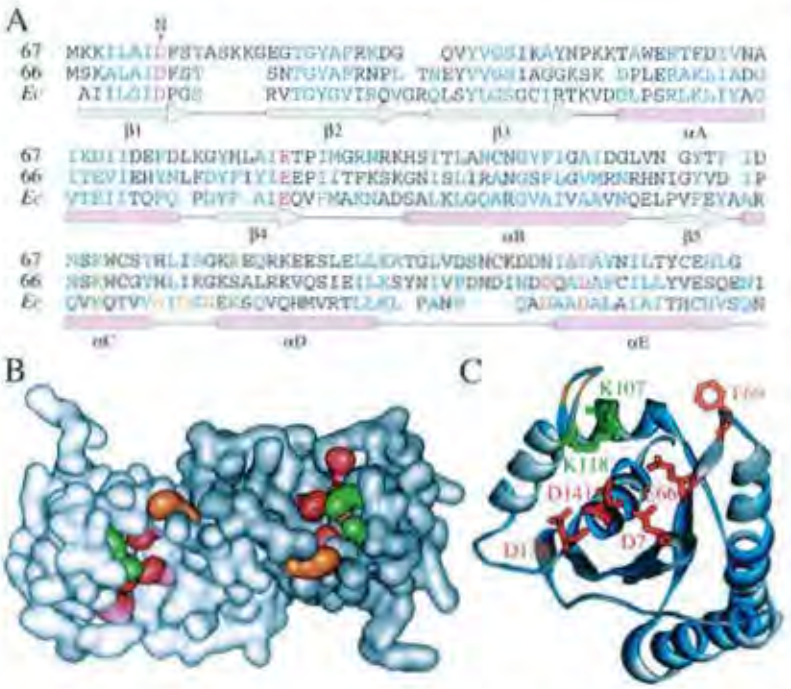

Fig. 9. Structural similarities between phage and *Ec*RuvC proteins.
A. The sequences of RuvC proteins from phages bIL66 and bIL67 are aligned with *Ec*RuvC. Structural features derived from the latter are indicated below the alignment. Conserved residues known to play important functions in *Ec*RuvC are highlighted; these include catalytic acidic residues (red), F69 (orange) implicated in base stacking interactions, K107 and K118 (green) that protrude into the DNA-binding cleft, and G112, G114 and A116 (yellow) that constitute a highly conserved loop in bacterial RuvC sequences. Other conserved residues between the three proteins are in blue.
B. Structure of the *Ec*RuvC dimer.
C. Ribbon structure of a single *Ec*RuvC subunit highlighting residues important in RuvC junction binding and resolution.
The colouring of residues in (B) and (C) matches those delineated in (A).

(Yoshikawa *et al.*, 2001). The yeast Cce1- and Ydc2-resolving enzymes also lack an appropriately positioned aromatic residue and, interestingly, neither requires a TT dinucleotide for resolution activity (White and Lilley, 1996; Whitby and Dixon, 1998).

Apart from these minor variations, the major difference between the RuvC proteins from phage and bacteria lies within the region of αC and αD (Fig. 9A). These two helices are strictly conserved in bacterial RuvC proteins and comprise the wall flanking the outer side of the DNA-binding cleft of *Ec*RuvC (Ariyoshi *et al.*, 1994; Hagan *et al.*, 1998). It is possible that changes in this region are largely responsible for the observed differences in structure recognition by the phage and *Ec*RuvC proteins and are also likely to influence sequence selectivity (Sharples *et al.*, 1999). Amino acids residing in these helices are already known to be important for DNA binding and sequence recognition by *Ec*RuvC (Hagan *et al.*, 1998; Ichiyanagi *et al.*, 1998; Yoshikawa *et al.*, 2000). Two invariant basic residues in *Ec*RuvC (K107 and K118), one in each alpha helix, have been identified as performing a key role in phosphodiester backbone recognition (Yoshikawa *et al.*, 2000) (Fig. 9C). In Ydc2, K107 is replaced by tyrosine (Y172) and probably helps coordinate metal binding (Ceschini *et al.*, 2001). The five-residue intervening loop between these helices is strikingly conserved in bacterial RuvC proteins; the consensus GxG/AxA/G loop (where x refers to any amino acid) was found in all but one of 137 bacterial RuvC sequences available in the protein database; the single exception from *Rubrobacter xylanophilus* has an SYGRA loop signature. These small glycine and alanine residues most probably maintain loop integrity. Mutations within this loop have previously been shown to alter the sequence specificity of resolution by *Ec*RuvC (Hagan *et al.*, 1998) and probably do so by altering the position of helices C and D and as a result the respective orientations of K107 and K118. This arrangement is not peculiar to the bacterial RuvC group as 67RuvC possesses potential equivalents of K107 (K110) in αC and K118 (K120), only one residue removed, in αD (Fig. 9A). These helices may well be conserved in 67RuvC but the configuration of the K110/K120 pairing may be significantly altered by the changes in the intervening loop structure, with important consequences for the architecture of the DNA-binding cleft. Caution is needed in interpreting the importance of these potential differences in the absence of structural information from a phage RuvC protein. For instance, the ancestrally related Ydc2/Cce1-resolving enzyme also differs from *Ec*RuvC in the region of αC and αD, yet retains specificity for Holliday junction DNA (Ceschini *et al.*, 2001). Structural data on phage RuvC proteins, preferably in complex with Holliday junction or fork DNA, are therefore needed to clarify these uncertainties. The 67RuvC D8N mutant is an ideal candidate for X-ray crystallography as it can be purified in large quantities and is incapable of degrading DNA. Such an approach should help reveal how the RuvC family of Holliday junction resolvases achieve structure and sequence selectivity of DNA cleavage.

Experimental procedures

Strains and UV sensitivity assays

Escherichia coli K12 *ruv* mutants strains, SR2210 (*ruvA200*), N1057 (*ruvB4*), N2057 (*ruvAB60*::Tn*10*), N4453 (Δ*ruvC*::*cat*) and N4454 (Δ*ruvABC*::*cat*), are derivatives of the *ruv*$^+$ wild-type strain, AB1157 (Sargentini and Smith, 1989; Sharples *et al.*, 1990; Seigneur *et al.*, 1998). Strains carrying appropriate clones were measured for UV resistance by growing cells to an A_{450} of 0.4 and spotting dilutions onto agar plates. These were exposed to UV light at a dose rate of 1 J m^{-2} s^{-1} and the fraction surviving was calculated with reference to an unirradiated control plate. *E. coli* BL21 (DE3) strains were employed as a host for expression of 67RuvC from appropriate constructs.

Phage bIL67 RuvC cloning

The *ruvC*-like gene (*ORF23*) from *L. lactis* phage bIL67 (Schouler *et al.*, 1994) was amplified from genomic DNA using Pfx DNA polymerase and primers 67-1 (5′-GGATAC ATATGAAGAAAATT-3) and 67-2 (5′-TTTATAAGGGATCCATT AAC-3′). The polymerase chain reaction (PCR) product was digested with *Nde*I and *Bam*HI (underlined) and inserted into pT7-7 to give pGS905. This construct failed to express protein at high levels and so was subcloned into pET24a (pFC105) and pET14b (pFC128), the latter creating an N-terminal histidine-tagged fusion. A 67RuvC D8N mutant was generated by PCR using 67-2 and 67-3 (5′-GATACATATG AAG AAAATTTTAGCTATT**A**ATTTCAG-3′) containing a single-nucleotide substitution within the *ruvC* coding region (bold). The restricted DNA product was inserted into pET24a to generate pFC182. Integrity of the cloned bIL67 *ruvC* gene and the D8N mutant was confirmed by DNA sequencing. Plasmid and chromosomal DNA samples were recovered from *E. coli* using kits purchased from Qiagen and Sigma respectively.

Proteins

*Ec*RuvC protein was purified as described (Dunderdale *et al.*, 1994). 67RuvC was purified from 750 ml of *E. coli* BL21 (DE3) pLysS carrying pFC105. Cells were grown to A_{600} of 0.5 in LB containing ampicillin (150 μg ml^{-1}) and expression was induced with 1 mM IPTG for 3 h at 37°C. Harvested cells were resuspended in 20 mM Tris-HCl, 1 mM EDTA, 1 M NaCl, lysed by sonication and dialysed in Buffer A (20 mM Tris-HCl pH 7, 1 mM EDTA, 0.5 mM DTT, 200 mM KCl). The lysate was applied to a 5 ml reactive blue-4 agarose column in Buffer A and the flow-through containing 67RuvC collected and loaded onto a 2 ml Q sepharose column. 67RuvC eluted between 0.40 and 0.55 M KCl and pooled fractions were dialysed in Buffer A at pH 8. This sample was applied to a 1 ml *ss*DNA cellulose column, eluting between 0.35 and 0.65 M KCl. The pooled fractions were diluted to 0.25 M KCl with Buffer A pH 8 without KCl and applied to a 2 ml phosphocellulose column. The 67RuvC protein eluted at 0.8–1 M KCl and was dialysed in Buffer A pH 8 containing 150 mM KCl and 50% glycerol. Samples were stored in aliquots at −80°C. A total of 86 μg was recovered at a concentration of 0.36 mg ml^{-1}. Larger quantities of purified 67RuvC (1.54 mg) were obtained from 6 l of cells utilizing interactions with phosphocellulose, reactive green-19 agarose and reactive blue-4 agarose matrices. 67RuvC associates with reactive blue-4 agarose prepared from the powdered matrix but not that purchased in liquid form. Both protein preparations displayed similar cleavage activities on X and fork DNA assays.

The 67RuvC D8N mutant was purified from 1 l of BL21 Codon+ carrying pFC182. Cells were disrupted as for the wild-type and the cleared lysate applied to a 3 ml phosphocellulose column. The D8N protein eluted between 0.6 and 0.9 M KCl and was dialysed against Buffer A with 0.2 M KCl causing the majority of the protein to precipitate. Centrifugation at 13 000 r.p.m. for 10 min yielded a protein pellet which was resolubilized in Buffer A containing 0.6 M KCl. This sample was sufficiently pure for biochemical analysis and so was stored at −80°C in Buffer A pH 8 containing 0.5 M KCl and 50% glycerol. A total of 8.54 mg 67RuvC D8N was recovered at a concentration of 1.22 mg ml^{-1}. A further 5.27 mg of pure 67RuvC D8N protein was obtained from 1 l of cells following protocols similar to those used with wild-type 67RuvC. Precipitation upon dialysis after an initial phosphocellulose column was avoided by keeping the KCl concentration above 0.35 M. All fractions of 67RuvC D8N behaved similarly in binding and resolution assays.

Protein concentrations were determined by a modified Bradford Assay (Bio-Rad) using bovine serum albumin (BSA) as a standard; amounts of protein are expressed as moles of dimeric protein. All experiments were performed with the untagged version of 67RuvC. Restriction endonucleases, T4 DNA ligase, T4 polynucleotide kinase and Pfx DNA polymerase were purchased from Invitrogen.

Construction of 50 bp DNA substrates

Synthetic Holliday junction structures were made by annealing oligonucleotides essentially as described (Parsons *et al.*, 1990). Each junction contained at least one oligonucleotide labelled with [γ-^{32}P]-ATP at its 5′ end using T4 polynucleotide kinase. The sequences of 50 base oligonucleotides used in the construction of mobile junctions harbouring 11 bp (J11) or 12 bp (J12) homologous cores are as published (Sharples *et al.*, 2004). F11 and F12 fork structures were made by annealing two strands from junction J11 (strands 2 and 3) and J12 (strands 1 and 2) respectively. The flap structure consisted of J11 with strand 4 omitted. The linear duplex was made by annealing strand 1 of J11 to 5′-GCTTCTCTTGCGG AAAGCATGCCTAGCAATCATCTGGTGATCACGTCGCC-3′.

DNA binding and cleavage assays

Band shift assays (20 μl) using ^{32}P-labelled junction DNA substrates were carried out in 50 mM Tris-HCl pH 8.0, 5 mM EDTA, 1 mM dithiothreitol, 5% glycerol, 100 μg ml^{-1} BSA. Samples were incubated on ice for 15 min before separation on 4% PAGE in 6.7 mM Tris-HCl pH 8.0, 3.3 mM sodium acetate, 2 mM EDTA at 10 V cm^{-1} for 2 h. Cleavage of ^{32}P-labelled DNA was assayed at 37°C, usually for 15 min, in 50 mM Tris-HCl pH 8.0, 1 mM dithiothreitol, 100 μg ml^{-1} BSA, 10 mM $MgCl_2$. Reactions (20 μl) were terminated by the addi-

tion of 5 µl of 100 mM Tris-HCl pH 8.0, 2.5% SDS, 100 mM EDTA, 10 mg ml^{-1} proteinase K and incubated for a further 10 min at 37°C. After addition of 5 µl of loading buffer (0.25% w/v bromophenol blue, 0.25% w/v xylene cyanol, 15% v/v Ficoll type 400), 15 µl was electrophoresed on 10% polyacrylamide gels in 90 mM Tris-borate, 2 mM EDTA at 12 V cm^{-1} for 2 h. Gels were dried onto filter paper and analysed by autoradiography and phosphorimaging.

Mapping DNA junction cleavage sites

Separate Holliday junction preparations were made so that each was ^{32}P-labelled in a different strand. Reactions were deproteinized in stop buffer and mixed with 0.3% (w/v) bromophenol blue, 0.3% (w/v) xylene cyanol, 10 mM EDTA, 97.5% formamide. Samples of 2–5 µl were boiled for 2 min and separated by gel electrophoresis on 15% polyacrylamide-7 M urea at 4.4 V cm^{-1} for 4–5 h. A+G Maxam-Gilbert sequencing reactions were performed on each oligonucleotide and run on the same gel to provide markers. Cleavage sites were mapped by reference to the sequencing ladder with a 1.5 base allowance made to compensate for the nucleoside eliminated in the sequencing reaction.

Acknowledgements

We thank Elena Bidnenko for the kind gift of bIL67 genomic DNA and John Rafferty for advice on structural differences between phage and bacterial RuvC proteins. This work was supported by the Biotechnology and Biological Sciences Research Council and the Royal Society.

References

Ariyoshi, M., Vassylyev, D.G., Iwasaki, H., Nakamura, H., Shinagawa, H., and Morikawa, K. (1994) Atomic structure of the RuvC resolvase: a Holliday junction-specific endonuclease from *E. coli*. *Cell* **78:** 1063–1072.

Ayora, S., Carrasco, B., Doncel, E., Lurz, R., and Alonso, J.C. (2004) *Bacillus subtilis* RecU protein cleaves Holliday junctions and anneals single-stranded DNA. *Proc Natl Acad Sci USA* **101:** 452–457.

Bennett, R.J., Dunderdale, H.J., and West, S.C. (1993) Resolution of Holliday junctions by RuvC resolvase: cleavage specificity and DNA distortion. *Cell* **74:** 1021–1031.

Benson, F.E., and West, S.C. (1994) Substrate specificity of the *Escherichia coli* RuvC protein. *J Biol Chem* **269:** 5195–5201.

Bidnenko, E., Ehrlich, S.D., and Chopin, M.-C. (1998) *Lactococcus lactis* phage operon coding for an endonuclease homologous to RuvC. *Mol Microbiol* **28:** 823–834.

Bolotin, A., Quinquis, B., Sorokin, A., and Ehrlich, D.S. (2004) Recent genetic transfer between *Lactococcus lactis* and enterobacteria. *J Bacteriol* **186:** 6671–6677.

Bolt, E.L., and Lloyd, R.G. (2002) Substrate specificity of RusA resolvase reveals the DNA structures targeted by RuvAB and RecG *in vivo*. *Mol Cell* **10:** 187–198.

Brussow, H. (2001) Phages of dairy bacteria. *Annu Rev Microbiol* **55:** 283–303.

Center, M.S., Studier, F.W., and Richardson, C.C. (1970) The structural gene for a T7 endonuclease essential for phage DNA synthesis. *Proc Natl Acad Sci USA* **65:** 242–248.

Ceschini, S., Keeley, A., McAlister, M.S., Oram, M., Phelan, J., Pearl, L.H., *et al.* (2001) Crystal structure of the fission yeast mitochondrial Holliday junction resolvase Ydc2. *EMBO J* **20:** 6601–6611.

Chan, S.N., Harris, L., Bolt, E.L., Whitby, M.C., and Lloyd, R.G. (1997) Sequence-specificity and biochemical characterization of the RusA Holliday junction resolvase of *Escherichia coli*. *J Biol Chem* **272:** 14873–14882.

Chan, S.N., Vincent, S.D., and Lloyd, R.G. (1998) Recognition and manipulation of branched DNA by the RusA Holliday junction resolvase of *Escherichia coli*. *Nucleic Acids Res* **26:** 1560–1566.

Connolly, B., Parsons, C., Benson, F.E., Dunderdale, H.J., Sharples, G.J., Lloyd, R.G., and West, S.C. (1991) Resolution of Holliday junctions *in vitro* requires *Escherichia coli ruvC* gene product. *Proc Natl Acad Sci USA* **88:** 6063–6067.

Dunderdale, H.J., Sharples, G.J., Lloyd, R.G., and West, S.C. (1994) Cloning, overexpression, purification and characterization of the *Escherichia coli* RuvC Holliday junction resolvase. *J Biol Chem* **269:** 5187–5194.

Fogg, J.M., Schofield, M.J., White, M.F., and Lilley, D.M. (1999) Sequence and functional-group specificity for cleavage of DNA junctions by RuvC of *Escherichia coli*. *Biochemistry* **38:** 11349–11358.

Giraud-Panis, M.J., and Lilley, D.M. (1998) Structural recognition and distortion by the DNA junction-resolving enzyme RusA. *J Mol Biol* **278:** 117–133.

Hagan, N.F.P., Vincent, S.D., Ingleston, S.M., Sharples, G.J., Bennett, R.J., West, S.C., and Lloyd, R.G. (1998) Sequence-specificity of Holliday junction resolution: identification of RuvC mutants defective in metal binding and target site recognition. *J Mol Biol* **281:** 17–29.

Hargreaves, D., Rice, D.W., Sedelnikova, S.E., Artymiuk, P.J., Lloyd, R.G., and Rafferty, J.B. (1998) Crystal structure of *E. coli* RuvA with bound DNA Holliday junction at 6 Å resolution. *Nat Struct Biol* **5:** 441–446.

Ichiyanagi, K., Iwasaki, H., Hishida, T., and Shinagawa, H. (1998) Mutational analysis on structure-function relationship of a Holliday junction specific endonuclease RuvC. *Genes Cells* **3:** 575–586.

Kemper, B., and Brown, D.T. (1976) Function of gene 49 of bacteriophage T4. II. Analysis of intracellular development and the structure of very fast sedimenting DNA. *J Virol* **18:** 1000–1015.

Kosak, H.G., and Kemper, B.W. (1990) Large-scale preparation of T4 endonuclease VII from over-expressing bacteria. *Eur J Biochem* **194:** 779–784.

Lilley, D.M.J., and White, M.F. (2001) The junction-resolving enzymes. *Nat Rev Mol Cell Biol* **2:** 433–443.

Lloyd, R.G. (1991) Conjugational recombination in resolvase-deficient *ruvC* mutants of *Escherichia coli* K-12 depends on *recG*. *J Bacteriol* **173:** 5414–5418.

Mahdi, A.A., Sharples, G.J., Mandal, T.N., and Lloyd, R.G. (1996) Holliday junction resolvases encoded by homologous *rusA* genes in *Escherichia coli* K-12 and phage 82. *J Mol Biol* **257:** 561–573.

Mandal, T.N., Mahdi, A.A., Sharples, G.J., and Lloyd, R.G. (1993) Resolution of Holliday intermediates in recombination and DNA repair: indirect suppression of *ruvA*, *ruvB* and *ruvC* mutations. *J Bacteriol* **175:** 4325–4334.

de Massy, B., Weisberg, R.A., and Studier, F.W. (1987) Gene 3 endonuclease of bacteriophage T7 resolves conformationally branched structures in double-stranded DNA. *J Mol Biol* **193:** 359–376.

Parsons, C.A., Kemper, B., and West, S.C. (1990) Interaction of a four-way junction in DNA with T4 endonuclease VII. *J Biol Chem* **265:** 9285–9289.

Saito, A., Iwasaki, H., Ariyoshi, M., Morikawa, K., and Shinagawa, H. (1995) Identification of four acidic amino acids that constitute the catalytic center of the RuvC Holliday junction resolvase. *Proc Natl Acad Sci USA* **92:** 7470–7474.

Sargentini, N.J., and Smith, K.C. (1989) Role of *ruvAB* genes in UV- and γ-radiation and chemical mutagenesis in *Escherichia coli*. *Mutation Res* **215:** 115–129.

Schouler, C., Ehrlich, S.D., and Chopin, M.C. (1994) Sequence and organization of the lactococcal prolate-headed bIL67 phage genome. *Microbiology* **140:** 3061–3069.

Seigneur, M., Bidneko, V., Ehrlich, S.D., and Michel, B. (1998) RuvAB acts at arrested replication forks. *Cell* **95:** 419–430.

Shah, R., Bennett, R.J., and West, S.C. (1994a) Genetic recombination in *E. coli*: RuvC protein cleaves Holliday junctions at resolution hotspots *in vitro*. *Cell* **79:** 853–864.

Shah, R., Bennett, R.J., and West, S.C. (1994b) Activation of RuvC Holliday junction resolvase *in vitro*. *Nucleic Acids Res* **22:** 2490–2497.

Sharples, G.J. (2001) The X-philes: structure-specific endonucleases that resolve Holliday junctions. *Mol Microbiol* **39:** 823–834.

Sharples, G.J., and Lloyd, R.G. (1991) Resolution of Holliday junctions in *Escherichia coli*: identification of the *ruvC* gene product as a 19-kilodalton protein. *J Bacteriol* **173:** 7711–7715.

Sharples, G.J., Benson, F.E., Illing, G.T., and Lloyd, R.G. (1990) Molecular and functional analysis of the *ruv* region of *Escherichia coli* K-12 reveals three genes involved in DNA repair and recombination. *Mol Gen Genet* **221:** 219–226.

Sharples, G.J., Ingleston, S.M., and Lloyd, R.G. (1999) Holliday junction processing in bacteria: insights from the evolutionary conservation of RuvABC, RecG, and RusA. *J Bacteriol* **181:** 5543–5550.

Sharples, G.J., Bolt, E.L., and Lloyd, R.G. (2002) RusA proteins from the extreme thermophile *Aquifex aeolicus* and lactococcal phage r1t resolve Holliday junctions. *Mol Microbiol* **44:** 549–559.

Sharples, G.J., Curtis, F.A., McGlynn, P., and Bolt, E.L. (2004) Holliday junction binding and resolution by the Rap structure-specific endonuclease of phage lambda. *J Mol Biol* **340:** 739–751.

Stasiak, A., Tsaneva, I.R., West, S.C., and Benson, C.J.B., Yu, X., and Egelman, E.H. (1994) The *E. coli* RuvB branch migration protein forms double hexameric rings around DNA. *Proc Natl Acad Sci USA* **91:** 7618–7622.

Takahagi, M., Iwasaki, H., and Shinagawa, H. (1994) Structural requirements of substrate DNA for binding to and cleavage by RuvC, a Holliday junction resolvase. *J Biol Chem* **269:** 15132–15139.

Tsujimoto, Y., and Ogawa, H. (1978) Intermediates in genetic recombination of bacteriophage T7 DNA. Biological activity and the roles of gene 3 and gene 5. *J Mol Biol* **125:** 255–273.

West, S.C. (1992) Enzymes and molecular mechanisms of genetic recombination. *Annu Rev Biochem* **61:** 603–640.

Whitby, M.C., and Dixon, J. (1998) Substrate specificity of the *Sp*CCE1 Holliday junction resolvase of *Schizosaccharomyces pombe*. *J Biol Chem* **273:** 35063–35073.

White, M.F., and Lilley, D.M.J. (1996) The structure-selectivity and sequence-preference of the junction-resolving enzyme CCE1 of *Saccharomyces cerevisiae*. *J Mol Biol* **257:** 330–341.

Yoshikawa, M., Iwasaki, H., Kinoshita, K., and Shinagawa, H. (2000) Two basic residues, Lys-107 and Lys-118, of RuvC resolvase are involved in critical contacts with the Holliday junction for its resolution. *Genes Cells* **5:** 803–813.

Yoshikawa, M., Iwasaki, H., and Shinagawa, H. (2001) Evidence that phenylalanine 69 in *Escherichia coli* RuvC resolvase forms a stacking interaction during binding and destabilization of a Holliday junction DNA substrate. *J Biol Chem* **276:** 10432–10436.

Functional similarities between phage λ Orf and *Escherichia coli* RecFOR in initiation of genetic exchange

Karen L. Maxwell*†, Patricia Reed‡‡, Rong-guang Zhang†§, Steven Beasley§, Adrian R. Walmsley†, Fiona A. Curtis†, Andrej Joachimiak§¶, Aled M. Edwards*, and Gary J. Sharples†||

†Centre for Infectious Diseases, Wolfson Research Institute, University of Durham, Queen's Campus, Stockton-on-Tees TS17 6BH, United Kingdom; ‡Ontario Cancer Institute and Department of Medical Biophysics, University of Toronto, Toronto, ON, Canada M5G 2M9; §Midwest Center for Structural Genomics and Structural Biology, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, IL 60440; and ¶Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, IL 60637

Edited by Charles M. Radding, Yale University School of Medicine, New Haven, CT, and approved June 22, 2005 (received for review April 25, 2005)

Genetic recombination in bacteriophage λ relies on DNA end processing by Exo to expose 3′-tailed strands for annealing and exchange by β protein. Phage λ encodes an additional recombinase, Orf, which participates in the early stages of recombination by supplying a function equivalent to the *Escherichia coli* RecFOR complex. These host enzymes assist loading of the RecA strand exchange protein onto ssDNA coated with ssDNA-binding protein. In this study, we purified the Orf protein, analyzed its biochemical properties, and determined its crystal structure at 2.5 Å. The homodimeric Orf protein is arranged as a toroid with a shallow U-shaped cleft, lined with basic residues, running perpendicular to the central cavity. Orf binds DNA, favoring single-stranded over duplex and with no obvious preference for gapped, 3′-tailed, or 5′-tailed substrates. An interaction between Orf and ssDNA-binding protein was indicated by far Western analysis. The functional similarities between Orf and RecFOR are discussed in relation to the early steps of recombinational exchange and the interplay between phage and bacterial recombinases.

bacteriophage | DNA repair | genetic recombination | NinB

Genetic recombination in bacteriophages preserves genomic integrity by repairing strand breakages. However, exchanges occasionally occur at inappropriate sites, leading to rearrangement of existing genes or acquisition of new ones. Thus, considerable diversity is generated amongst phage populations, and this has a major impact on bacterial pathogen evolution by facilitating dissemination of virulence genes (1). The mechanism of phage recombination has been studied in some detail by using phage λ as a model system, partly as a consequence of its exploitation for *in vivo* genetic engineering (2). Two pathways of exchange predominate in λ depending on whether a DNA strand is used to invade a homologous duplex or is annealed to a complementary single-strand. Both envisage the restoration of a genomic dsDNA break by exchange with a second λ chromosome (3). The invasion reaction is typical of models for *Escherichia coli* recombination at a break and requires host RecA to bind ssDNA, locate a homologous duplex, and promote strand exchange to create a recombinant joint (4). The second pathway functions independently of RecA and involves annealing of homologous ssDNA partner sequences. This reaction requires the breaks to be located at different sites in separate λ genomes and depends on the annealing properties of phage β protein (3, 5). Recombination in both reaction pathways is initiated by the coupled action of phage Exo and β proteins, collectively termed the Red system (2, 5). Exo is a 26-kDa exonuclease, degrading ssDNA in the 5′ to 3′ direction from a duplex DNA end to produce 3′ overhangs (6). The 30-kDa β protein can generate recombinants by annealing the 3′-tailed product generated by Exo to complementary ssDNA sequences (3, 7, 8). Recent experiments suggest that β can also perform strand invasion reactions similar to those mediated by RecA (9).

In addition to Exo and β, λ encodes a third protein, Orf (NinB), which is influential in the initial phase of genetic exchange (10–12). The 17-kDa Orf substitutes for a complex of three *E. coli* proteins (RecF, RecO, and RecR) in a *red* mutant crosses but not during host conjugational exchange (10, 11). However, Orf can replace *recFOR* function in *E. coli* recombination when Exo and β are present (12). Mutations in any of the *recFOR* genes normally confer modest deficiencies in DNA repair and recombination, phenotypes that can be partially suppressed by RecA mutations improving its ability to nucleate on ssDNA and displace ssDNA-binding (SSB) protein (13, 14). Individually and corporately, RecFOR exhibit a remarkable array of *in vitro* activities. RecF binds ssDNA or dsDNA and possesses a weak ATPase activity important in dissociation from dsDNA (15–17). RecO binds DNA and promotes strand invasion and annealing of homologous sequences, especially those already bound by SSB (18–20). The monomeric RecO is composed of three domains: an N-terminal oligonucleotide/oligosaccharide binding fold, a central α-helical bundle, and a C-terminal zinc-binding motif (21). RecR proteins from *Bacillus subtilis* and *Deinococcus radiodurans* bind DNA (22, 23), although *E. coli* RecR apparently does not (16, 24). Four RecR subunits are arranged as a toroid with a central hole of 30–35 Å in diameter, which probably accommodates dsDNA (23). There is evidence for RecF·O, RecFR, RecOR, and RecFOR interactions and an association between RecO and SSB (24–26). RecFR limits extension of RecA on duplex DNA after initial filament assembly on ssDNA, whereas RecOR helps RecA gain access to ssDNA blocked by the presence of SSB protein (27). All three together are required to load RecA at ssDNA–dsDNA junctions containing a 5′ end when SSB coats ssDNA at gaps (28). Finally, RecFOR appears to safeguard the nascent strands at stalled replication forks (29) and is toxic in certain mutant backgrounds where DNA replication is impaired, suggesting that inappropriate loading of RecA at replication forks is detrimental to survival (30, 31).

To elucidate how a single phage protein, Orf, can substitute for three much larger host polypeptides, we purified λ Orf, determined its crystal structure, and characterized its interactions with DNA substrates designed to mimic the early inter-

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: SSB, ssDNA-binding protein; MBP, maltose-binding protein.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 1PC6).

‡K.L.M., P.R., and R.-g.Z. contributed equally to this work.

||To whom correspondence should be addressed. E-mail: gary.sharples@durham.ac.uk.

www.pnas.org/cgi/doi/10.1073/pnas.0503399102

mediates of genetic recombination. A physical interaction between Orf and SSB suggests that the λ protein fulfils an important role in helping recombinases gain access to the DNA template, a function conserved throughout biology.

Materials and Methods

Purification of Orf Protein for DNA-Binding Analysis. The *orf* (*ninB*) gene was amplified by PCR from λ genomic DNA and inserted into pET14b (pPR100) or pET15b (pKM123) to generate an N-terminal His_6 affinity tag fusion containing a thrombin cleavage site. The integrity of the cloned *orf* gene from both constructs was confirmed by DNA sequencing. Orf was overexpressed from 2 liters of *E. coli* BL21 pLysS carrying pPR100 and purified on nickel-iminodiacetic acid Sepharose and heparin agarose columns (Sigma). Orf was stored in 20 mM Tris·HCl, pH 8.0/1 mM EDTA/0.5 mM DTT/500 mM KCl/50% glycerol at −80°C. A total of 2.6 mg of purified Orf was recovered at 1.3 mg/ml. Protein concentrations were determined by a modified Bradford Assay (Bio-Rad) using BSA as a standard; amounts of protein are expressed as moles of dimeric protein.

Orf Sample Preparation for Crystallography. *E. coli* BL21-Gold cells carrying pKM123 were grown at 37°C in TB medium supplemented with 4% glucose to an A_{600} of 1.0. Protein expression was induced with 1 mM IPTG, followed by 4 h incubation at 37°C. Orf was purified and prepared for crystallization studies as described (32). As a final purification step, gel filtration was performed by FPLC on a 25-ml Superdex-75 column (Amersham Pharmacia). Fractions containing pure protein as assayed by SDS/PAGE were concentrated by ultrafiltration. Selenomethionine-enriched Orf was expressed in the methionine auxotroph *E. coli* B834(DE3) in supplemented M9 medium and was purified as the native protein, except 5 mM 2-mercaptoethanol was added to the purification buffers.

Crystallization of Orf. Orf crystals were obtained by vapor diffusion in hanging drops (3 μl of protein to 3 μl of precipitant) over 1.4 M sodium acetate/0.05 M sodium cacodylate, pH 6.9, for 2–4 days at 21°C. For diffraction studies, the crystals were flash-frozen with the crystallization buffer plus 25% ethylene glycol. The crystals formed hexagonal rods reaching dimensions of 500 × 150 × 150 μm and belong to trigonal space group $P3_121$, with unit cell dimensions $a = b = 76.78$ Å, and $c = 107.39$ Å.

Crystal Structure Determination. Diffraction data were collected at 100K at the 19ID beamline of the Advanced Photon Source (Argonne National Laboratory) as described (33). The three-wavelength anomalous diffraction (MAD) data (peak, 0.9795 Å; inflection point, 0.9797 Å; high remote, 0.94656 Å) were collected to 2.5 Å from the selenomethionine-substituted protein crystal by using an inverse-beam strategy. All data were processed and scaled by using HKL2000 (34). Patterson searches, MAD phasing, density modification, initial map calculation, and structure refinement were carried out by using the CNS suite (35). The initial model was built automatically by using ARP/WARP (36), and consists of 75% main chain and 45% side chains. The model was rebuilt manually by using QUANTA and improved through several cycles of refinement and model building. The final model was refined to 2.50 Å by using CNS against the peak data. The final *R*-factor was 0.234, and the R_{free} was 0.294.

DNA Substrates. Oligonucleotides used are listed in Table 1, which is published as supporting information on the PNAS web site. For band shift assays, one oligonucleotide was labeled with [γ-^{32}P]ATP at its 5′ end by using T4 polynucleotide kinase.

DNA-Binding Assays. Band shift assays using ^{32}P-labeled DNA substrates were performed in 50 mM Tris·HCl, pH 8.0/5 mM EDTA/1 mM DTT/5% glycerol/100 μg/ml BSA. Samples were incubated on ice for 15 min before separation on 4% PAGE in 6.7 mM Tris·HCl, pH 8.0/3.3 mM sodium acetate/2 mM EDTA. Gels were dried onto filter paper and analyzed by autoradiography and PhosphorImaging. The interaction between Orf and DNA was time-resolved on an Applied Photophysics SX.18MV stopped-flow instrument. For measurements of the change in tryptophan fluorescence, samples were excited with light at 285 nm, selected with a monochromator, and emission monitored at wavelengths >335 nm by using a cutoff filter. Routinely, equal volumes of the reactants were mixed together in the stopped-flow instrument. To set up the instrument, 1.25 μM Orf protein was mixed with 50 mM Tris·HCl (pH 8.0), and the photomultiplier tube voltage was increased to 4 V and then backed-off to zero, and changes about this zero baseline recorded.

Far Western Blotting. SSB protein was purified as described (37) apart from omission of the gel-filtration step. An N-terminal maltose-binding protein (MBP)–Orf fusion expressed from pMALc2 was purified on amylose resin. SSB was separated on SDS/15% PAGE and transferred to poly(vinylidene difluoride) membrane by electroblotting in 10 mM cyclohexylaminopropanesulfonic acid/20% methanol. A prestained protein molecular weight standard served as a size marker. Blots were probed with MBP–Orf protein, and interactions were detected with monoclonal anti-MBP antibodies and mouse IgG peroxidase conjugate (Sigma). Chemiluminescence was observed by exposure to x-ray film after treatment with enhanced chemiluminescence reagents (Amersham Pharmacia).

Results and Discussion

Interaction of Orf with DNA. The λ *orf* gene was cloned into pET expression vectors and purified as an N–His_6 fusion by nickel-affinity chromatography followed by further fractionation on heparin agarose. Because all three components of the RecFOR complex show some association with DNA, we investigated the ability of Orf to bind different DNA substrates in gel mobility-shift assays. Orf formed at least two discrete complexes with a 50-nt DNA substrate, but bound much less well to a duplex of the same sequence (Fig. 1*A*). The obvious band smearing with the duplex indicates that Orf forms an unstable association with dsDNA (Fig. 1*A*, lane h). A clear preference for binding ssDNA was evident when a larger range of protein concentration was used (Fig. 1*B*). Samples of Orf protein lacking the N-terminal tag gave very similar DNA binding profiles on these two substrates (data not shown).

To further probe the interaction of Orf with DNA, we used stopped-flow fluorescence spectroscopy. This approach can be used to evaluate interactions between proteins and their substrates assuming an appropriately situated tryptophan residue is available and is sufficiently close to the ligand binding site. Orf contains seven tryptophans, and the fluorescence of one or more of these is quenched when protein and DNA are mixed on a stopped-flow device (Fig. 2*A*). The fluorescence decrease is most likely a consequence of DNA blocking the exposed tryptophan(s) or a conformational change in protein architecture resulting in burying of these residues. The rate and differences in amplitude of the fluorescence quench show that Orf binds to ssDNA (either 25 or 50 nt) in preference to dsDNA of the same length and sequence (Fig. 2*A*). Binding of Orf to ssDNA appears to occur in two phases with a rapid initial quench in fluorescence followed by a slower change between 10 and 30 s. Very similar binding profiles were obtained by using a 50-base oligonucleotide containing a 25-bp duplex section at either the 5′ or 3′ end, or in the center of the molecule (Fig. 2*B*, see Fig. 5*B*). Thus, Orf does not appear to favor loading at the intersection between single- and double-stranded DNA as noted with RecF (38) and RecFOR (28).

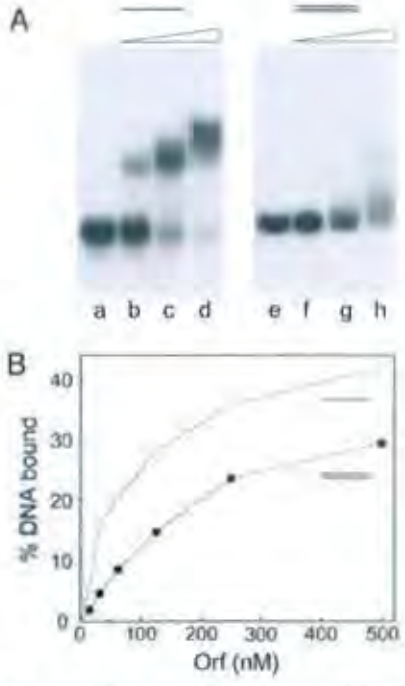

Fig. 1. Orf binding to DNA in gel mobility shift assays. (*A*) Samples contained 0.3 nM ssDNA (lanes a–d) or dsDNA (lanes e–h) with 0, 5, 50, and 500 nM His-Orf protein. (*B*) Binding reactions contained 0.3 nM ssDNA and dsDNA as indicated with 0, 15.625, 31.25, 62.5, 125, 250, and 500 nM His-Orf. Data are the mean of two independent experiments.

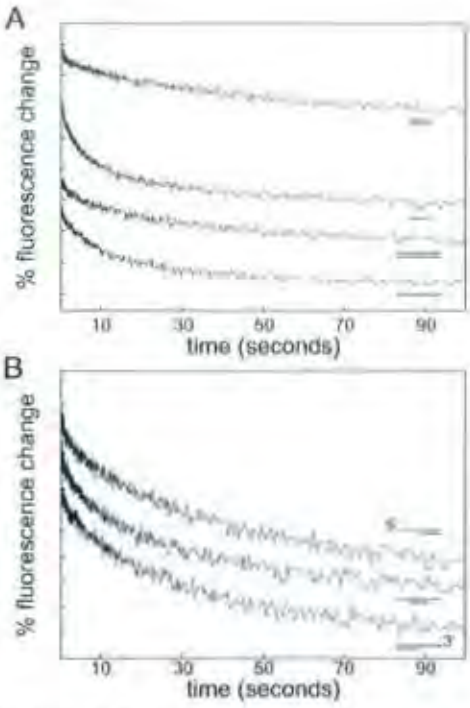

Fig. 2. Orf binding to DNA in stopped-flow assays. Samples were mixed on the stopped-flow instrument to give final concentrations of 625 nM His-Orf and 5 μM DNA as indicated. (*A*) Binding to 25- and 50-mer ssDNA and dsDNA substrates. (*B*) Binding to 50-mer 5′ and 3′ overhangs and tailed duplex DNA substrates. Each division on the *y* axis represents a fluorescence change of 5% (*A*) and 2% (*B*).

One of the components of the RecFOR complex, RecO, has the capacity to anneal homologous single-stranded DNA molecules (18, 20), and so we assayed Orf for strand annealing activity. Using polyacrylamide gels to monitor duplex formation of ^{32}P-labeled DNA, we found no evidence that Orf could anneal either homologous oligonucleotides (50 nt) or those (25 and 50 nt) matching ϕX174 circular ssDNA. Similarly, Orf failed to stimulate strand annealing in stopped-flow assays using DAPI to quantify formation of dsDNA from oligonucleotide pairs (data not shown). Because λ β is a single-strand annealing protein (7, 8), it is perhaps unsurprising that Orf lacks this activity. Nevertheless, the ssDNA-binding properties of Orf clearly correlate with its predicted role in initiation of exchange at gaps or ends where β and RecA compete with SSB for access.

Crystal Structure of Orf Protein. To help ascertain how Orf interacts with DNA, we determined its atomic structure by x-ray crystallography using multiwavelength anomalous diffraction data from a single crystal of selenomethionine-substituted protein. Statistics describing the quality of x-ray diffraction data are summarized in Table 2, which is published as supporting information on the PNAS web site. There are two 146-residue chains in the asymmetric unit; the final model includes residues 1–141 for chain A, and 1–133 for chain B. The Ramachandran plot produced by PROCHECK (39) shows that, of 274 residues in the model, 88.1% are in the most favored regions, 11.1% are in additional allowed regions, and no residues fall in the disallowed regions. Statistics related to the final refined protein model are shown in Table 3, which is published as supporting information on the PNAS web site.

The 2.51-Å resolution structure reveals that Orf belongs to the $\alpha + \beta$ protein class, and possesses a unique fold as determined by the DALI algorithm (40). Two Orf monomers are arranged as an interlwined dimer, forming a ring with a funnel-like channel through the center (Fig. 3 *A* and *B*). The Orf structure can be loosely divided into two components: a larger C-terminal region, composed of two three-helix bundles and two antiparallel β-sheets, and a smaller N-terminal region comprising a single antiparallel β-sheet topped by two α-helices. Each Orf monomer consists of two separate regions that interact to form the oligomeric structure (Fig. 3*C*). The first region comprises the initial 38 residues of the protein and includes three segments that contribute to the dimer interface: residues 3–9 (β1), 14–22 (αA), and 26–38 (β2 loop and β2). The β-strands form a single intertwined β-sheet, and the two α-helices pack across the top of the β-sheet in a parallel manner. The second region comprises residues 67–97, and also has three segments contributing to the interface: residues 67–80 (αC), 83–89 (β3), and 95–97 (β4). In this region, the α-helix of one monomer packs on top of an antiparallel β-sheet formed by the two strands from the other monomer. Nine hydrophobic residues from each monomer (Ile-8, Val-21, Val-33, Ile-36, Val-75, Val-86, Leu-89, Val-95, and Ile-97) are highly buried in the dimer interface. With the exception of Val-33, the hydrophobic character of each of these residues is highly conserved between Orf homologs (Fig. 3*D*). There are also six salt bridges and numerous ionic interactions mediated through water molecules that contribute to the stability and specificity of the dimer interface. The size of the oligomeric interface (1,407 Å^2 buried per subunit) is close to the average of 1,700 ± 1,100 Å^2 observed for homodimers (41), and accounts for ≈20% of the accessible surface of each subunit. To investigate the role of the dimer conformation in the crystal structure, size exclusion chromatography and cross-linking experiments were performed. The homogenous Orf protein preparation formed an apparent dimer of 33 kDa on size exclusion chromatography and when exposed to the cross-linking agent glutaral-

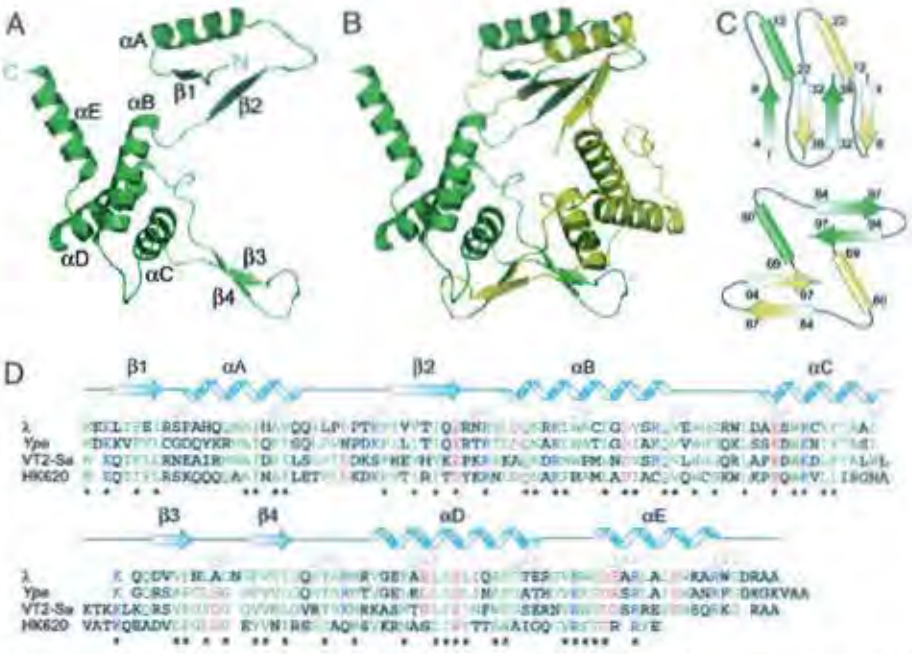

Fig. 3. Primary, secondary, and tertiary structure of Orf. (*A*) Ribbon diagram of an Orf subunit with secondary structure elements labeled. (*B*) Overall view of the dimeric structure of Orf with monomer A shown in green, and monomer B in yellow. (*C*) Schematic representation of the N- and C-terminal dimerization regions of Orf. The strands and helices are colored as in *B*. (*D*) Sequence alignment of selected phage Orf homologs. Conserved residues are highlighted: acidic (red), basic (blue), and others (green). Asterisks denote residues conserved in all four sequences. The representative homologs come from a *Yersinia pestis* prophage (*Ype*), *E. coli* O157:H7 phage VT2-Sa, and *E. coli* H phage HK620. Secondary structure elements are indicated above the residue numbers given for the λ Orf sequence.

dehyde (Fig. 7, which is published as supporting information on the PNAS web site).

The Orf dimer exhibits asymmetry due to a twist in the backbone at residues Asn-40–Ser-42. When the two monomers are examined, residues Lys-3–Glu-38 can be aligned with an rms deviation (rmsd) of 0.78 Å over 144 backbone atoms, and residues Ser-42–Ser-128 can be aligned with an rmsd of 0.65 Å over 352 backbone atoms. The hinge point involves residues Asn-40 and Arg-41, which have ϕ/ψ angles of $-104/-45.5$ and $-117/124$, respectively, in monomer A and $-47/150$ and $-43/-57$, respectively, in monomer B. Additional asymmetry is provided by the C-terminal tail, which extends away from the protein out into solution. After the fourth α-helix (residues 106–121), monomer A has an α-helix comprised of residues 128–140, whereas monomer B has random coil until residue 133, when the electron density is missing, indicating that this region is disordered. Examination of the crystal data shows that the C terminus of monomer A packs against three neighboring proteins in the crystal, stabilizing αE. The extended conformation of the helix in monomer A and disorder observed at the C terminus of monomer B suggests that the last 20 residues are flexible and could adopt different conformations in Orf complexed to DNA or other proteins. There are several reports in the literature of asymmetry in homodimer–nucleic acid interactions, including nonstructural protein B (NSP3) from Rotavirus, which forms a heart-shaped, asymmetric homodimer when bound to RNA, creating a single RNA-binding site (42). The HIV-1 reverse transcriptase also forms an asymmetric dimer, creating only one RNA-dependent DNA polymerase active site, one RNase H active site, one tRNA-binding site, and one noncompetitive inhibitor-binding site (43). A similar situation may arise in the Orf dimer, allowing it to interact with DNA as well as providing binding sites for other recombinases.

Calculation of the molecular surface of the protein performed by the program SPOCK (http://quorum.tamu.edu) reveals a central channel that traverses the dimer (Fig. 4*A*). Electrostatic calculations show that the interior of this channel is very positively charged (Fig. 4*B*), due to several highly conserved residues (Lys3A/B, Arg41A, Lys73A/B, and Lys81A; Fig. 3*D*). The channel ranges in diameter from ≈20 Å at the top to only 8 Å near the bottom, and could potentially accommodate the passage of ssDNA like a bead on a string. The distal end of the cavity is partially occluded by the first two residues of monomer B. However, the presence of either a polyhistidine or maltose-binding protein tag at the N terminus had no significant impact on Orf DNA binding activity (data not shown), suggesting that either DNA does not pass through this channel, or that the N terminus of monomer B is able to change conformation to accommodate it. Alternatively, binding to ssDNA could occur within a shallow U-shaped cleft that traverses the top of the central opening and is lined with a number of positively charged residues (Fig. 4*C*). In addition, the aromatic residue Trp50A is also present in the putative DNA binding region, and could act to stabilize the ssDNA–protein complex by stacking against the nucleotide bases.

Binding to ssDNA Across the Surface Cleft. To determine whether Orf threads ssDNA through the central cavity or across the surface cleft, we examined binding to a gapped duplex substrate based on the premise that dsDNA would be unable to penetrate the narrow aperture. We used a 26-nt stretch of ssDNA flanked at each end by 17-bp duplexes. Orf readily binds 25-mer oligonucleotides (Fig. 2*B*), so this gapped region should be sufficiently large to accommodate at least one Orf dimer. Control tailed substrates were made by simply omitting either of the short strands making up the duplex sections. In band shift assays with ^{32}P-radiolabeled DNA, Orf bound to all three of the substrates with similar affinities (Fig. 5*A*), indicating that ssDNA is ac-

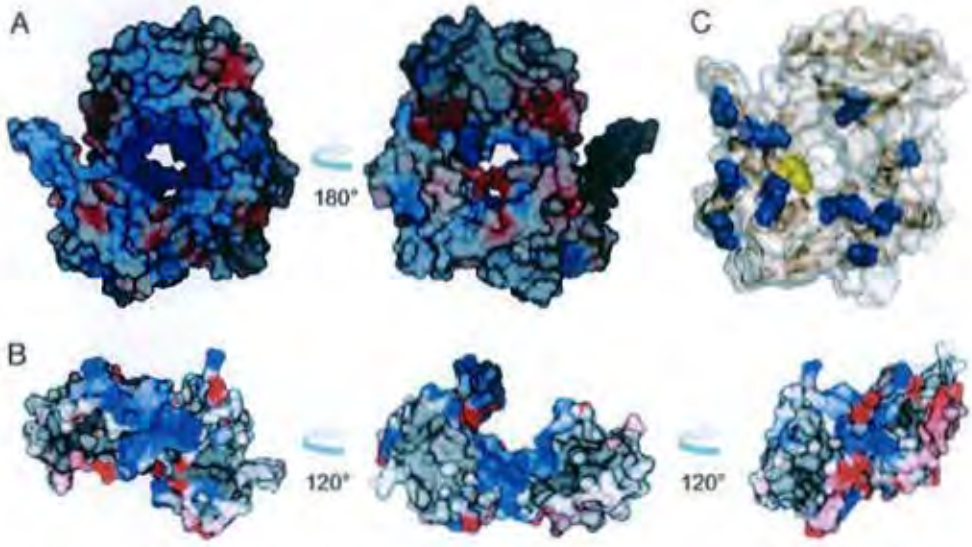

Fig. 4. Potential Orf DNA-binding sites. (*A*) Electrostatic surface representation of chemical properties of the solvent accessible surface of the Orf dimer calculated by using a probe radius of 1.4 Å. The surface is colored red and blue, representing electrostatic potentials less than −20 *kBT*, or greater than +20 *kBT*, where *B* is the Boltzman coefficient and *T* is temperature. A view of the potential DNA binding site looking through the central channel from the top (*Left*) and bottom (*Right*) of the protein is shown. (*B*) This view is rotated by 90° about the *x* axis from *A*. The front half of the protein is removed to show the contours and electrostatic charge present in the interior of the channel that runs the length of the protein. Each of the three representations are rotated about the *y* axis by 120°. (*C*) Molecular surface representation illustrating the potential DNA binding cleft that traverses the top of the Orf dimer. Basic residues are shown in blue, Trp50A is shown in yellow.

commodated within the surface cleft. Binding to the gapped duplex was also monitored by stopped-flow spectroscopy (Fig. 5*B*). Very similar fluorescence change profiles were detected by using the gapped duplex, ssDNA, or 5′ and 3′ overhangs (Fig. 5*B*), implying that Orf binds each of these substrates equally well. At this concentration of DNA, very little decrease in protein fluorescence was detected with dsDNA. Taken together, the data suggest that ssDNA is accommodated within the U-shaped cleft of Orf rather than passing through the central channel of the ring (Fig. 4). This finding fits with the ability of Orf to associate with both ssDNA and dsDNA, albeit weakly with the latter.

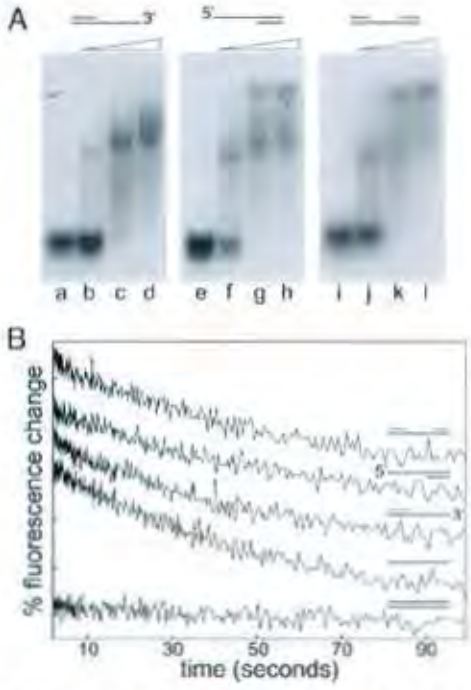

Fig. 5. Defining the Orf DNA-binding site. (*A*) Gel-shift assays of binding to 0.3 nM 60-mer 3′ overhang (lanes a–d), 5′ overhang (lanes e–h), and gapped duplex (lanes i–l) with His-Orf protein at 0, 5, 50, and 500 nM. (*B*) Stopped-flow assays of His-Orf (625 nM) binding to DNA (0.5 μM). Each division on the *y* axis represents a fluorescence change of 5%.

Orf and SSB Interact. RecFOR promote assembly of RecA filaments on SSB-coated DNA (28). Therefore, we investigated

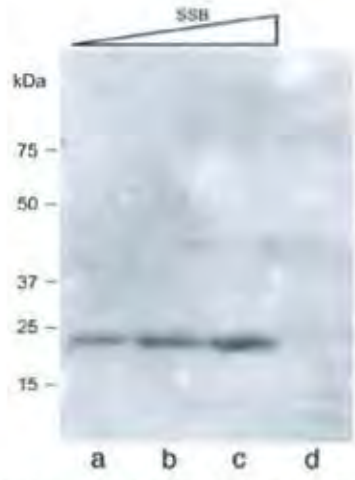

Fig. 6. Interaction between Orf and SSB. Far Western analysis on purified SSB (lanes a–c) at 0.5, 1.9, and 3.8 μg separated on SDS/15% PAGE. Lane d contained no protein. Blotted SSB protein was probed with 20 μg of MBP-Orf and interactions detected with anti-MBP antibodies.

whether Orf could interact directly with SSB by using far Western blotting. Wild-type SSB protein on a poly(vinylidene difluoride) membrane was probed with purified MBP–Orf and antibodies directed against the tag. The 19-kDa SSB protein was detected in purified (Fig. 6) and overexpressed samples, indicating that Orf and SSB interact. Similar results were obtained by using MBP–Orf on gels with purified His–SSB as a probe with anti-His antibodies. No signal was detected in the absence of MBP–Orf protein or using MBP alone (data not shown).

Potential Function of Orf in Initiation of Genetic Exchange. The *in vitro* ssDNA and SSB binding properties of Orf clearly fit with a role in the initial steps of genetic exchange as predicted by its ability to substitute for RecFOR in recombination reactions *in vivo* (10, 12). As with the RecFOR complex, Orf could aid RecA nucleation on ssDNA prebound by SSB (12), although why a phage system should mimic so closely that of its host is unclear. It is possible, because Orf shows no apparent preference for the 5′ ended structure recognized by RecFOR (28), that loading RecA onto any available ssDNA has specific advantages for phage recombination. Alternatively, Orf could encourage loading of β protein to overcome the inhibitory presence of SSB; this may explain why Orf seems to function better in recombination reactions mediated by the Red system (12) rather than those supported by its host (11). It has always been assumed that Exo aids loading of β onto ssDNA during end degradation, because the two recombinases are known to form a complex (2). Orf may be needed to help β gain access to ssDNA in other situations, for example those exchanges occurring at gaps or replication forks (44). The fluorescence quench observed when Orf associates with DNA could indicate a conformational change that helps either displace SSB or allows access of RecA or β protein to ssDNA. This change could occur by reinstating symmetry to the Orf dimer or alterations in the order/disorder of helix E. Further work is needed to establish whether bacterial or phage recombination pathways are stimulated by Orf. Many of the structurally distinct systems for overcoming the barrier of ssDNA binding protein show species-specific interactions with the recombination enzymes they load (28, 45). The known bacterial and eukaryotic mediators are annealing proteins in their own right, so phage λ Orf and its unrelated analogue from T4 (UvsY) may be unusual amongst this group in functioning primarily in a facilitatory role.

We thank Peter McGlynn (University of Aberdeen, Aberdeen, U.K.) for the SSB overexpression construct. This work was supported by the Biotechnology and Biological Sciences Research Council, the Canadian Institutes of Health Research (CIHR), National Institutes of Health Grant GM62414 (to A.J.), the Ontario Research and Development Challenge Fund, and the United States Department of Energy, Office of Biological and Environmental Research under contract W-31-109-ENG-38. K.L.M. was supported by a CIHR Postdoctoral Fellowship.

1. Brüssow, H., Canchaya, C. & Hardt, W. D. (2004) *Microbiol. Mol. Biol. Rev.* **68,** 560–602.
2. Court, D. L., Sawitzke, J. A. & Thomason, L. C. (2002) *Annu. Rev. Genet.* **36,** 361–388.
3. Stahl, M. M., Thomason, L., Poteete, A. R., Tarkowski, T., Kuzminov, A. & Stahl, F. W. (1997) *Genetics* **147,** 961–977.
4. Kowalczykowski, S. C. (2000) *Trends Biochem. Sci.* **25,** 156–165.
5. Poteete, A. R. (2001) *FEMS Microbiol. Lett.* **201,** 9–14.
6. Little, J. W. (1967) *J. Biol. Chem.* **242,** 679–686.
7. Kmiec, E. & Holloman, W. K. (1981) *J. Biol. Chem.* **256,** 12636–12639.
8. Muniyappa, K. & Radding, C. M. (1986) *J. Biol. Chem.* **261,** 7472–7478.
9. Rybalchenko, N., Golub, E. I., Bi, B. & Radding, C. M. (2004) *Proc. Natl. Acad. Sci. USA* **101,** 17056–17060.
10. Sawitzke, J.-A. & Stahl, F. W. (1992) *Genetics* **130,** 7–16.
11. Sawitzke, J. A. & Stahl, F. W. (1994) *J. Bacteriol.* **176,** 6730–6737.
12. Poteete, A. R. (2004) *J. Bacteriol.* **186,** 2699–2707.
13. Madiraju, M. V., Lavery, P. E., Kowalczykowski, S. C. & Clark, A. J. (1992) *Biochemistry* **31,** 10529–10535.
14. Wang, T. V., Chang, H. & Hung, J. (1995) *Mutat. Res.* **294,** 157–166.
15. Griffin, T. J. & Kolodner, R. D. (1990) *J. Bacteriol.* **172,** 6291–6299.
16. Webb, B. L., Cox, M. M. & Inman, R. B. (1995) *J. Biol. Chem.* **270,** 31397–31404.
17. Webb, B. L., Cox, M. M. & Inman, R. B. (1999) *J. Biol. Chem.* **274,** 15367–15374.
18. Luisi-DeLuca, C. & Kolodner, R. (1994) *J. Mol. Biol.* **236,** 124–138.
19. Luisi-DeLuca, C. (1995) *J. Bacteriol.* **177,** 566–572.
20. Kantake, N., Madiraju, M. V., Sugiyama, T. & Kowalczykowski, S. C. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 15327–15332.
21. Leiros, I., Timmins, J., Hall, D. R. & McSweeney, S. (2005) *EMBO J.* **24,** 906–918.
22. Alonso, J. C., Stiege, A. C., Dobrinski, B. & Lurz, R. (1993) *J. Biol. Chem.* **268,** 1424–1429.
23. Lee, B. I., Kim, K. H., Park, S. J., Eom, S. H., Song, H. K. & Suh, S. W. (2004) *EMBO J.* **23,** 2029–2038.
24. Umezu, K. & Kolodner, R. D. (1994) *J. Biol. Chem.* **269,** 30005–30013.
25. Umezu, K., Chi, N. & Kolodner, R. D. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 3875–3879.
26. Hegde, S. P., Qin, M. H., Li, X. H., Atkinson, M. A., Clark, A. J., Rajagopalan, M. & Madiraju, M. V. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 14468–14473.
27. Webb, B. L., Cox, M. M. & Inman, R. B. (1997) *Cell* **91,** 347–356.
28. Morimatsu, K. & Kowalczykowski, S. C. (2003) *Mol. Cell.* **11,** 1337–1347.
29. Chow, K. H. & Courcelle, J. (2004) *J. Biol. Chem.* **279,** 3492–3496.
30. Petit, M. A. & Ehrlich, D. (2002) *EMBO J.* **21,** 3137–3147.
31. Moore, T., McGlynn, P., Ngo, H. P., Sharples, G. J. & Lloyd, R. G. (2003) *EMBO J.* **22,** 735–745.
32. Zhang, R., Andersson, C. E., Savchenko, A., Skarina, T., Evdokimova, E., Beasley, S., Arrowsmith, C. H., Edwards, A. M., Joachimiak, A. & Mowbray, S. L. (2003) *Structure (Cambridge, U.K.)* **11,** 31–42.
33. Walsh, M. A., Evans, G., Sanishvili, R., Dementieva, I. & Joachimiak, A. (1999) *Acta Crystallogr. D* **55,** 1726–1732.
34. Otwinowski, Z. & Minor, W. (1997) *Methods Enzymol.* **276,** 307–326.
35. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., *et al.* (1998) *Acta Crystallogr. D* **54,** 905–921.
36. Perrakis, A., Sixma, T. K., Wilson, K. S. & Lamzin, V. S. (1997) *Acta Crystallogr. D* **53,** 448–455.
37. Cadman, C. J. & McGlynn, P. (2004) *Nucleic Acids Res.* **32,** 6378–6387.
38. Hegde, S. P., Rajagopalan, M. & Madiraju, M. V. (1996) *J. Bacteriol.* **178,** 184–190.
39. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26,** 283–291.
40. Holm, L. & Sander, C. (1995) *Trends Biochem. Sci.* **20,** 478–480.
41. Jones, S. & Thornton, J. M. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 13–20.
42. Deo, R. C., Groft, C. M., Rajashankar, K. R. & Burley, S. K. (2002) *Cell* **108,** 71–81.
43. Wang, J., Smerdon, S. J., Jager, J., Kohlstaedt, L. A., Rice, P. A., Friedman, J. M. & Steitz, T. A. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 7242–7246.
44. Ellis, H. M., Yu, D., DiTizio, T. & Court, D. L. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 6742–6746.
45. Beernink, H. T. & Morrical, S. W. (1999) *Trends Biochem. Sci.* **24,** 385–389.

DOI: 10.1534/genetics.105.045906

The RuvAB Branch Migration Translocase and RecU Holliday Junction Resolvase Are Required for Double-Stranded DNA Break Repair in *Bacillus subtilis*

Humberto Sanchez,* Dawit Kidane,†,1 Patricia Reed,‡ Fiona A. Curtis,‡ M. Castillo Cozar,* Peter L. Graumann,†,1 Gary J. Sharples‡ and Juan C. Alonso*,2

*Department of Microbial Biotechnology, Centro Nacional de Biotecnología, Consejo Superior de Investigaciones Científicas, 28049 Madrid, Spain, †Biochemie, Fachbereich Chemie, Philipps-Universität Marburg, 35032 Marburg, Germany and ‡Centre for Infectious Diseases, Wolfson Research Institute, University of Durham, Stockton-on-Tees TS17 6BH, United Kingdom

Manuscript received May 20, 2005
Accepted for publication July 13, 2005

ABSTRACT

In models of *Escherichia coli* recombination and DNA repair, the RuvABC complex directs the branch migration and resolution of Holliday junction DNA. To probe the validity of the *E. coli* paradigm, we examined the impact of mutations in Δ*ruvAB* and Δ*recU* (a *ruvC* functional analog) on DNA repair. Under standard transformation conditions we failed to construct Δ*ruvAB* Δ*recG*, Δ*recU* Δ*ruvAB*, Δ*recU* Δ*recG*, or Δ*recU* Δ*recJ* strains. However, Δ*ruvAB* could be combined with *addAB* (*recBCD*), *recF*, *recH*, Δ*recS*, Δ*recQ*, and Δ*recJ* mutations. The Δ*ruvAB* and Δ*recU* mutations rendered cells extremely sensitive to DNA-damaging agents, although less sensitive than a Δ*recA* strain. When damaged cells were analyzed, we found that RecU was recruited to defined double-stranded DNA breaks (DSBs) and colocalized with RecN. RecU localized to these centers at a later time point during DSB repair, and formation was dependent on RuvAB. In addition, expression of RecU in an *E. coli ruvC* mutant restored full resistance to UV light only when the *ruvAB* genes were present. The results demonstrate that, as with *E. coli* RuvABC, RuvAB targets RecU to recombination intermediates and that all three proteins are required for repair of DSBs arising from lesions in chromosomal DNA.

IN all organisms, structural aberrations in the DNA template or strand breaks induce arrest or collapse of replication forks and their restoration relies on recombination functions (Haber 1999; Kuzminov 1999; Cox *et al.* 2000; Michel *et al.* 2004). In *Escherichia coli*, stalled forks can reverse to form a four-stranded Holliday junction (HJ) intermediate (Seigneur *et al.* 1998). Fork regression, which might also occur spontaneously, involves RecG or potentially RecA, the latter loaded onto single-stranded DNA (ssDNA) by the RecFOR complex (Robu *et al.* 2001, 2004; Singleton *et al.* 2001; McGlynn and Lloyd 2002a,b). Once formed, the HJ can be processed in a number of ways: (i) The extruded duplex end can be removed by either RecBCD or RecQ and RecJ to reset the fork; (ii) DNA synthesis on the extruded partial duplex end followed by restoration of the fork by RecG or RuvAB branch migration provides a means of translesion bypass; and (iii) branch migration away from a block or lesion and HJ resolution by RuvC generates a broken fork (as is the case when the replisome encounters a strand break). RecA then mediates invasion of this broken end into the intact chromosome arm to rebuild the replication fork (Haber 1999; Kuzminov 1999; Cox *et al.* 2000; Michel *et al.* 2004). Mechanisms for direct fork rescue, which do not invoke the formation of a HJ intermediate, have also been proposed and rely on the action of RecFOR, RecJ, and RecQ recombinases (Courcelle and Hanawalt 1999; Courcelle *et al.* 2001; Donaldson *et al.* 2004).

The models for recombination-dependent replication highlight the important role played by RecBCD, RecQ, RecJ, and RecFOR in processing the ends of collapsed forks and loading of RecA (Clark and Sandler 1994; Kowalczykowski and Eggleston 1994; Kuzminov 1999; Amundsen and Smith 2003; Michel *et al.* 2004). RecBCD preferentially degrades double-stranded DNA (dsDNA) ends to expose a 3′ single-stranded tail. Similar reactions can be catalyzed by unwinding the end using RecQ helicase coupled with strand removal by the RecJ 5′-3′ exonuclease (Courcelle *et al.* 2001; Amundsen and Smith 2003). RecA can be loaded directly onto this resected ssDNA by RecBCD (Anderson and Kowalczykowski 1997, 2000; Chedin and Kowalczykowski 2002; Amundsen and Smith 2003; Xu and Marians 2003) or by the RecFOR complex when the strand is coated with Single-stranded DNA binding (SSB) protein (Umezu and Kolodner 1994; Shan *et al.* 1997; Kantake *et al.* 2002; Ivancic-Bace *et al.* 2003). Formation of a RecA nucleoprotein filament allows homologous pairing and strand

[1]*Present address:* Institut für Mikrobiologie, Universität Freiburg, Verfügungsgebäude, Stefan-Meier-Strasse 17, 79104 Freiburg, Germany.

[2]*Corresponding author:* Departmento de Biotecnología Microbiana, Centro Nacional de Biotecnología, CSIC, Campus Universidad Autónoma de Madrid, C/Darwin 3, Cantoblanco, 28049 Madrid, Spain. E-mail: jcalonso@cnb.uam.es

exchange between the broken end and its undamaged partner. Invasion of the homologous duplex by the processed 3′-tail creates a D-loop upon which the replication apparatus can be reloaded by PriA (Kowalczykowski 2000; Marians 2000; McGlynn and Lloyd 2002a,b). At this stage the chromosomes are still interconnected so further extension of the DNA joint to form a HJ is needed so that RuvABC resolution can complete the fork restoration process.

In *Bacillus subtilis* the recombination genes, other than *recA*, which is central to all pathways of recombinational repair, have been placed into six different epistatic groups: α [comprising *recF*, *recL*, *recO*, and *recR* (*recFLOR*) and *recN*], β (*addA* and *addB*), γ (*recP* and *recH*), ϵ (*ruvA*, *ruvB*, *recD*, and *recU*), ζ (*recS*), and η (*recG*) (Alonso *et al.* 1993; Fernandez *et al.* 1998, 1999, 2000; Chedin *et al.* 2000; Ayora *et al.* 2004; Carrasco *et al.* 2004). Throughout this article, unless stated otherwise, the indicated genes and products refer to those of *B. subtilis* origin. The *recA*, *recF*, *recG*, *recJ*, *recN*, *recO*, *recQ*, *recR*, *ruvA*, and *ruvB* genes each have a homologous counterpart in *E. coli* with the same gene designation. In addition, the *addAB* and *recU* genes encode functional equivalents of *E. coli recBCD* ($recBCD_{Eco}$) and $ruvC_{Eco}$ genes, respectively (Fernandez *et al.* 2000; Ayora *et al.* 2004). However, several recombination genes (*recL*, *recD*, *recH*, and *recP*) have no known equivalent in *E. coli* and, along with *recS* (a RecQ-like helicase), remain uncharacterized (Fernandez *et al.* 1998). Hence products classified within the α, β, ϵ, and η groups have their *E. coli* counterparts in RecN-FOR, RecBCD, RuvABC, and RecG, respectively (Ayora *et al.* 2004; Carrasco *et al.* 2004; Kidane *et al.* 2004), while those within the γ and ζ epistatic groups have yet to be assigned a function in DNA repair and recombination (Fernandez *et al.* 2000). Additionally, genetic analysis has not been undertaken with RecJ and RecQ and so neither has been assigned to any of these groupings.

Many of these functions and pathways of recombination resemble those encountered in the *E. coli* system. In wild-type cells, the loading of RecA protein onto SSB-coated ssDNA (presynaptic step) relies on AddAB or RecN-RecFLOR proteins (Chedin *et al.* 2000; Kidane *et al.* 2004). Recently it has been shown that: (i) ~35% of the cells in a Δ*recA* and ~5% in a Δ*recU* mutant contain unrepaired DSBs under normal growth conditions (Kidane *et al.* 2004); (ii) the *ruvA* gene complements the defect of the *recB2* mutation classified within the ϵ epistatic group [hence *recB2* was renamed *ruvA2* (Ayora *et al.* 2004)]; (iii) purified RecU protein binds preferentially to three- and four-strand junctions and cleaves Holliday junction substrates to produce nicked duplexes (Ayora *et al.* 2004); and (iv) in the absence of the RuvAB or RecG branch-migration activities, RecU and the poorly characterized RecD bias exchange toward crossovers (CO) (Carrasco *et al.* 2004).

To shed light on the importance of HJ branch migration and resolution in *B. subtilis*, we constructed a *ruvAB* null mutant (Δ*ruvAB*) and analyzed its sensitivity to different DNA-damaging agents. The Δ*ruvAB*, Δ*recG*, and *recF* strains showed similar sensitivities to DNA damage, which were significantly increased when Δ*ruvAB* was combined with *add*AB, *recF*, *recH*, or Δ*recJ*. Previously it was shown that in the absence of RuvAB, RecU, or RecG a clear chromosomal segregation defect was observed (Carrasco *et al.* 2004). Under standard transformation conditions we failed to construct Δ*ruvAB* Δ*recG*, Δ*ruvAB* Δ*recU*, and Δ*recU* Δ*recG* double-mutant strains. Expression of RecU could replace the repair function of $RuvC_{Eco}$ in the heterologous *E. coli* system if the $RuvAB_{Eco}$ complex were present. Formation of RecU foci on nucleoids also required the presence of RuvAB. Our data support the notion that RuvAB works in concert with the junction-resolving enzyme RecU, in a similar manner to the $RuvABC_{Eco}$ resolvasome complex, and that RuvAB, RecD, and RecU play a vital role in DNA DSB repair.

MATERIALS AND METHODS

Bacterial strains and plasmids: All *B. subtilis* strains used in this study are listed in Table 1 and are isogenic to strain YB886 (*rec*$^+$ control). A 2-kb *six-cat-six* cassette containing two directly repeated copies of the β-site-specific recombinase target site (*six*) surrounding the chloramphenicol acetyl transferase gene (*cat*) (Carrasco *et al.* 2004) was introduced within the coding sequence of *recJ*, *recQ*, *recG*, and *ruvA ruvB* to generate the *recJ::six-cat-six*, *recQ::six-cat-six*, *recG::six-cat-six*, and *ruvAB::six-cat-six* disruptions. These disruptions were transferred into the chromosome of wild-type cells to generate Δ*recJ*, Δ*recQ*, Δ*recG*, and Δ*ruvAB* strains, and expression of the β-gene provoked deletion of the *cat* gene. The null Δ*recU* or Δ*ruvAB* mutation was transferred into the isogenic *rec*-deficient derivatives and the double mutants generated by a double CO event as previously described (Alonso *et al.* 1993).

Chromosomal DNA from Δ*ruvAB* (*ruvAB::six-cat-six*), Δ*recU* (*recU::six-spc-six*), or Δ*recS* (*recS::cat*) strains were used to transform the wild type and the mutants Δ*ruvAB*, Δ*recU*, Δ*recG*, or Δ*recJ* strains with selection for chloramphenicol (conferred by the *cat* gene) or spectinomicyn (conferred by the *spc* gene) resistance. The Δ*recS* mutation could be transferred into all transformed strains, showing that the mutant strains are competent for transformation. An equivalent amount of chromosomal DNA from Δ*ruvAB* or Δ*recU* mutants transforms wild-type cells with similar efficiency, but no *bona fide* transformants were obtained for the Δ*ruvAB*, Δ*recU*, Δ*recG*, or Δ*recJ* mutant strain. Few tiny colonies after 72 hr of incubation times were obtained, detailed analysis of several of these transformants suggested that some of them contained a single CO with one copy of the wild type and one of the mutant gene, and a few with a double CO contained suppressor mutations (see below).

E. coli K12 *ruv* mutants strains, SR2210 (*ruvA200*), N1057 (*ruvB4*), N2057 (*ruvAB60*::Tn*10*), GS1481 (Δ*ruvC*::*kan*), GS85 (*ruvC53 eda*::Tn*10*), AM888 (Δ*ruvAC65* Δ*ruvA*::*kan*), and N4454 (Δ*ruvABC*::*cat*) are derivatives of the *ruv*$^+$ wild-type strain, AB1157 (Sargentini and Smith 1989; Sharples *et al.* 1990; Mandal *et al.* 1993; Mahdi *et al.* 1996; Seigneur *et al.* 1998). Plasmid pCB564 was constructed by transferring the 1.2-kb *Bsp*EI-*Bam*HI DNA segment containing the *recU* gene into pHP13 (pRecU). pCB593 contains the 4.9-kb *Stu*I (*ruvA*) fragment inserted into pUC18 (pRuvA). pCB594 the 4.6-kb

*Hin*dIII (*ruvB*) fragment inserted into pUC18 (pRuvB), and pCB550 the 5.5-kb *Bam*HI-*Eco*RI (*ruvAB*) fragment inserted into pUC18 (pRuvAB). A pUC18 clone carrying RecU (pFC204) was obtained by polymerase chain reaction (PCR) from pCB564 using 5′-AGAATTCTAAGGAGGATGAGATAATGATTC-3′ and 5′-TCTGACATAGGATCCCAACCTTTCG and *Eco*RI and *Bam*HI restriction endonucleases (underlined). To create a C-terminal fusion of RecU with YFP for single crossover integration into the chromosome, the 3′ region (500 bp) of *recU* was amplified by PCR using primers 5′-ATCGGGCCCTCGCGGAATGACCCTCG-3′ and 5′-CTAGAATTCACCTTTCGCACCAGATGATG-3′ and was cloned into *Apa*I and *Eco*RI sites of plasmid pYSG that carries *yfp* and *cat* genes and a xylose promoter for transcription of downstream genes (D. Kidane and P. L. Graumann, unpublished data), resulting in pYDK6. By transforming pyDK6 into PY79, we created the strain DK53. To move the *recU-yfp* fusion in different mutant backgrounds, DK53 was transformed with chromosomal DNA from the Δ*recN* strain, giving strain DK55, and the Δ*ruvAB* strain was transformed with chromosomal DNA of strain DK53, resulting in strain DK56. For the colocalization experiments, strain DK53 was transformed with chromosomal DNA from *recN-cfp*, generating strain DK54. Plasmid pGS739 was constructed by transferring the 1.7-kb *Bgl*II (*Sst*II)-*Eco*RV fragment containing the *ruvC* gene (obtained from a derivative of pFB512; Benson *et al.* 1988) into pACYC184 cleaved with *Bam*HI and *Hin*cII. This clone does not fully restore UV resistance to *ruv* mutants, possibly due to a slight negative effect on cell survival after UV exposure. Other plasmids used were pGS762, pGS711, and pPVA101 (Sharples *et al.* 1990; Sharples and Lloyd 1991).

Viability test: *B. subtilis* recombination-deficient strains were plated and incubated in Luria broth (LB) medium overnight. At least six independent colonies from each strain were resuspended in fresh LB medium and shaken for 30 min to minimize aggregation. Appropriate dilutions were plated on LB and colony-forming units (CFUs) were counted or stained with membrane-permeable SYTO 9 and membrane-impermeable propidium iodide and subjected to conventional direct count of total cells. SYTO 9, which labels bacteria with green fluorescence, and propidium iodide, which stains membrane-compromised bacteria with red fluorescence, were purchased from Molecular Probes (Leiden, The Netherlands).

DNA repair survival studies: Exponentially growing *B. subtilis* cells were obtained by inoculating overnight cultures in fresh LB media and grown to an A_{560nm} of 0.4 at 37°. These were exposed to 10 mM methyl methanesulfonate (MMS) and the fraction surviving was determined with reference to an unexposed control plate. Alternatively, the sensitivity to MMS, 4-nitroquinoline-1-oxide (4NQO), or mitomycin C (MMC) was determined by growing cultures to an A_{560nm} of 0.4 and spotting 10 μl of serial 10-fold dilutions (1×10^{-2} to 1×10^{-5}) on LB agar supplemented with the indicated concentrations of the DNA-damaging agent and incubating overnight at 37°.

E. coli strains carrying appropriate clones were measured for UV resistance by growing cells in LB media to an A_{650nm} of 0.4 and spotting appropriate dilutions onto agar plates. These were exposed to UV light at a dose rate of 1 J/m²/sec and the fraction surviving was determined with reference to an unirradiated control plate.

DNA lesions generated by MMS, which reacts with single reactive groups in adenine (N3-alkyladenine), guanine (N7-alkylguanine) and 4NQO, which is a potent mutagen that induces two main guanine adducts at positions C_8 and N_2 in damaged dsDNA or ssDNA. MMC results in the formation of interstrand crosslinks and UV light primarily induces pyrimidine dimers. All of these lesions act as DNA replication roadblocks, inducing replication fork arrest and DSBs.

Image acquisition: Fluorescence microscopy was performed on an Olympus AX70 microscope. Cells were grown in minimal medium and were mounted on agarose pads containing S750 medium on object slides as described in Kidane *et al.* (2004). Images were acquired with a digital MicroMax CCD camera; signal intensities were measured using the META-MORPH 4.6 program. DNA was stained with 4′,6 diamidino-2-phenylindole (DAPI; final concentration 0.2 ng/ml) and membrane with FM4-64 (final concentration 1 nM).

RESULTS

Defects in the α, ε, and η epistatic groups render cells extremely sensitive to DNA-damaging agents: To gain insight into the involvement of HJ processing in the repair of DNA damage, we constructed a null Δ*ruvAB* mutant strain and analyzed its phenotype in parallel with mutations in the α (*recF*15, *recL*16, Δ*recO*, Δ*recR*), β (*addA*5 *addB*72, termed here *addAB*), γ (*recH*342), ε (*recD*1, Δ*recU*, Δ*ruvAB*), ζ (Δ*recS*, Δ*recQ*, Δ*recJ*), and η (Δ*recG*) epistatic groups as well as in the Δ*recA* strain (Table 1).

The recombination-deficient cells, when present in an otherwise Rec⁺ strain, were exposed to the killing action of alkyl groups generated by MMS and their phenotypes were recorded. The Δ*recS*, Δ*recQ*, and Δ*recJ* cells (epistatic group ζ) showed a similar degree of sensitivity to MMS; hence, only the former is shown. Figure 1 shows that Δ*recS*, *addAB*, and *recH*342 cells displayed a moderate and/or sensitive phenotype to the killing action of 10 μM MMS when compared to the wild-type control.

The *recF*15, *recL*16, Δ*recO*, and Δ*recR* cells (group α) showed a similar degree of sensitivity to MMS or 4NQO (Alonso *et al.* 1993); hence, only the sensitivity of the former mutant strain is shown. The *recF*15, Δ*recG*, *recD*1, Δ*recU*, Δ*ruvAB*, and Δ*recA* cells were extremely sensitive to the killing action of 10 mM MMS when compared to the wild-type control. The *recF*15 (group α), *recD*1, Δ*recU*, Δ*ruvAB* (group ε), and Δ*recG* (group η) strains were less sensitive to 10 mM MMS than the Δ*recA* strain (Figure 1).

Branch migration and resolution of Holliday junctions is essential for DNA repair: Previously, it was shown that *addAB*, *recH*342, and Δ*recS* mutations increased the sensitivity of Δ*recU* cells to DNA damage (Fernandez *et al.* 1998). Furthermore, strains lacking RecF, RecU, or both are extremely sensitive to MMS and 4NQO (Alonso *et al.* 1993). These results indicate that RecA assembly factors (*e.g.*, RecFLOR) and the RecU HJ resolvase facilitate repair of DSBs (Fernandez *et al.* 1998; Ayora *et al.* 2004). We therefore investigated whether RecU was required for repair of DSBs generated by different DNA-damaging agents. The Δ*recU* null mutation was transferred into representatives from each of the epistatic groups (α, *recF*15 and Δ*recO*; β, *addA*5 *addB*72; γ, *recH*342; and ζ, Δ*recS* and Δ*recQ* strains), but we were unable to recover the Δ*recU* allele in Δ*ruvAB* (epistatic group ε), Δ*recG* (η), and Δ*recJ* (ζ) backgrounds without the appearance of undesired mutations. In our

TABLE 1

***B. subtilis* strains used in this study**

Bacterial strain	Epistatic group	Relevant genotype	Source or reference
YB886	NA	*trpC2 metB5 amyE sigB37 xin-1 att*SPβ	Yasbin *et al.* (1980)
BG190	NA	Δ*recA*	Ceglowski *et al.* (1990)
BG129	α	*recF15*	Alonso *et al.* (1988)
BG189	β	*addA5 addB72*	Alonso *et al.* (1993)
BG119	γ	*recH342*	Alonso *et al.* (1988)
BG633	ε	Δ*recU*	Fernandez *et al.* (1998)
BG121	ε	*recD41*	Alonso *et al.* (1988)
BG703	ε	Δ*ruvAB*	This work
BG707	η	Δ*recG*	This work
BG425	ζ	Δ*recS*	Fernandez *et al.* (1998)
BG705	ζ	Δ*recQ*	This work
BG675	ζ	Δ*recJ*	This work
BG501	NA + ε	Δ*sms* Δ*recU*	Carrasco *et al.* (2002)
BG651	NA + ε	Δ*recA* Δ*recU*	Carrasco *et al.* (2004)
BG703	NA + ε	Δ*recA* Δ*ruvAB*	This work
BG817	NA + η	Δ*recA* Δ*recG*	This work
BG717	α + ε	*recF15* Δ*ruvAB*	This work
BG735	β + ε	*addA5 addB72* Δ*ruvAB*	This work
BG783	γ + ε	*recH342* Δ*ruvAB*	This work
BG711	ε + ζ	Δ*ruvAB* Δ*recJ*	This work
BG709	ε + ζ	Δ*ruvAB* Δ*recQ*	This work
DK1	NA	*recN-cfp*	Kidane *et al.* (2004)
DK35	α	ΔrecN	Kidane *et al.* (2004)
DK53	NA	*recU-gfp*	This work
DK54	NA	*recU-yfp recN-cfp*	This work
DK55	α	*recU-gfp* ΔrecN	This work
DK56	ε	*recU-gfp* Δ*ruvAB*	This work

NA, not applied.

attempt to construct Δ*recU* Δ*ruvB*, Δ*recU* Δ*recG*, and Δ*recU* Δ*recJ* double mutants, we obtained a few colonies after prolonged incubation. Analysis of several of these transformants suggested that they contained either single CO or suppressor mutations (*e.g.*, Δ*recU* Δ*recG sms*; results not shown). To confirm that no other unselected mutations accumulate in these strains, DNA from a plasmid-borne *recG:six-cat-six* was used to transform *B. subtilis* BG501 (Δ*recU* Δ*sms*) competent cells selecting for chloramphenicol resistance. Using this approach, we succeeded in making a Δ*sms* Δ*recU* Δ*recG* strain. This fits with the earlier observations that Δ*sms* (also termed Δ*radA*) partially suppresses the Δ*recU* defect (Carrasco *et al.* 2002).

Unlike *Streptococcus pneumoniae* in which the *recU* gene (TIGR SP0370) is apparently essential (Thanassi *et al.* 2002), a *B. subtilis* Δ*recU* mutant is viable, although it grows poorly (Table 2) and accumulates suppressor mutations at a high frequency (Pedersen and Setlow 2000; Carrasco *et al.* 2002, 2004). Therefore, the Δ*ruvAB* null mutation was transferred into representatives from the different epistatic groups (α, *recF15* and Δ*recO*; β, *addA5 addB72*; γ, *recH342*; ε, *recD41*; and ζ, Δ*recS*, Δ*recQ*, and Δ*recJ* strains), the double and triple mutants were exposed to MMS, 4NQO, or MMC, and their phenotypes were recorded.

In the absence of any DNA-damaging agent, the number of viable cells per colony of strains grouped in the α, β, γ, or ζ epistatic group was affected <1.5-fold when compared to wild-type cells (data not shown), whereas the Δ*ruvAB* and Δ*recA* strains showed a similar reduced number of viable cells per colony (4- to 5-fold) when compared to the wild-type strain (Table 2). Exponentially growing cells were stained with SYTO 9, and only ~4% of these wild-type cells were also stained with propidium iodide (an indicator of membrane-compromised "dead" bacteria). The proportion of Δ*ruvAB* and Δ*recA* cells stained with propidium iodide increased 3- to 4-fold when compared with wild-type cells (Table 2). A similar reduced number of viable cells per colony was observed when the Δ*recU* or Δ*recG* cells were analyzed (Table 2).

The Δ*ruvAB* cells were extremely sensitive to 10 μg/ml of MMS, 0.75 μg/ml of 4NQO, or 12 ng/ml of MMC (Figure 2), whereas the wild-type strain showed a minimal defect in the presence of 250 μg/ml MMS, 24 μg/ml of 4NQO, or 150 ng/ml of MMC relative to an unexposed control (data not shown). The DNA damage sensitivity of Δ*ruvAB recD41* cells (epistatic group ε) was similar to that obtained with the Δ*ruvAB* mutant strain (Figure 2). The recombination mutants classified within the β (*addAB*) and ζ (Δ*recS*) groups marginally increased

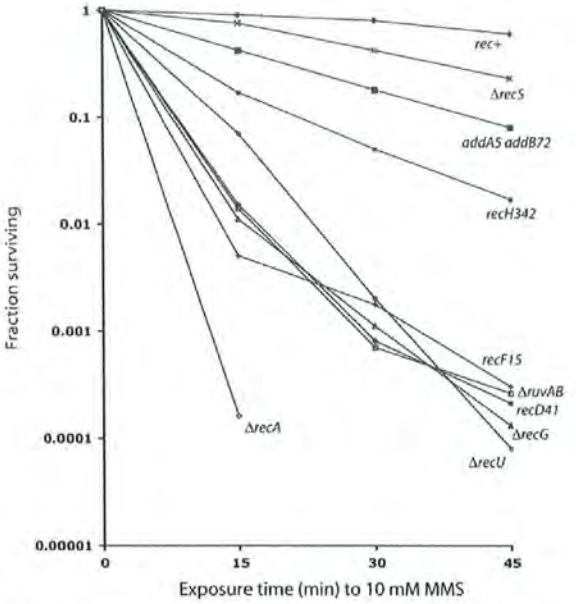

FIGURE 1.—Survival of strains after exposure to MMS. The strains used, which are identified by the indicated relevant genotype, were exposed to 10 mM MMS for a variable time.

the sensitivity of Δ*ruvAB* cells following exposure to 10 μg/ml MMS, 0.75 μg/ml of 4NQO, or 12 ng/ml of MMC (Figure 2), but they were slightly less sensitive than Δ*recA* cells. It is likely that acting, in concert, RecQ and RecJ initiate DNA recombination in Δ*ruvAB* cells.

The *recF*15 and the poorly characterized *recH*342 mutation reduced the survival of Δ*ruvAB* cells to a level comparable to the Δ*recA* strain following exposure to 5 μg/ml MMS, 0.35 μg/ml of 4NQO, or 6 ng/ml of MMC (Figure 2). The Δ*ruvAB* mutation did not increase the sensitivity of the Δ*ruvAB* Δ*recA* strain (Figure 2).

RecU resolves a HJ by endonucleolytic cleavage (AYORA *et al.* 2004), while it is believed that RuvAB, perhaps in concert with the unknown activity associated with RecD, recognizes and branch migrates HJs. The *recU*, *recD*, *ruvA*, and *ruvB* mutants all belong to the ϵ epistatic group. Since we failed to construct a Δ*recU* Δ*ruvAB* strain, yet successfully made Δ*recA* Δ*recU* (CARRASCO *et al.* 2004), Δ*recA* Δ*ruvAB* (this work), and Δ*recA* Δ*recG* strains (H. SANCHEZ and J. C. ALONSO, unpublished results), we propose that the Δ*recU*Δ*ruvAB* combination leads to accumulation of "toxic" recombination intermediates during strain construction.

RecU restores UV resistance to *E. coli ruvC* mutants: Recently it was demonstrated that RecU protein binds three- and four-stranded DNA branches, resolves Holliday junctions, and promotes joint molecule and D-loop formation *in vitro* (AYORA *et al.* 2004). Furthermore, the structure of RecU, which shows a striking similarity to a class of resolvase enzymes found in archaea and members of the type II restriction endonuclease family, was determined (MCGREGOR *et al.* 2005). To confirm the involvement of RuvAB and RecU in HJ processing, we examined their ability to replace the activities of their counterparts in the well-characterized *E. coli* system. Plasmid-borne *recU*, *ruvA*, *ruvB*, or *ruvAB* genes were introduced into various *E. coli ruv* mutant combinations and exposed to varying doses of UV light (Figure 3 and Table 3). Plasmids carrying RecU restored full UV resistance to strains (Δ$ruvC_{Eco}$ and $ruvC53_{Eco}$) deficient in the $RuvC_{Eco}$ HJ resolvase (Figure 3; Table 3; data not shown). RecU also conferred resistance to MMC at 0.2 and 0.5 μg/ml in these strains (data not shown). The results reveal for the first time that RecU functions as a HJ resolvase *in vivo*. Significantly, RecU only partially improved the UV sensitive phenotype in *E. coli ruvA*, *ruvB*, *ruvAB*, *ruvAC*, and *ruvABC* mutants (Figure 3;

TABLE 2

Viability of Δ*recA* recombination-deficient mutants

Strain	Relevant genotype	CFU per colony[a]	CFU per colony relative to wild type[b]	% propidium-iodide-stained cells[c]
YB886	Wild type	$2.2 \times 10^7 \pm 0.1$	1	4 (96)
BG190	Δ*recA*	$3.6 \times 10^6 \pm 0.2$	0.20	21 (79)
BG633	Δ*recU*	$3.7 \times 10^6 \pm 0.2$	0.19	21 (79)
BG703	Δ*ruvAB*	$3.7 \times 10^6 \pm 0.2$	0.24	28 (72)
BG707	Δ*recG*	$4.2 \times 10^6 \pm 0.1$	0.25	26 (73)
BG651	Δ*recA* Δ*recU*	$1.5 \times 10^6 \pm 0.1$	0.07	90 (9)
BG703	Δ*recA* Δ*ruvAB*	$1.3 \times 10^6 \pm 0.2$	0.06	93 (6)
BG817	Δ*recA* Δ*recG*	$1.0 \times 10^6 \pm 0.1$	0.04	94 (6)

Cells were grown to midexponential phase in LB medium and plated. Individual colonies after overnight incubation were resuspended in LB and appropriate dilutions were plated.

[a] The number of CFUs (viable cells) per colony, reported as the mean ± standard deviation, averaged from 5 to 10 colonies.

[b] Ratio of CFUs per total number of cells for mutant strains relative to wild type.

[c] Percentage of propidium-iodide-stained cells per colony, averaged from more than three colonies. The percentage of *only* SYTO 9-stained cells is indicated within parentheses.

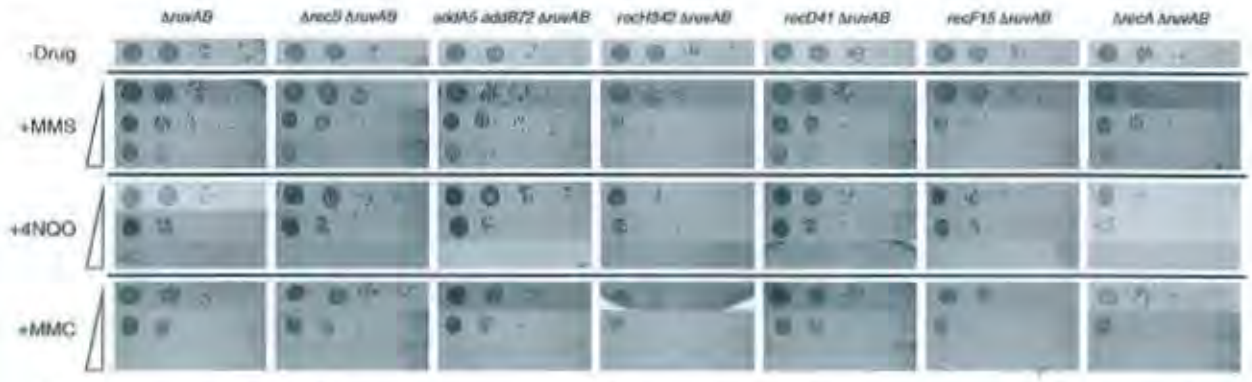

FIGURE 2.—Survival of Δ*ruvAB* mutants in combination with mutations from other epistatic groups after exposure to DNA-damaging agents. The strains used are identified by the indicated relevant genotype. A serial 10-fold dilution (10-μl sample) of a culture of each strain was spotted on LB agar plates containing the indicated concentration of the DNA-damaging agents. Dilution fractions were from 0.01 (left) to 0.00001 (right). Cells were exposed to different concentrations of MMS (2.5, 5, or 10 μg/ml), 4NQO (0.12, 0.35, or 0.75 μg/ml), or to MMC (6, 12, or 24 ng/ml) or plated in the absence of any drug (−drug).

Table 3; data not shown). Because RecU is as effective as $RuvC_{Eco}$ at promoting repair in a $ruvC_{Eco}$ mutant, the results establish that, as with $RuvC_{Eco}$ RecU depends on RuvAB branch migration for efficient HJ resolution.

The plasmid constructs carrying RuvA, RuvB, or RuvAB were unable to improve the UV sensitivity of the relevant ruv_{Eco} mutants (Table 3). Both RuvA and RuvAB clones produced an obvious negative effect on wild-type *E. coli* cells exposed to UV light. The results suggest that unlike RecU, RuvAB cannot replace the function of $RuvAB_{Eco}$ and that this may, in part, be due to a detrimental effect of RuvA expression. In fact, plasmids expressing high levels of $RuvA_{Eco}$ are known to confer an extreme negative effect on the UV sensitivity of wild-type cells (SHARPLES *et al.* 1990). This is probably a consequence of RuvA binding HJ DNA and preventing access of alternative junction processing enzymes such as $RuvC_{Eco}$ or $RecG_{Eco}$. These effects strengthen the argument that RuvAB cannot work properly with $RuvC_{Eco}$ rather than an expression and/or stability problem with the heterologous RuvAB complex.

We attempted to construct a plasmid carrying all three *B. subtilis* HJ processing genes to test RuvAB and RecU functionality directly in an *E. coli ruvABC*-deficient strain. However, the clones obtained had suffered substantial deletions, indicating that this combination is highly deleterious. Similarly, we were unable to maintain pCB559 (RuvAB) and pCB564 (RecU) jointly in a strain lacking *ruvAC* (effectively a *ruvABC* mutant) and the cryptic HJ resolvase *rusA* (MAHDI *et al.* 1996; data not shown). It seems likely that too much RuvAB and RecU generates frequent lethal DSBs at regressed replication forks or additionally blocks their processing even in the absence of DNA-damaging agents.

RecU forms discrete foci on nucleoids after induction of DSBs and colocalizes with RecN: Previously, it was shown that RecN, RecO, and RecF proteins accumulate in discrete foci following induction of DSBs (KIDANE *et al.* 2004). RecN foci were detected 15–20 min after treatment with MMC, RecO foci were first visible 30 min after induction, while RecF foci were not observed until after ~60–90 min (KIDANE *et al.* 2004; our unpublished results).

A RecU-GFP fusion strain was constructed and localization of the RecU protein was investigated in the presence or absence of MMC. The fusion was the sole source of RecU in the cell and fully supported repair of DNA following addition of MMC, showing that the fusion retained activity. In exponentially growing cells, RecU-GFP was present throughout the cells (Figure 4A), with fluorescence levels barely above background. However, after addition of 100 ng/ml of MMC, RecU-GFP formed discrete foci in up to 45% of the >500 cells analyzed (Figure 4C). RecU-GFP foci were always present on the nucleoids (see arrowheads in Figure 4C) and cells generally contained a single focus; only 2.7% of the cells contained two foci. One hour after the addition of MMC, clear foci were observed in only 1.5% of the cells (Figure 4B), with the highest number of foci observed 120 min after induction of DSBs (Figure 4C), and became increasingly fewer and fainter thereafter. The foci therefore occurred at a later point during DSB repair than RecN or RecO foci.

To establish that RecU is recruited to the RecNOF DSB repair centers (RCs), we generated a RecU-YFP variant and combined it with a RecN-CFP fusion, such that both were simultaneously expressed within cells. Many dually labeled cells showed rather patchy areas on the nucleoids (shaded arrow, Figure 4F); only 10% of the cells showed clear RecU-YFP and RecN-CFP foci after MMC treatment (open arrowheads, Figure 4F), mostly because RecN-CFP fluorescence was extremely low. The formation of patches in many cells suggests that GFP labels on both proteins slightly interfere with the proper function of the proteins, although the single labels are fully functional. However, in all of the cells

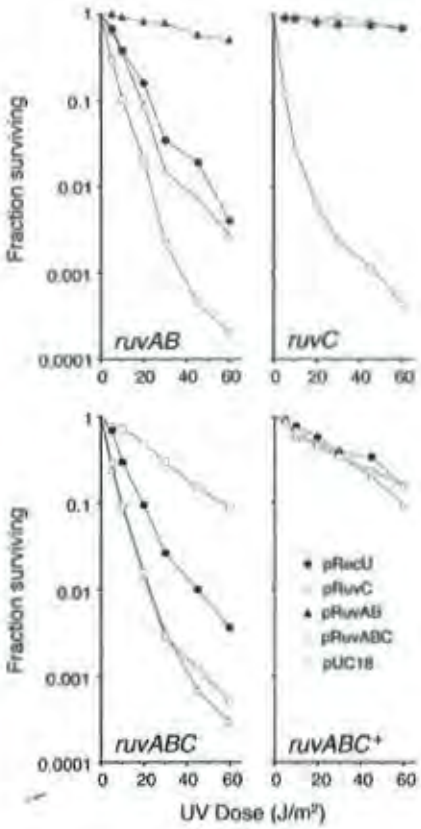

FIGURE 3.—Survival of UV-irradiated *E. coli ruv* mutants carrying the *recU* gene. The strains used were N2057 (*ruvAB*), GS1481 (*ruvC*), N4454 (*ruvABC*), and AB1157 (*ruv*$^+$). Symbols for plasmids carrying RecU (pFC204), $RuvC_{Ec}$ (pGS762), $RuvAB_{Ec}$ (pGS711), $RuvABC_{Ec}$ (pPVA101), and the pUC18 vector are shown (bottom). The relevant vector control for pPVA101, pHSG415 (not shown), was UV sensitive.

with clear foci, both RecU-YFP and RecN-CFP signals were coincident, and in most cells containing fluorescent patches these signals were likewise at similar places within the cells, showing that RecN and RecU colocalize within the DSB RCs.

To test if accumulation of RecU requires other proteins, we moved the RecU-YFP fusion into a *ruvAB* or *recN* mutant background. Only 0.4% of the cells showed RecU-YFP foci in the absence of *ruvAB* (Figure 4D), while 37% of the *recN* mutant cells contained RecU-YFP foci after addition of MMC (Figure 4E). Interestingly, 25% of the *recN* mutant cells contained two foci, rather than one (Figure 4E). As with RecU-YFP, RecN-GFP forms two foci in only 4% of the cells and a single focus in the remaining cells (KIDANE *et al.* 2004). These experiments demonstrate that RecU is part of a dynamic response to DSBs in *B. subtilis* cells and is recruited into defined RCs at a late stage in a reaction dependent on RuvAB. RecU is recruited to RCs independently of RecN in agreement with data showing that different avenues can lead to the formation of crossovers that are the substrate for RecU. However, on the basis of our previous findings suggesting that several DSBs are recruited to and repaired within a single RC (KIDANE *et al.* 2004), it is clear that RecN is a candidate for a factor combining different breaks into a single RC, because of the increase in the number of RecU-YFP foci in the absence of RecN.

DISCUSSION

This work provides evidence that inactivation of genes in epistatic group β (*addAB*) or in epistatic group ζ (*recQ*, *recS*, or *recJ*), when present in otherwise Rec$^+$ cells, have rather modest effects on sensitivity to MMS. In contrast, elimination of those functions classified within the α (namely *recF*15, *recL*16, Δ*recO*, or Δ*recR*), ε (Δ*ruvAB*, *recD*41, and Δ*recU*) or η (Δ*recG*) epistatic groups shows a dramatic reduction in ability to repair DNA damage mediated by these agents, showing only slightly more resistance than the recombination-defective Δ*recA* strain. Previously it was shown that RecS shares 36 and 34% identity with RecQ and $RecQ_{Ec}$ proteins, respectively (FERNANDEZ *et al.* 1998). This homology is significantly greater (43 and 40% identity) if only the regions containing the seven conserved DExH-box DNA helicase motifs (the first 330 residues of RecQ and RecS) are compared (FERNANDEZ *et al.* 1998). It was shown that Δ*recS* does substitute for the Δ*recQ* defect as the double mutant is as sensitive as the single parent mutant (our unpublished results). $RecQ_{Ec}$ unwinds both partially dsDNA and fully duplex DNA with a 3′ to 5′ polarity, while $RecJ_{Ec}$ is a 5′ to 3′ ssDNA exonuclease, generating a 3′-terminated end subsequently coated with SSB_{Ec} (KOWALCZYKOWSKI and EGGLESTON 1994; COURCELLE and HANAWALT 1999; AMUNDSEN and SMITH 2003). We have therefore tentatively placed *recJ* and *recQ* within the ζ epistatic group, together with *recS*, in a recombination pathway that can generate 3′-tailed ssDNA at broken forks akin to the activities of AddAB or RecBCD.

Since the DNA-damaging agents used in this report act as replication roadblocks, inducing replication fork arrest and single-strand nicks or DSBs, we considered the possibility that replication restart in Δ*ruvAB* cells relies on the processing of DNA ends by the action of AddAB (CHEDIN *et al.* 2000) or by the combined action of the RecQ or RecS helicase and the RecJ ssDNA exonuclease. RecA protein could be loaded on the 3′-ssDNA by the AddAB enzyme (CHEDIN *et al.* 2000) or RecN-RecFLOR complex (KIDANE *et al.* 2004). This is consistent with the observation that the Δ*recU* Δ*recJ* strain did not seem to be viable and that RecN forms RCs at DSBs in concert with RecO and RecF (KIDANE *et al.* 2004). RecA bound to ssDNA promotes homologous pairing, D-loop formation, and strand exchange between one or both of the broken ends with an intact DNA molecule to generate HJs. RecG or RuvAB (alone

TABLE 3

Effect of plasmids carrying RecU on the survival of UV-irradiated ruv_{Eco} mutants

		Fraction surviving (60 J/m²)				
Plasmid	*E. coli* strain: Genotype:	SR2210 *ruvA*	N1057 *ruvB*	N2057 *ruvAB*	GS1481 *ruvC*	AB1157 *ruv*⁺
pRecU[a]		0.0014	0.000057	0.0013	0.19	0.26
Vector[a]		0.00077	0.000074	0.0014	0.00010	0.56
pRuvA[b]		0.000068	0.000022	0.000075	ND	0.0033
pRuvB[b]		0.00014	0.000098	0.00032	ND	0.138
pRuvAB[b]		Sensitive	0.000025	0.0010	ND	0.0055
Vector[b]		0.000047	0.000010	0.00064	ND	0.57
pRuvA, B, AB_{Eco}[c]		0.0080	0.079	0.091	ND	0.17
$pRuvC_{Eco}$[d]		0.013	0.00045	0.00048	0.036	0.075
Vector[d]		0.0011	0.00050	0.00058	0.00016	0.48

ND, not done.

[a] pRecU (pCB564: *recU* in pHP13).

[b] pRuvA (pCB593: *ruvA* in pUC18), pRuvB (pCB594: *ruvB* in pUC18), and pRuvAB, pCB559 (*ruvAB* in pUC18).

[c] $pRuvA_{Eco}$ (pGT14: $ruvA_{Eco}$ in pUC18, control for SR2210), $pRuvB_{Eco}$ (pGT119: $ruvB_{Eco}$ in pUC18, control for N1057), and $pRuvAB_{Eco}$ (pGS711: $ruvAB_{Eco}$ in pUC18, control for N2057).

[d] $pRuvC_{Eco}$ (pGS739: $ruvC_{Eco}$ in pACYC184). Relevant strain genotypes are indicated.

or in concert with RecD) branch migrates these junctions for RecU resolution. This model fits with the observations that: (i) *recF addAB* cells are impaired in DNA repair and genetic recombination to the level of *recA* cells and showed a similarly reduced viability in the absence of external damage, together with extreme sensitivity to the killing action of MMS, 4NQO, or MMC (Alonso *et al.* 1993), and (ii) *recF* or Δ*recO* mutations reduce the viability of Δ*ruvAB* cells to a greater extent than do *addAB* mutations (see Figure 2; our unpublished results).

To confirm the functionality of RuvAB and RecU in HJ processing, we studied their ability to complement the DNA repair defect *of E. coli ruv* mutants. We found that expression of the *recU* gene restores UV-light resistance to $ruvC_{Eco}$ strains to a level similar to that of clones carrying $RuvC_{Eco}$. In contrast, RecU conferred only a slight improvement in UV survival of $ruvA_{Eco}$ $ruvB_{Eco}$ $ruvAC_{Eco}$ $ruvAB_{Eco}$ or $ruvABC_{Eco}$ mutants. Since the *E. coli ruv* system is well defined, we can conclude that RecU does indeed function as a HJ resolvase as demonstrated by *in vitro* data (Ayora *et al.* 2004; McGregor *et al.* 2005). The improvement in resistance to UV when RecU is present in strains lacking $ruvAB_{Eco}$ indicates that it can function to some extent in the absence of RuvAB. However, this may be artificially high due to overexpression of RecU, since clones carrying $RuvC_{Eco}$ also improve the survival of $ruvAB_{Eco}$ strains following exposure to UV light. The dependence on RuvAB for full DNA repair activity does suggest that RecU normally functions together with the branch migration complex as is the case with *E. coli* RuvABC (Zerbib *et al.* 1998; van Gool *et al.* 1999). Any contacts that stabilize a $RuvABC_{Eco}$ or a RuvAB-RecU complex must be conserved between these heterologous systems, if indeed they are important for stability of the tripartite complex. Consequently, the resolvasome model, where the resolution endonuclease (either RecU or $RuvC_{Eco}$) scans the junction for preferred target sequences, appears to be widely conserved in bacteria.

$RuvAB_{Eco}$ or $RecG_{Eco}$ catalyze replication fork regression *in vivo* and play a critical role in promoting the recovery of replication when it is blocked by DNA damage (Bolt and Lloyd 2002; Gregg *et al.* 2002; Meddows *et al.* 2004). Other studies, however, indicate that $RuvAB_{Eco}$ or $RecG_{Eco}$-catalyzed fork regression is not essential for DNA synthesis to resume following arrest by UV-induced DNA damage *in vivo* (Donaldson *et al.* 2004). In this work, we also show that it is possible to visualize the place of action of RecU in live cells. We have found that RecU forms a single discrete center on the nucleoid upon induction of DSBs, as previously observed with RecNOF proteins (Kidane *et al.* 2004). RecN is the first to form the RCs, within 15–20 min with foci visible at defined DSBs in live cells (Kidane *et al.* 2004). RecU is recruited into RCs, since it colocalizes with RecN. Consistent with a role in resolution of HJs, RecU accumulated within the RCs after the formation of RecN, RecO, or RecF foci; RecU foci were clearly visible 120 min after induction of DSBs. These data indicate that repair of DSBs occurs over a long period of time during which several sequential processes take place. Recruitment of RecU was dependent on RuvAB proteins, strengthening the view that these proteins form a resolvasome complex. Interestingly, the number of RCs was increased in the absence of RecN protein, supporting our suggestion that RecN might organize different recombination events within a single center (Kidane *et al.* 2004).

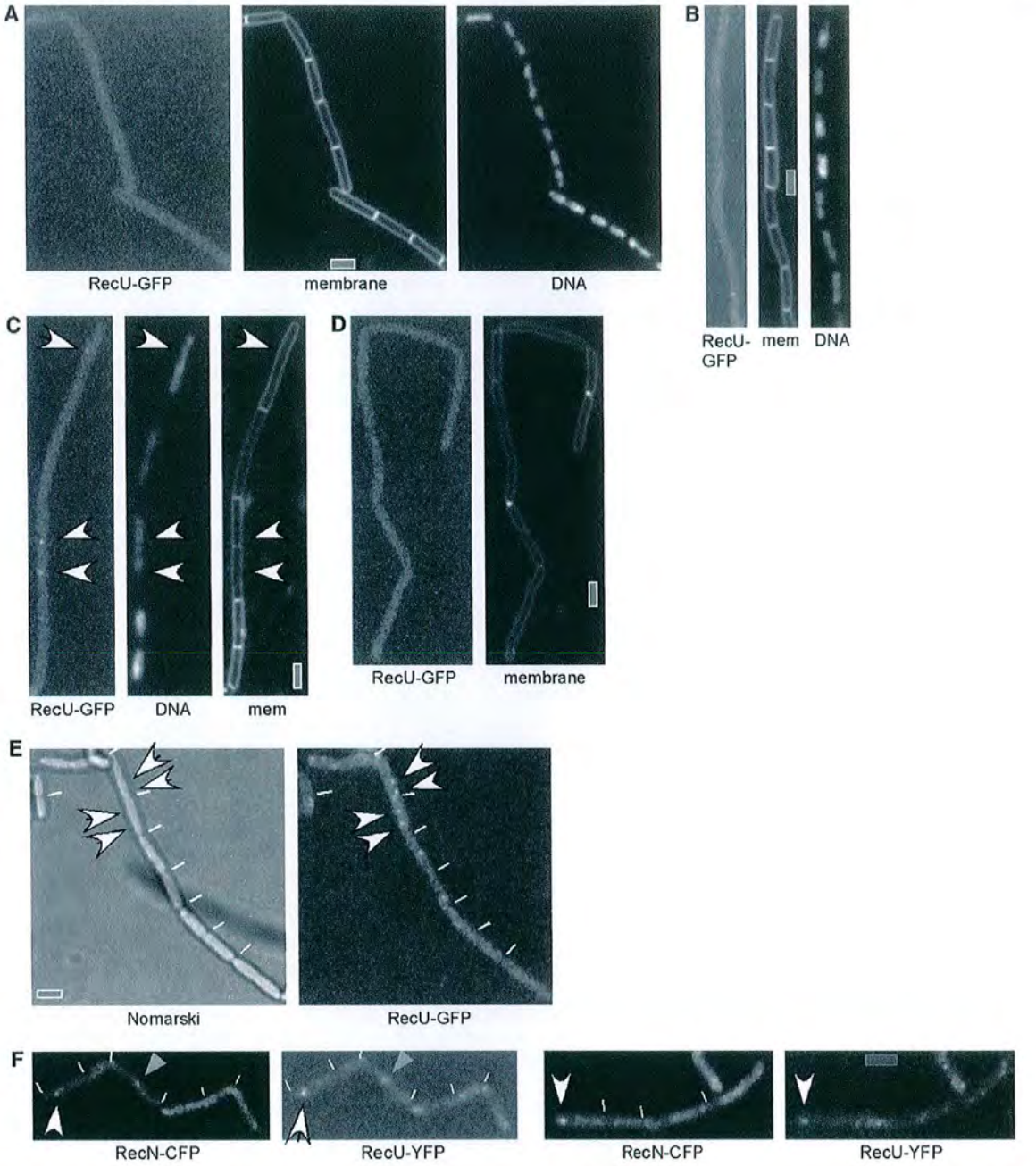

FIGURE 4.—Subcellular localization of RecU in live *B. subtilis* cells. (A) RecU-GFP in exponentially growing cells; (B) 60 min after addition of MMC; (C) 120 min after addition of MMC. Open arrowheads in C indicate RecU-GFP foci. (D) RecU-GFP in Δ*ruvAB* cells, 120 min after addition of MMC. (E) RecU-YFP in Δ*recN* cells, 120 min after addition of MMC; open arrowheads indicate two RecU-YFP foci per cell. (F) Colocalization of RecU-YFP and RecN-CFP 120 min after addition of MMC, indicated by open arrowheads; shaded arrowheads indicate the patch formed by RecN and RecU and open lines indicate the ends of cells. DNA is stained with DAPI and membranes (mem) by FM4-64. Bar, 2 μm.

Under standard transformation conditions we failed to construct Δ*ruvAB* Δ*recG*, Δ*ruvAB* Δ*recU*, and Δ*recU* Δ*recG* mutant strains. Previously we reported the construction of *ruvA*2 *recU*40 double-mutant strains (ALONSO *et al.* 1992) and here report the construction of the *recD*41 Δ*ruvAB* strain. However, the *recU*40 strain is proficient in plasmid transformation and shows a doubling time similar to the wild-type strain, whereas a Δ*recU* is significantly impaired in plasmid transformation and shows a marked growth defect (FERNANDEZ

et al. 1998; Pedersen and Setlow 2000; Carrasco *et al.* 2002; Table 2). It is likely that the *recU40* allele may encode only a partially defective HJ resolvase. Very little information is available concerning the *recD41* strain. These two pieces of apparently conflicting data argue that RecU resolves HJ intermediates branch migrated by either RuvAB(RecD) or RecG DNA helicases. The improvements in UV resistance conferred upon $ruvAB_{Eco}$ mutants by clones carrying $RuvC_{Eco}$ and RecU support this idea, confirming that both these HJ resolvases can function *in vivo* without RuvAB branch migration. Alternatively, the apparent lethality of Δ*recG* Δ*ruvAB*, Δ*recG* Δ*recU*, and Δ*ruvAB* Δ*recU* double mutants arises from accumulation of "toxic" recombination intermediates (Gangloff *et al.* 2000). This is consistent with the observation that Δ*recA* Δ*recU*, Δ*recA* Δ*recG* (Carrasco *et al.* 2004), and Δ*recA* Δ*ruvAB* were viable, albeit with a 14- to 25-fold reduced plating efficiency (see Table 2). In *E. coli*, *recG ruvAB*, *recG ruvC*, and *ruvABC* mutants are viable, the latter indistinguishable from single-mutant strains (Lloyd 1991; Mandal *et al.* 1993). There are clearly important differences between Gram-negative and Gram-positive recombinational repair processes despite the apparent similarities in coordination of HJ resolution by RuvAB-RecU and $RuvABC_{Eco}$. This serves to highlight the importance of having more than one model system to evaluate the mechanics of complex repair, replication, and recombination processes.

We are grateful to Begoña Carrasco for the communication of unpublished results. This work was supported by grants BMC2003-00150 and BIO2001-4342-E from Dirección General de Investigación (J.C.A.), from the Deutsche Forschungsgemeinschaft (P.L.G.), and from the Biotechnology and Biological Sciences Research Council (G.J.S.).

LITERATURE CITED

Alonso, J. C., R. H. Tailor and G. Luder, 1988 Characterization of recombination-deficient mutants of *Bacillus subtilis*. J. Bacteriol. **170:** 3001–3007.

Alonso, J. C., G. Luder and T. A. Trautner, 1992 Intramolecular homologous recombination in *Bacillus subtilis* 168. Mol. Gen. Genet. **236:** 60–64.

Alonso, J. C., A. G. Stiege and G. Luder, 1993 Genetic recombination in *Bacillus subtilis* 168: effect of *recN*, *recF*, *recH* and *addAB* mutations on DNA repair and recombination. Mol. Gen. Genet. **239:** 129–136.

Amundsen, S. K., and G. R. Smith, 2003 Interchangeable parts of the *Escherichia coli* recombination machinery. Cell **112:** 741–744.

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

Arnold, D. A., and S. C. Kowalczykowski, 2000 Facilitated loading of RecA protein is essential to recombination by RecBCD enzyme. J. Biol. Chem. **275:** 12261–12265.

Ayora, S., B. Carrasco, E. Doncel, R. Lurz and J. C. Alonso, 2004 *Bacillus subtilis* RecU protein cleaves Holliday junctions and anneals single-stranded DNA. Proc. Natl. Acad. Sci. USA **101:** 452–457.

Benson, F. E., G. T. Illing, G. J. Sharples and R. G. Lloyd, 1988 Nucleotide sequencing of the *ruv* region of *Escherichia coli* K-12 reveals a LexA regulated operon encoding two genes. Nucleic Acids Res. **16:** 1541–1549.

Bolt, E. L., and R. G. Lloyd, 2002 Substrate specificity of RusA resolvase reveals the DNA structures targeted by RuvAB and RecG *in vivo*. Mol. Cell **10:** 187–198.

Carrasco, B., S. Fernandez, K. Asai, N. Ogasawara and J. C. Alonso, 2002 Effect of the *recU* suppressors *sms* and *subA* on DNA repair and homologous recombination in *Bacillus subtilis*. Mol. Genet. Genomics **266:** 899–906.

Carrasco, B., M. C. Cozar, R. Lurz, J. C. Alonso and S. Ayora, 2004 Genetic recombination in *Bacillus subtilis* 168: contribution of Holliday junction-processing functions in chromosome segregation. J. Bacteriol. **186:** 5557–5566.

Ceglowski, P., G. Luder and J. C. Alonso, 1990 Genetic analysis of *recE* activities in *Bacillus subtilis*. Mol. Gen. Genet. **222:** 441–445.

Cheldin, F., and S. C. Kowalczykowski, 2002 A novel family of regulated helicases/nucleases from Gram-positive bacteria: insights into the initiation of DNA recombination. Mol. Microbiol. **43:** 823–834.

Cheldin, 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:** 7–20.

Clark, A. J., and S. J. Sandler, 1994 Homologous genetic recombination: the pieces begin to fall into place. Crit. Rev. Microbiol. **20:** 125–142.

Courcelle, J., and P. C. Hanawalt, 1999 RecQ and RecJ process blocked replication forks prior to the resumption of replication in UV-irradiated Escherichia coli. Mol. Gen. Genet. **262:** 543–551.

Courcelle, J., A. K. Ganesan and P. C. Hanawalt, 2001 Therefore, what are recombination proteins there for? BioEssays **23:** 463–470.

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

Donaldson, J. R., C. T. Courcelle and J. Courcelle, 2004 RuvAB and RecG are not essential for the recovery of DNA synthesis following UV-induced DNA damage in *Escherichia coli*. Genetics **166:** 1631–1640.

Fernandez, S., A. Sorokin and J. C. Alonso, 1998 Genetic recombination in *Bacillus subtilis* 168: effects of *recJ* and *recS* mutations on DNA repair and homologous recombination. J. Bacteriol. **180:** 3405–3409.

Fernandez, S., Y. Kobayashi, N. Ogasawara and J. C. Alonso, 1999 Analysis of the *Bacillus subtilis* recO gene: RecO forms part of the RecFLOR function. Mol. Gen. Genet. **261:** 567–573.

Fernandez, S., S. Ayora and J. C. Alonso, 2000 *Bacillus subtilis* homologous recombination: genes and products. Res. Microbiol. **151:** 481–486.

Gangloff, S., C. Soustelle and F. Fabre, 2000 Homologous recombination is responsible for cell death in the absence of the Sgs1 and Srs2 helicases. Nat. Genet. **25:** 192–194.

Gregg, A. V., P. McGlynn, R. P. Jaktaji and R. G. Lloyd, 2002 Direct rescue of stalled DNA replication forks via the combined action of PriA and RecG helicase activities. Mol. Cell **9:** 241–251.

Haber, J. E., 1999 DNA recombination: the replication connection. Trends Biochem. Sci. **24:** 271–275.

Ivancic-Bace, I., P. Peharec, S. Moslavac, N. Skrobot, E. Salaj-Smic *et al.*, 2003 RecFOR function is required for DNA repair and recombination in a RecA loading-deficient recB mutant of *Escherichia coli*. Genetics **163:** 485–494.

Kantake, N., M. V. Madiraju, T. Sugiyama and S. C. Kowalczykowski, 2002 *Escherichia coli* RecO protein anneals ssDNA complexed with its cognate ssDNA-binding protein: a common step in genetic recombination. Proc. Natl. Acad. Sci. USA **99:** 15327–15332.

Kidane, D., H. Sanchez, J. C. Alonso and P. L. Graumann, 2004 Visualization of DNA double strand breaks repair in live bacteria reveals dynamic recruitment of *Bacillus subtilis* RecF, RecO and RecN protein to distinct sites on the nucleoids. Mol. Microbiol. **52:** 1627–1639.

Kowalczykowski, S. C., 2000 Initiation of genetic recombination and recombination-dependent replication. Trends Biochem. Sci. **25:** 156–165.

Kowalczykowski, S. C., and A. K. Eggleston, 1994 Homologous pairing and DNA strand-exchange proteins. Annu. Rev. Biochem. **63:** 991–1043.

KUZMINOV, A., 1999 Recombinational repair of DNA damage in *Escherichia coli* and bacteriophage lambda. Microbiol. Mol. Biol. Rev. **63:** 751–813.

LLOYD, R. G., 1991 Conjugational recombination in resolvase-deficient *ruvC* mutants of *Escherichia coli* K-12 depends on recG. J. Bacteriol. **173:** 5414–5418.

MAHDI, A. A., G. J. SHARPLES, T. N. MANDAL and R. G. LLOYD, 1996 Holliday junction resolvases encoded by homologous rusA genes in *Escherichia coli* K-12 and phage 82. J. Mol. Biol. **257:** 561–573.

MANDAL, T. N., A. A. MAHDI, G. J. SHARPLES and R. G. LLOYD, 1993 Resolution of Holliday intermediates in recombination and DNA repair: indirect suppression of *ruvA*, *ruvB*, and *ruvC* mutations. J. Bacteriol. **175:** 4325–4334.

MARIANS, K. J., 2000 Replication and recombination intersect. Curr. Opin. Genet. Dev. **10:** 151–156.

MCGLYNN, P., and R. G. LLOYD, 2002a Genome stability and the processing of damaged replication forks by RecG. Trends Genet. **18:** 413–419.

MCGLYNN, P., and R. G. LLOYD, 2002b Recombinational repair and restart of damaged replication forks. Nat. Rev. Mol. Cell Biol. **3:** 859–870.

MCGREGOR, N., S. AYORA, S. SEDELNIKOVA, B. CARRASCO, J. C. ALONSO *et al.*, 2005 The structure of *Bacillus subtilis* RecU Holliday junction resolvase and its role in substrate selection and sequence specific cleavage. Structure **13:** 1341–1351.

MEDDOWS, T. R., A. P. SAVORY and R. G. LLOYD, 2004 RecG helicase promotes DNA double-strand break repair. Mol. Microbiol. **52:** 119–132.

MICHEL, B., G. GROMPONE, M. J. FLORES and V. BIDNENKO, 2004 Multiple pathways process stalled replication forks. Proc. Natl. Acad. Sci. USA **101:** 12783–12788.

PEDERSEN, L. B., and P. SETLOW, 2000 Penicillin-binding protein-related factor A is required for proper chromosome segregation in *Bacillus subtilis*. J. Bacteriol. **182:** 1650–1658.

ROBU, M. E., R. B. INMAN and M. M. COX, 2001 RecA protein promotes the regression of stalled replication forks in vitro. Proc. Natl. Acad. Sci. USA **98:** 8211–8218.

ROBU, M. E., R. B. INMAN and M. M. COX, 2004 Situational repair of replication forks: roles of RecG and RecA proteins. J. Biol. Chem. **279:** 10973–10981.

SARGENTINI, N. J., and K. C. SMITH, 1989 Role of *ruvAB* genes in UV- and gamma-radiation and chemical mutagenesis in Escherichia coli. Mutat. Res. **215:** 115–129.

SEIGNEUR, M., V. BIDNENKO, S. D. EHRLICH and B. MICHEL, 1998 RuvAB acts at arrested replication forks. Cell **95:** 419–430.

SHAN, Q., J. M. BORK, B. L. WEBB, R. B. INMAN and M. M. COX, 1997 RecA protein filaments: end-dependent dissociation from ssDNA and stabilization by RecO and RecR proteins. J. Mol. Biol. **265:** 519–540.

SHARPLES, G. J., and R. G. LLOYD, 1991 Resolution of Holliday junctions in *Escherichia coli*: identification of the *ruvC* gene product as a 19-kilodalton protein. J. Bacteriol. **173:** 7711–7715.

SHARPLES, G. J., F. E. BENSON, G. T. ILLING and R. G. LLOYD, 1990 Molecular and functional analysis of the *ruv* region of *Escherichia coli* K-12 reveals three genes involved in DNA repair and recombination. Mol. Gen. Genet. **221:** 219–226.

SINGLETON, M. R., S. SCAIFE and D. B. WIGLEY, 2001 Structural analysis of DNA replication fork reversal by RecG. Cell **107:** 79–89.

THANASSI, J. A., S. L. HARTMAN-NEUMANN, T. J. DOUGHERTY, B. A. DOUGHERTY and M. J. PUCCI, 2002 Identification of 113 conserved essential genes using a high-throughput gene disruption system in *Streptococcus pneumoniae*. Nucleic Acids Res. **30:** 3152–3162.

UMEZU, K., and R. D. KOLODNER, 1994 Protein interactions in genetic recombination in *Escherichia coli*. Interactions involving RecO and RecR overcome the inhibition of RecA by single-stranded DNA-binding protein. J. Biol. Chem. **269:** 30005–30013.

VAN GOOL, A. J., N. M. HAJIBAGHERI, A. STASIAK and S. C. WEST, 1999 Assembly of the *Escherichia coli* RuvABC resolvasome directs the orientation of Holliday junction resolution. Genes Dev. **13:** 1861–1870.

XU, L., and K. J. MARIANS, 2003 PriA mediates DNA replication pathway choice at recombination intermediates. Mol. Cell **11:** 817–826.

YASBIN, R. E., P. I. FIELDS and B. J. ANDERSEN, 1980 Properties of *Bacillus subtilis* 168 derivatives freed of their natural prophages. Gene **12:** 155–159.

ZERBIB, D., C. MEZARD, H. GEORGE and S. C. WEST, 1998 Co-ordinated actions of RuvABC in Holliday junction processing. J. Mol. Biol. **281:** 621–630.

Communicating editor: S. T. LOVETT