View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

Cell, Vol. 107, 79–89, October 5, 2001, Copyright ©2001 by Cell Press

Structural Analysis of DNA Replication Fork Reversal by RecG

Martin R. Singleton, Sarah Scaife,[2] and Dale B. Wigley[1]
ICRF Clare Hall Laboratories
Blanche Lane
South Mimms
Potters Bar
Hertfordshire EN6 3LD
United Kingdom

Summary

The stalling of DNA replication forks that occurs as a consequence of encountering DNA damage is a critical problem for cells. RecG protein is involved in the processing of stalled replication forks, and acts by reversing the fork past the damage to create a four-way junction that allows template switching and lesion bypass. We have determined the crystal structure of RecG bound to a DNA substrate that mimics a stalled replication fork. The structure not only reveals the elegant mechanism used by the protein to recognize junctions but has also trapped the protein in the initial stage of fork reversal. We propose a mechanism for how forks are processed by RecG to facilitate replication fork restart. In addition, this structure suggests that the mechanism and function of the two largest helicase superfamilies are distinct.

Introduction

Chromosomal replication in eubacteria typically begins at a unique site, termed the "origin" (Kornberg and Baker, 1992). DNA replication then proceeds bidirectionally to produce so-called "theta" structures (Cairns, 1963) which contain two replication forks, each progressing away from the replication origin via the action of a multiprotein complex called the replisome. The replisome contains the DNA polymerase III holoenzyme, as well as the primosome, another multiprotein complex that includes the hexameric ring helicase DnaB, PriA protein, DNA primase, and several other proteins. At a position roughly equidistant from the origin in either direction around the circular chromosome, the two forks meet at a sequence that directs the termination of replication and disassembly of the replication apparatus via the action of a terminator protein that is bound at this site. For many years, it was envisaged that replication of the leading strand at each replication fork would be a continuous, highly processive process that would continue uninterrupted around the chromosome. Replication of the lagging strand would take place at regular intervals from newly synthesized RNA primers that initiate Okazaki fragments. However, more recently, it has become apparent that this is a naïve view of the process (Kowalczykowski, 2000; Cox et al., 2000). A problem arises when the fork encounters DNA damage in the form of either a base lesion or a single-stranded nick in the DNA (Figure 1). In either of these situations, the replication fork stalls, and the replisome disassembles. Fork progression can be recovered via any of a number of parallel pathways, probably dependent upon the nature of the DNA damage (Figure 1). In all of these cases, the primosomal protein, PriA, directs the reestablishment of the replisome, thereby allowing replication to restart (Marians, 2000). Although the idea of recombination-dependent replication-restart is not a new one (Skalka, 1974; Mosig, 1987; Asai et al., 1993; Kuzminov, 1995), the frequency and extent of this process has only recently been appreciated. It now seems that most, perhaps all, replication forks encounter some form of damage and have to restart during each cycle of chromosomal replication. It is not surprising, therefore, that there are a large number of proteins associated with this essential function. Recent estimates for *E. coli* are that at least 26 proteins are implicated, in addition to the multisubunit DNA polymerase III holoenzyme (Cox et al., 2000). At the present time there is structural information for only a handful of these proteins.

Recombination-mediated repair of stalled replication forks appears to take place by several distinct mechanisms (Figure 1). Two of these are thought to involve RecA in forming a four-way (Holliday junction) intermediate that is migrated and subsequently resolved to recreate a repaired fork. One system known to be able to migrate Holliday junctions is the RuvAB complex (reviewed in West, 1996). RuvA is a tetrameric protein that recognizes and binds tightly to Holliday junctions. The RuvB protein is a hexameric helicase that is recruited to the RuvA/Holliday junction complex and drives junction migration. Finally, the junction is resolved by the RuvC endonuclease. Crystal structures of the RuvA tetramer (Rafferty et al., 1996), the RuvA:Holliday junction complex (Hargreaves et al., 1998; Roe et al., 1998; Ariyoshi et al., 2000), and the RuvB monomer (Yamada et al., 2001; Putnam et al., 2001) have revealed how RuvAB interacts with Holliday junctions. The RuvA tetramer binds to the junction so that each of the four arms are arranged symmetrically around a central acidic "pin" region. The RuvB protein is then thought to act by pulling two of the arms of the junction, forcing the DNA strands to separate either side of the acidic pins of the RuvA protein, thereby inducing the strand exchange required for junction migration. Despite the availability of several crystal structures and considerable biochemical data (reviewed in West, 1996), the molecular details of the process by which the RuvAB protein complex drives the migration of Holliday junctions are still not well understood.

Genetic and biochemical evidence appeared to have revealed another system for branch migration (Lloyd, 1991; Whitby et al., 1993). This parallel system involves a single protein, RecG, that is also able to migrate Holliday junctions prior to their resolution by RuvC (McGlynn and Lloyd, 2000). However, while RecG is a Superfamily 2

[1] Correspondence: d.wigley@icrf.icnet.uk
[2] Present address: Yamanouchi Research Institute, Littlemore Park, Littlemore, Oxfordshire OX4 4SS, United Kingdom

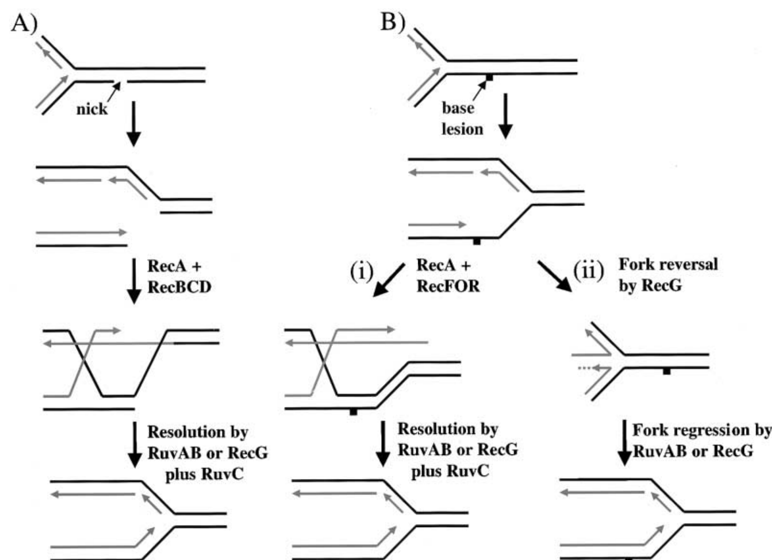

Figure 1. Potential Mechanisms for Recovery of Stalled Replication Forks

(A) When a single-strand nick is present in the template, this results in a double-strand break after passage of the replication fork. This is thought to be repaired by the actions of RecA and RecBCD proteins to form a Holliday junction intermediate, which is then migrated and resolved by either the RuvABC complex or RecG. PriA protein then mediates reassembly of the replisome.

(B) When a base lesion is encountered, the replisome stalls and disassembles. The fork is then repaired either by (i) RecA and RecFOR, followed by Holliday junction resolution and reestablishment of the fork by PriA, or (ii) RecG-mediated fork reversal to a "chicken foot" intermediate to allow template switching, followed by regression of the fork and PriA-mediated replisome assembly. It is this latter pathway that is thought to be the principal role of RecG in vivo. Adapted from Cox et al. (2000) and McGlynn and Lloyd (2000).

(SF2) DNA helicase (Gorbalenya and Koonin, 1993), RuvB is a member of the AAA^{+} protein family (Neuwald et al., 1999), and outside of the helicase motifs, there is no sequence homology between RuvB and RecG. Furthermore, whereas RuvB is a hexameric protein that requires RuvA for junction migration activity, RecG functions alone as a monomer (McGlynn et al., 2000). The relationship between these proteins was therefore uncertain. Despite the considerable recent advances in our understanding of RecG (Mahdi et al., 1997; Whitby and Lloyd, 1998; McGlynn et al., 2000; McGlynn and Lloyd, 2001), the details about how this protein is able to recognize junctions and drive branch migration remain unclear. Furthermore, it has been shown that RecG is also efficient at unwinding DNA:RNA hybrid structures (so-called R loops (Hong et al., 1995; Vincent et al., 1996)).

Given the important role of RecG in bacteria, there is a surprising lack of RecG homologs in other organisms, at least at the sequence level. However, recent evidence reveals that the phage T4 protein UvsW can complement a recG defect in *E. coli*, showing that it is a functional homolog of RecG (Dudas and Kreuzer, 2001). In vitro, the protein is also capable of unwinding R loops. Outside of the helicase motifs, there is no detectable sequence homology between UvsW and RecG, and the proteins are of very different sizes. The conservation of function rather than sequence raises the possibility that there may be proteins with RecG-like functions in other organisms. Indeed, it has been suggested that a number of helicases (e.g., Werner's, BLM, and Sgs1) may play a role in the recovery of stalled replication forks in eukaryotes (Frei and Gasser, 2001).

Although initially identified as a Holliday junction migrating protein in vitro, more recent evidence favors a different role for RecG in vivo (McGlynn and Lloyd, 2000, 2001; McGlynn et al., 2001). RecG is able to convert three-way (fork) junctions into Holliday junctions, particularly those in which DNA synthesis on the leading strand has stalled prematurely. The function of this process appears to be to allow reversal of a fork that has stalled on the leading strand so that replication can continue by template switching (Higgins et al., 1976; McGlynn and Lloyd, 2000). Fork reversal results in the formation of "chicken foot" intermediates (Figure 1). Formation of these intermediates can be driven by supercoiling (Postow et al., 2001), or can be catalysed by RecG (McGlynn and Lloyd, 2000, 2001). The observation that RecG can operate upon stalled forks within negatively supercoiled DNA supports this role for the enzyme in vivo (McGlynn et al., 2001). Once the chicken foot has been formed, DNA synthesis on the stalled strand can continue by switching template strands to bypass the lesion. Subsequently, the Holliday junction can be regressed past the lesion to recover the fork, presumably by either RecG or RuvAB, with repair of the lesion taking place at a later stage.

To understand more about the mechanism by which RecG is able to reverse replication forks, we have determined the crystal structure of the *Thermatoga maritima* RecG protein complexed with ADP and a synthetic three-way DNA junction that resembles a leading strand stalled replication fork. The protein is monomeric with the DNA bound mainly to the large N-terminal domain of RecG, a domain that is not found in other DNA helicases. This region of the protein not only clamps onto and splits open the junction, but also stabilizes unwinding of the fork. In the structure we have determined, the junction has already begun to unwind, catching the complex in the initial stages of fork reversal. The template arm of the DNA (i.e., the region that would precede the moving replication fork) is bound across the interface between the N- and C-terminal domains, suggesting a novel mechanism for DNA unwinding.

Results and Discussion

Our aim was to determine the structure of a complex between RecG and a synthetic DNA substrate that would mimic a stalled replication fork. Although a wide range of three-way junctions were tried, and several produced crystals, the only one to produce diffraction-quality crys-

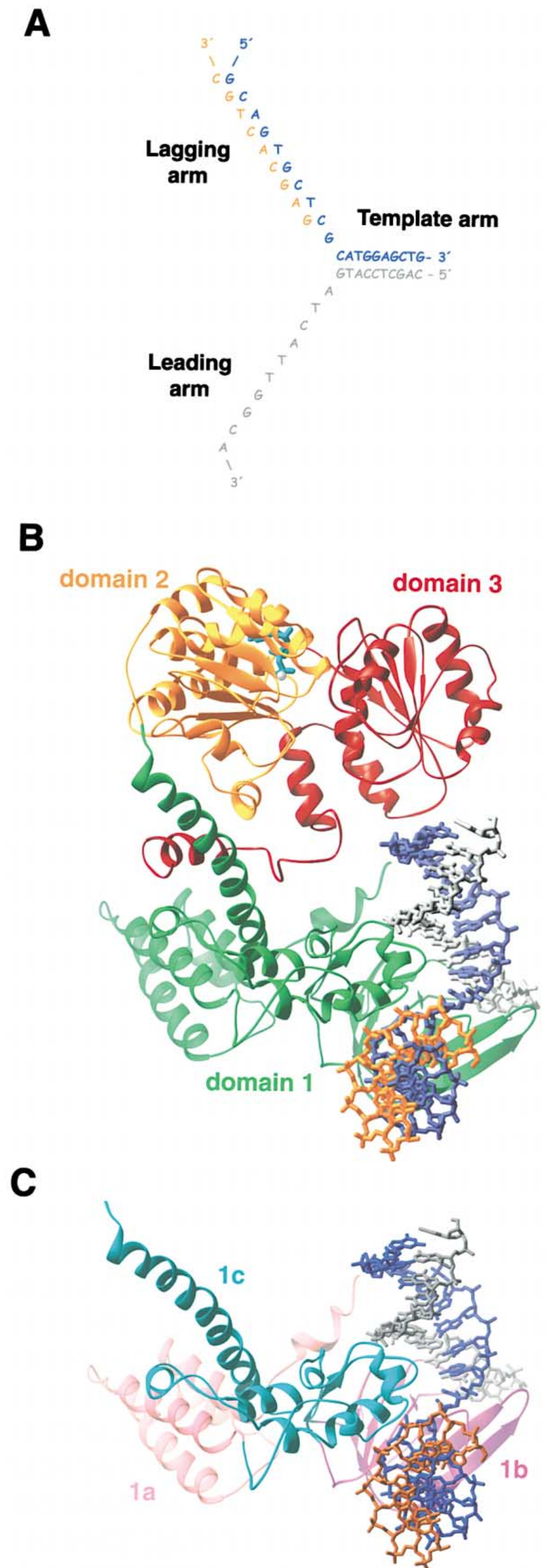

Figure 2. The Architecture of the RecG Protein:DNA Complex

(A) Stalled replication fork junction used for crystallization. This substrate has a single-stranded leading strand arm, mimicking a fork in which DNA synthesis on the leading strand has terminated prematurely. Forks of this kind are the preferred substrate for RecG (McGlynn and Lloyd, 2001). The three oligonucleotides that comprise the fork are labeled chain A (template-leading strand, colored gray), chain B (template-lagging strand, colored blue), and chain C (nascent lagging strand, colored orange) in Figure 3, and are numbered from the 5′ end. Unless otherwise stated, this nomenclature and color scheme is retained throughout the text and figures.

(B) The protein comprises three domains shown in a ribbon representation, with the bound DNA fork in stick representation. The bound ADP is shown in cyan, and the bound magnesium ion as a silver sphere. This figure and several others were prepared using RIBBONS (Carson, 1991).

(C) Structural motifs of Domain 1. The N-terminal helical domain (1a) is colored pink, the wedge domain (1b) is purple and the remainder of the domain (1c) is in cyan.

tals was that shown in Figure 2. To simplify discussion, the nomenclature used to describe the junction in Figure 2 will be used throughout. This junction has a ten base pair duplex template arm ahead of the fork, a one base gap at the junction on the lagging arm followed by nine base pairs of duplex, and a ten base single-stranded leading arm. Forks with a single-stranded leading arm have been shown to be the preferred substrate for RecG in vitro (McGlynn and Lloyd, 2001). This substrate mimics a stalled replication fork in which leading strand synthesis is behind that of the lagging strand, a situation that is likely to occur if the polymerase hits a base lesion on the leading strand template (McGlynn and Lloyd, 2000). Although DNA synthesis on the leading and lagging strands is thought to be tightly coupled in *E. coli*, RecG will unwind both leading and lagging stalled forks in vitro, but with a distinct preference for unwinding the lagging strand arm. The structure discussed below is of RecG in a complex with ADP and the leading strand stalled fork shown in Figure 2.

Overall Fold of RecG

RecG comprises three structural domains (Figure 2). The largest domain (Domain 1, residues 1–350) is at the N terminus and consists of about half of the protein. Consistent with biochemical studies (Mahdi et al., 1997), this domain makes the most extensive interaction with the DNA junction (see below). A significant feature in this domain is a long α helix that runs diagonally across the center of the domain and appears to provide a rigid structural foundation upon which the rest of the domain is folded. The overall fold of this domain is different from that seen in other helicases. There are, however, at least two common structural motifs present within this domain. Residues 21–99 form an antiparallel four helix bundle, preceded by an additional α helix, although this entire region is absent in many RecG sequences (such as *E. coli* RecG) that are shorter than the *T. maritima* protein. Within Domain 1, there is also a greek key motif (residues 154–252) that is conserved in all RecG sequences and which we shall refer to as the "wedge" domain. These residues form a significant part of the surface that contacts the bound DNA (see below). The remainder of the domain is wrapped around the long α helix and appears to have a unique fold as assessed by the DALI server (Holm and Sander, 1993). An alignment of the structural features against the protein sequence

is included as Supplemental Data on the *Cell* website (http://www.cell.com/cgi/content/full/107/1/79/DC1).

The remainder of the protein is split approximately equally between the two C-terminal domains (Domain 2 (residues 351–549) and Domain 3 (residues 550–780)). Domains 2 and 3 contain the characteristic motifs that identify RecG as an SF2 helicase (Gorbalenya and Koonin, 1993). Consequently, this part of the protein has a similar structure to the equivalent domains in other SF2 helicases such as NS3 (Yao et al., 1997; Kim et al., 1998), UvrB (Theis et al., 1999; Machius et al., 1999), and Eif4A (Caruthers et al., 2000; Story et al., 2001), which in turn are related to equivalent domains in Superfamily 1 (SF1) helicases such as PcrA (Subramanya et al., 1996) and Rep (Korolev et al., 1997). We therefore refer to these as the "helicase domains." One point of interest in this region is the relative orientation of these domains, which has been shown to alter when ATP binds in the cleft between these domains in the SF1 helicases (Velankar et al., 1999), a mechanism that is thought to be similar in SF2 helicases (Kim et al., 1998). This region is certainly flexible in both SF1 and SF2 helicases, and several different relative orientations of these domains have been observed in different crystal structures (Subramanya et al., 1996; Korolev et al., 1997; Yao et al., 1997; Velankar et al., 1999; Story et al., 2001). The C-terminal 50 or so residues of the protein extend from the end of Domain 3 and cross back to Domain 1, forming a hook that wraps around the extended α helix. This interaction provides a link between Domains 1 and 3 that is likely to be affected by nucleotide binding.

Consistent with biochemical data (McGlynn et al., 2000), RecG is a monomer in the crystals, in common with other crystal structures of both SF1 and SF2 helicases in a variety of liganded states (Subramanya et al., 1996; Korolev et al., 1997; Velankar et al., 1999; Theis et al., 1999; Machius et al., 1999). The only exception is the *M. jannaschi* Eif4A, in which crystal contacts are proposed to be a dimer interface (Story et al., 2001), although this interface is not retained in yeast Eif4A (Caruthers et al., 2000). Monomeric helicases likely utilize an "inchworm" rather than an "active rolling" mechanism (Bird et al., 1998a; Velankar et al., 1999).

Interactions with the Bound Replication Fork

The site of interaction between the protein and DNA is located primarily on Domain 1 (Figure 3). The most intimate region of contact between the protein and DNA is at the junction itself, indicating how this protein interacts specifically with stalled replication forks. The interaction is mediated on one side of the junction largely through the wedge domain, and on the other by an extended β hairpin (residues 259–266). The junction is gripped between these structural features, which serve to split the two duplex arms of the fork simultaneously. The template strands run along either side of the wedge domain in grooves that are too narrow to accommodate a DNA duplex. Therefore, both the lagging and leading duplex arms of a junction would be split across the wedge domain as the template strands run into these grooves.

On the template duplex arm, there are nine base pairs of DNA that are in regular B-form conformation. At the tenth base pair, the junction has been split open so that the single-stranded lagging strand arm of the junction begins one base earlier than expected. Therefore, the structure has caught the enzyme in an initial act of strand displacement, revealing some of the details of this process. The breaking of the duplex is stabilized in a number of ways. The orphan base of the template arm (a10) is sandwiched between the ninth base (a9) on one side and Phe204 from the protein on the other. The protein therefore substitutes for the base stacking that this base would experience in a duplex by capping the end with an aromatic residue. The partner to this tenth base is flipped out of the duplex at the center of the junction. This base (b11) appears to have been caught in the middle of flipping from one duplex arm of the fork to another. There are no contacts to stabilize this base, presumably because this needs to be a transient state to facilitate flipping between the duplexes during junction migration. Although we have crystallized the protein with a fork rather than a four-way junction, it is evident from the structure that the displaced strands on each arm of the fork would be suitably disposed to allow them to associate and form the fourth arm of the chicken foot as they emerged from the surface of the wedge domain (discussed below).

On moving from the junction down along the lagging strand arm of the fork, there is a very similar interaction to that described above for the template arm (Figure 3). The junction that we used to crystallize the complex lacks a base on the nascent strand at the junction, thereby creating an orphan base. This base (b10) is sandwiched between the end of the lagging strand duplex and Tyr208 of the protein, again mimicking the contacts that the base would experience in a nucleic acid duplex as observed in the template arm. Thus, the fork is bound with an internal symmetry of contacts to stabilize the junction as it is split across the surface of the protein.

The phosphodiester backbones of the duplex arms of the junction are also contacted by the protein. Although both strands of the lagging strand arm are contacted, extending to the fourth base pair counting away from the junction, the contacts are such that either a B-form (DNA) or A-form (e.g., DNA/RNA hybrid) duplex could probably be accommodated. This might explain why RecG can unwind substrates with either DNA or RNA strands (e.g., Holliday junctions and R loops). The template arm duplex ahead of the fork extends across Domain 1 and onto the surface of Domain 3. If extended further than in the present structure, the template arm would run into a part of Domain 3, although a relatively small reorientation of the DNA would allow a longer arm to run across the surface of the domain. In a real fork, of course, the template arm would be a much longer section of duplex DNA, which is likely to make additional contacts with the protein beyond those we observe in the present crystal structure. We are therefore cautious in interpreting details of the interaction between the template arm and Domain 3 at this stage.

There is, however, one other crystal structure of a SF2 helicase family member in a complex with nucleic acid. The hepatitis C virus NS3 RNA helicase has been crystallized with an eight base single-stranded deoxyuridylate oligonucleotide (Kim et al., 1998), providing some information about the interaction between this protein and nucleic acid. In the structure, the DNA binds in a groove

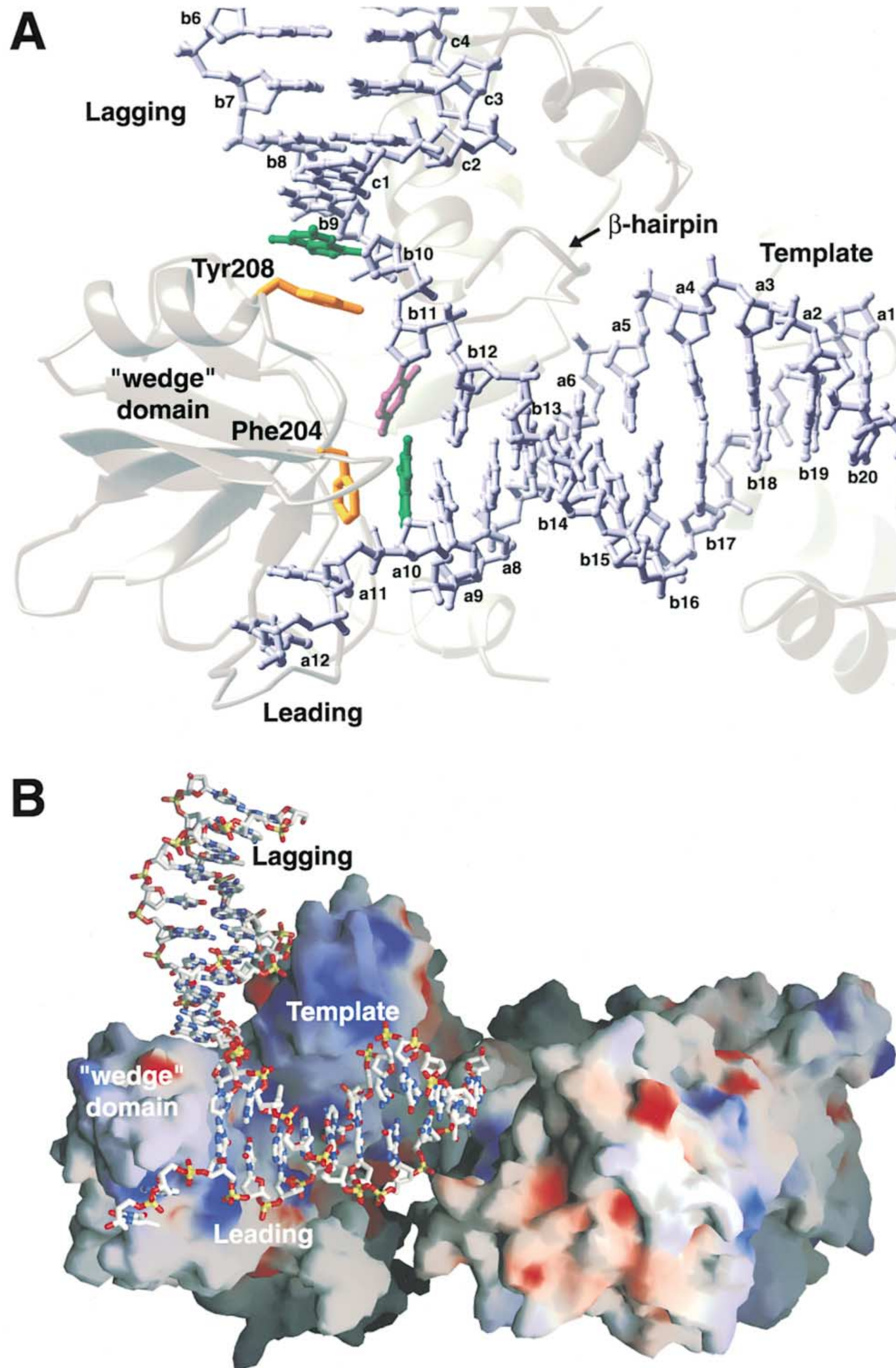

Figure 3. Interaction between the Protein and the Junction

(A) Details of the interaction between the protein (silver) and DNA (pale blue) showing how aromatic interactions (Phe204 and Tyr208, colored gold) stabilize the orphan bases (a10 and b10, colored green) at the junction. The flipped out base (b11) is shown in magenta.

(B) Surface representation of the interaction with the DNA substrate in the same orientation as (A) illustrating how the fork is split across the surface of the wedge domain. Positive potential on the surface is colored blue and negative potential in red. The DNA is shown overlaid in atom colors in stick representation. This figure was prepared using GRASP (Nicholls and Honig, 1991).

between the surface of the helicase domains (equivalent to Domains 2 and 3 of RecG) and the C-terminal domain for which there is no equivalent in RecG. Interestingly, the conformation of this single-stranded DNA is very similar to that observed for each strand within regular B-form duplex DNA. The region of RecG that would be

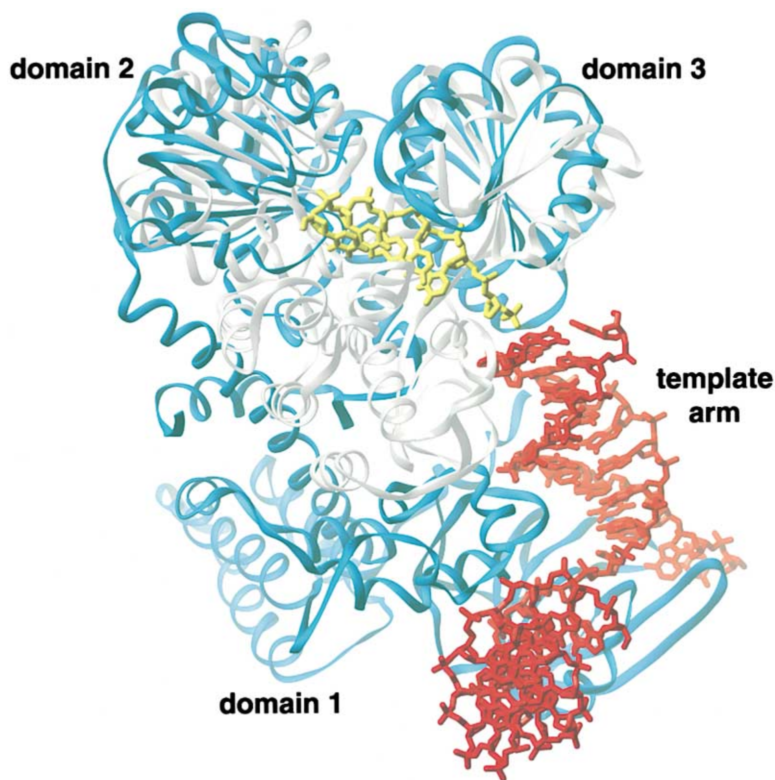

Figure 4. Model for DNA Duplex Binding

Superposition of the NS3 helicase (silver) with its bound ssDNA (yellow), and RecG (blue) with the bound junction (red). A longer template arm of the fork would extend across the region occupied by the DNA in the NS3 structure.

the equivalent DNA binding site lies at the center of the structure as viewed in Figure 2, and a superposition of the two structures is shown in Figure 4. This superposition reveals that a longer template strand arm could bind across this site. However, an important difference is that because RecG lacks the C-terminal domain present in NS3 protein, the groove in NS3 becomes an open surface in RecG, providing sufficient room for a duplex to bind. Furthermore, because RecG is missing this C-terminal domain, several contacts that are important for the interaction of NS3 helicase with ssDNA are absent in RecG, particularly the critical "bookend" tryptophan and valine residues (Kim et al., 1998). Consequently, the mechanism for DNA translocation in RecG cannot be the same as that proposed for NS3.

The single-stranded portion of the DNA corresponding to the leading strand behind the fork runs along a groove along the side of the wedge domain before extending away from the surface of the protein. Only the first two or three bases are contacted by the protein, and even if the arm were duplex rather than single-stranded, the contacts with the protein would appear to be very limited. This observation explains why RecG is able to accommodate substrates with either single-stranded or duplex DNA in the leading arm of the fork, and why RecG is able to unwind both leading and lagging strand stalled forks, albeit with a preference for forks stalled on the leading strand (McGlynn and Lloyd, 2001).

Interestingly, recent biochemical data (McGlynn and Lloyd, 2001) have revealed a requirement for a correlation between the length of the leading and lagging arms in order to allow displacement of the nascent strand from the lagging arm. By using different fork substrates in which the leading and lagging strand arm lengths were altered, the highest activity was observed for substrates in which the lengths of the arms were more closely matched. Although complicated by issues of processivity, these preliminary data do at least suggest that the enzyme requires substrate arms of similar lengths to enable efficient strand displacement. This proposal would certainly be consistent with the structure, because it would seem that for efficient binding and translocation, both arms would have to be bound onto the surface of the protein to facilitate splitting of the junction across the wedge domain. Without these interactions, the torsional strain which must arise during translocation along the template arm could force the substrate to lift away from the surface of the protein.

Previous biochemical data (Mahdi et al., 1997) have demonstrated the importance of the N-terminal region of RecG for specific binding of junctions. Truncation mutants in which either the first 60 or 144 residues of the *E. coli* RecG protein were absent failed to bind junctions. By contrast, C-terminal truncations (Mahdi et al., 1997) and point mutants that have disrupted ATPase activity (McGlynn et al., 2000) bind junctions with the same avidity as wild-type protein. The C-terminally truncated proteins are, however, defective helicases, indicating a role for these residues in catalysis. From sequence analysis (not shown), these four deletions of *E. coli* RecG correspond to residues 1–166 or 1–250 at the N terminus and 757–780 or 742–780 at the C terminus of the *T. maritima* protein, due to the additional residues at the N terminus of *T. maritima* RecG. Deletion of the first 166 or 250 residues of the protein would severely

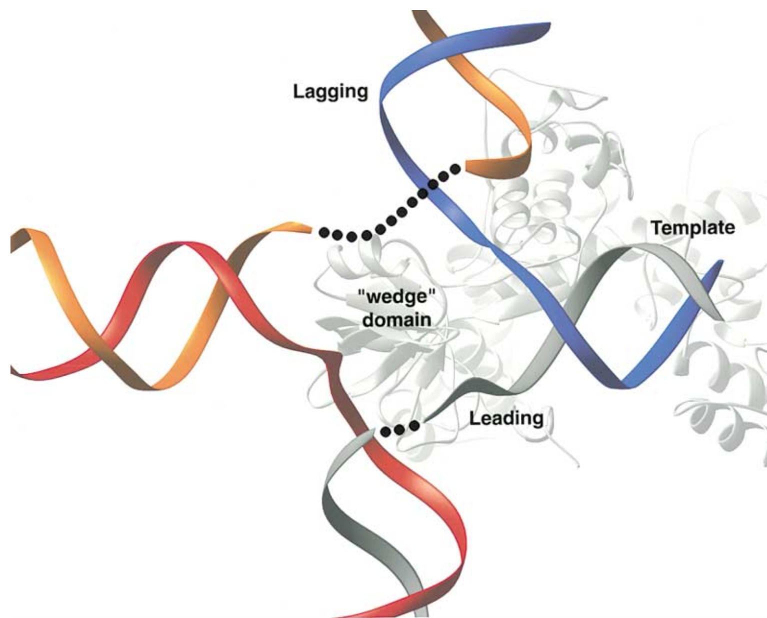

Figure 5. Holliday Junction Formation

Model illustrating how a Holliday junction might form during catalysis by the enzyme. The Holliday junction DNA is colored using the same scheme as in Figure 6. The view is the same as that in Figure 3.

disrupt DNA binding by removing a significant portion of the surface that interacts with the junction, including a part or all of the wedge domain. The role of the C-terminal 40 or so residues is more enigmatic. These residues cross back from Domain 3 and make contacts with Domain 1. However, much of this region is poorly ordered in the structure, particularly the last 20 or so residues. Precisely how these residues contribute to helicase activity is not clear, but it is likely that they are involved in conformational changes associated with domain movements arising from ATP binding and hydrolysis.

Activity of RecG on Different Substrates

The two duplex arms of the junction are arranged at an angle close to 90° (Figure 3). The product of reversal of stalled forks by RecG is a four-way Holliday junction (McGlynn and Lloyd, 2000). To date, there have been crystal structures of several different proteins bound to Holliday junctions, including RuvA (Hargreaves et al., 1998; Roe et al., 1998; Ariyoshi et al., 2000) as well as the Cre and Flp recombinases (Guo et al., 1997; Chen et al., 2000). In all of these cases, the conformation of the junction when bound to protein has been square planar rather than the stacked X structure that is thought to be more stable in solution in the presence of physiological concentrations of magnesium ions (Duckett et al., 1990). It has been shown that magnesium ions have a drastic effect upon binding of RecG to Holliday junctions, which was interpreted to be due to altering the conformation of the junction from a planar to stacked X structure (Whitby and Lloyd, 1998). It is probably not a coincidence, therefore, that the two duplex arms of the DNA bound to RecG superimpose very well with the square planar conformation of a Holliday junction, suggesting that it would be easy to bind two of the arms of a four-way junction in a manner similar to that observed for this fork (Figure 5). The conformation of the fork is stabilized by a combination of the interactions at the junction and the interactions with the arms of the fork. In this way the enzyme:DNA complex is set up appropriately to facilitate formation of four-way junctions. To reverse the reaction (i.e., fork regression), RecG would simply have to bind the junction in a different manner, so that either the leading or lagging arm would be situated in the position occupied by the template strand in our structure.

Assignment of the physiological role of RecG has been complicated by the plethora of substrates that can be acted upon by the enzyme in vitro. In addition to reversing leading strand stalled forks, RecG can also reverse lagging strand stalled forks (albeit less effectively), as well as unwind Holliday junctions, R loops, and D loops. The promiscuity of the protein can be explained by the crystal structure, because the central core of each of these substrates is equivalent and probably interacts with RecG in a manner similar to that observed for the case of the lagging strand stalled fork (Figure 6).

A Mechanism for Junction Unwinding by RecG

There has been a considerable development of our understanding of helicases in recent years. In terms of the mechanism of unwinding, perhaps one of the best understood enzymes is the SF1 helicase, PcrA. Crystal structures of the protein alone and in two different DNA complexes (Subramanya et al., 1996; Velankar et al., 1999) suggested an inchworm mechanism akin to that proposed 20 years previously (Yarranton and Gefter, 1979), but with some modifications. In this model, PcrA unwinds duplex DNA at a junction and then translocates along one of the resulting single strands of DNA. There were two important corollaries for this proposed mechanism. The first of these was that the enzyme would consume one ATP molecule for every base on the DNA along which it translocated. Using pre-steady-state kinetic analysis of phosphate release, it was shown that PcrA does indeed hydrolyze one ATP per base during translocation along single-stranded DNA (Dillingham et

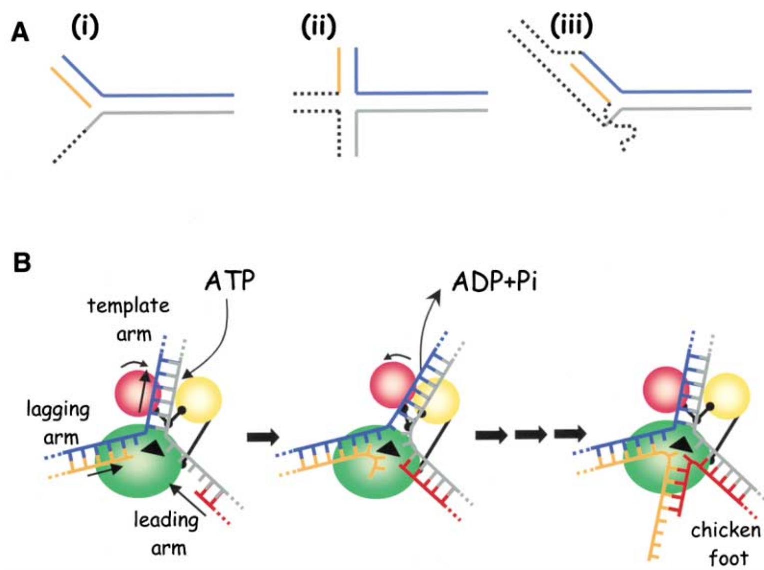

Figure 6. Mechanism for Fork Unwinding by RecG

(A) Different DNA substrates that are unwound by RecG, with the common core of the substrates shown using the colors in Figure 2, and the remainder of the substrate shown in dotted lines. The substrates are (i) stalled fork, (ii) Holliday junction, and (iii) R loop.

(B) Domain 1 is colored green, Domain 2 is yellow, and Domain 3 is magenta. The template strands of the DNA fork are colored blue and gray, with the nascent strands colored orange (lagging strand) and red (leading strand). The wedge domain is represented by a black triangle. The leading strand stalled junction is bound to the protein prior to the binding of ATP. When ATP binds, the cleft between Domains 2 and 3 closes, inducing a conformational change that pulls on the template arm, splitting the duplex on the lagging arm across the wedge domain. When ATP is hydrolyzed, the cleft between Domains 2 and 3 reopens, and Domain 3 slides along the template arm, which is held onto Domain 1 by the tight binding at the junction. The cycles are repeated multiple times until the leading strand duplex is encountered, at which point the two displaced nascent strands can associate to form a chicken foot structure.

al., 2000). The second implication of the model was that the enzyme distorted the DNA duplex ahead of the junction, thus facilitating fork progression. This point was demonstrated by a combination of nuclease protection, DNA footprinting, and site-specific mutagenesis (Soultanas et al., 2000). Thus, the model has been substantiated at both the structural and biochemical levels. However, it is not clear whether this mechanism applies to other DNA helicases, in particular those of other helicase superfamilies. One helicase superfamily is the group of enzymes that are related to the bacterial replicative helicase, DnaB. These enzymes form hexameric rings of RecA-like domains with a nucleotide binding site situated at the interface between the subunits (Singleton et al., 2000). Biochemical data (reviewed in Patel and Pichia, 2000) demonstrate extreme negative cooperativity in nucleotide binding sites such that only a subset of the six potential nucleotide binding sites contains bound nucleotide at any one time. Although there are several similarities between the enzymes, in the chemistry of nucleotide hydrolysis for example, it is evident that the hexameric helicases cannot utilize the same mechanism as PcrA for unwinding DNA duplexes.

For SF2 helicases the situation is even less clear. Crystal structures of members of this family reveal a structure that is closely related to SF1 helicases with two RecA-like domains, although with a connectivity that is like the nucleotide binding domain of adenylate kinase (Bird et al., 1998a). In common with SF1 helicases, the nucleotide binding site in SF2 enzymes has been shown to be in the cleft between these domains (Theis et al., 1999) and, although there is still no direct structural evidence, it has been proposed that cleft closure might be associated with nucleotide binding in a manner analogous to that demonstrated for PcrA, and that this conformational change would drive helicase activity (Kim et al., 1998; Story et al., 2001). However, there are interesting differences between SF1 and SF2 helicases in regard to their NTPase properties. The NTPase activity of SF1 helicases shows a marked dependence upon ssDNA such that the rate is stimulated by around three orders of magnitude by binding of ssDNA (Bird et al., 1998b). By contrast, it has been shown that this stimulation of NTPase activity in SF2 helicases is less marked (Preugschat et al., 1996; Whitby and Lloyd, 1998). Furthermore, the stimulation of RecG, for example, is dependent upon dsDNA rather than ssDNA (Whitby and Lloyd, 1998). Since the stimulation of the NTPase activity of PcrA has been shown to be greatest when associated with translocation along ssDNA (Dillingham et al., 2000), this difference between the enzymes raises an interesting possibility, namely that RecG might translocate along dsDNA rather than along ssDNA. This proposal allows us to suggest a model for fork unwinding by RecG.

It is has been shown that RecG unwinds both the leading and lagging strand duplex arms of a three-way junction and that the unwinding of these arms appears to be coordinated (McGlynn and Lloyd, 2001). The crystal structure of RecG that we present here suggests a simple mechanism for this process (Figure 6). RecG binds initially to the junction with the arms of the fork disposed around the wedge domain of the protein. The structure shows how the base pairs at the junction are split across the protein and stabilized by aromatic interactions with the orphan bases. It therefore seems that in order to unwind both arms simultaneously, the protein would simply have to pull on the template arm, thereby dragging the junction across the wedge domain. As described above, it is not possible for a DNA duplex to pass though the junction binding site; only the template strands on the leading and lagging arms would be able to pass through the grooves on either side of the wedge domain. Consequently, the nascent DNA strand(s) would be stripped off the template by a simple steric interaction and would be displaced to either side, but their proximity would allow association of these complementary strands to form a four-way junction (i.e., chicken foot).

Table 1. Crystallographic Statistics

(a) Data collection				
	Native	Pb	Hg	Se
Resolution (Å)	20–3.25	20–3.5	20–3.5	20–3.25
Completeness (%)	98.3	99.3	99.5	99.0
R_{symm} (%)	3.9	5.9	7.4	5.3
R_{deriv} (%)	—	13.7	21.7	15.2
Number of sites	—	2	2	21
Phasing power	—	2.2	1.4	1.1
Overall mean figure of merit	0.48			
(b) Final model				
Rfactor (%) (All data)	27.5			
Rfree (%) (5% of data)	32.7			
Rmsd bond length (Å)	0.021			
Rmsd bond angle (°)	2.6			

This model would require that RecG be a dsDNA translocase, a proposal that would be consistent with the NTPase kinetics (Whitby and Lloyd, 1998). A simple model for translocation would be driven by opening and closing of the cleft between Domains 2 and 3 as ATP binds and hydrolyzes, in a manner similar to that demonstrated for PcrA, allowing the protein alternately to bind and release the template duplex region, thereby walking along the DNA. The interaction between the protein and the duplex need not involve both strands. Indeed, the 3′-5′ polarity that has been demonstrated for RecG suggests that the major contacts will be with only one of the strands.

This proposed mechanism has parallels with that proposed for the RuvAB complex (West, 1996). In RuvAB, the two RuvB protein rings are thought to act as double-strand DNA translocation motors that each pull one arm of the Holliday junction in order to split the junction across the RuvA protein tetramer. Interestingly, the DNA junction is split across a domain of the RuvA protein that has the same fold as that of the wedge domain of RecG, although the acidic pin region of RuvA that is responsible for splitting the junction (Rafferty et al., 1996) is not present in RecG. For the RecG:junction complex, a single motor pulls one arm of the junction, which is then split across one RuvA-like domain, albeit in a slightly different manner to that used by RuvAB. However, there is one potential difference between these systems that might be of biological significance. Since hexameric ring helicases, such as RuvAB, are generally more processive than monomeric enzymes such as RecG, this raises the possibility that RecG and RuvAB might differ in their respective processivities. This difference might be important in relation to their respective roles in the processing of stalled replication forks. RecG, for example, may only be required to reverse a stalled fork a short distance beyond the site of DNA damage to provide a primer for template switching, whereas RuvAB might be required to migrate Holliday junctions over greater distances and/or through regions of heterology.

The first crystal structure of a helicase (Subramanya et al., 1996) unexpectedly revealed a tandem repeat of RecA-like domains with the ATP binding site situated in a cleft between them. A series of seven conserved sequence motifs thought to be characteristic of helicases (Gorbalenya and Koonin, 1993) were all located within these domains. This initial observation led to the suggestion that helicases might be modular, with the two RecA-like domains being the core structure and conferring helicase activity to which other domains might be added to provide specificity for different nucleic acid substrates (Subramanya et al., 1996; Bird et al., 1998a). Subsequent crystal structures appear to have confirmed this idea (Yao et al., 1997; Korolev et al., 1997; Theis et al., 1999; Machius et al., 1999; Caruthers et al., 2000; Singleton et al., 2000). However, biochemical data have revealed that although the basic premise seems to be correct, the situation is a little more complex, in that not all proteins that contain the so-called "helicase motifs" are in fact helicases. There are now several examples of these proteins which show little or no helicase activity in vitro. Furthermore, it has been shown that helicase activity can be severely disrupted by point mutations in regions outside of the helicase domains (Soultanas et al., 2000). However, these mutant proteins are able to translocate along single-stranded DNA as proficiently as the wild-type enzyme. This latter observation led to the proposal that the "helicase" motifs should more appropriately be referred to as "translocase" domains. Thus, the modularity of these proteins couples a nucleic acid translocation motor function to other domains that provide specificity for different forms of nucleic acids but which may also contribute to the helicase activity itself by destabilizing the duplex(es). The structure that we present here reveals another extension of this family. In RecG, we see a conserved structure for the translocase domains, to which is attached a novel domain that confers both a specificity for stalled replication forks and a mechanism for splitting two duplexes simultaneously.

Our proposal that RecG might be a double-strand translocase raises the question whether other SF2 helicases might also be dsDNA (or dsRNA) translocases. It is evident that many of the family members would be able to perform their tasks without needing to separate the strands of the duplex, at least in a permanent fashion (e.g., type I restriction enzymes, UvrB, Swi/Snf2), and it may be that this is one functional reason behind the evolution of these two helicase families.

The structure presented here provides detailed information about the recognition of stalled fork structures by RecG, but there are still many unanswered questions. Although we propose a general model for how RecG unwinds junctions, details about the mechanism, such as the step size (i.e., how many base pairs of DNA are unwound by RecG for each ATP that is hydrolyzed), remain to be elucidated. Future biochemical and structural work should help to clarify these issues. Although there is no sequence homolog of RecG in eukaryotes, several proteins have been proposed to play a role in the recovery of stalled replication forks such as BLM, Werner's, and Sgs1 helicases (Frei and Gasser, 2001). It is likely that many aspects of the structure we present here will be a paradigm for these eukaryotic enzymes.

Experimental Procedures

The gene encoding *T. maritima* RecG protein was cloned by PCR from genomic DNA and overexpressed in pET28a in *E. coli* B834

pLysS cells carrying a plasmid (pSJS1240, gift from S.J. Sandler) to express low abundance tRNAs (Del Tito et al., 1995). Protein was purified using HiTrap heparin-Sepharose (Pharmacia), isopropyl-Source (Pharmacia), and F3GA-blue Sepharose. Prior to crystallization, protein was concentrated to 15 mg/ml in 1 mM DTT, 200 mM NaCl, and 10 mM Tris.HCL [pH 7.5], and DNA was added to a molar ratio of 1:1.2. Crystals were grown in hanging drops from conditions of 500 mM KH_2PO_4/K_2HPO_4 [pH 5.0], 5 mM $MgCl_2$, and 1 mM ADP, and were of the monoclinic spacegroup C2 with unit cell dimensions a = 133.7 Å, b = 144.6 Å, c = 84.0 Å, β = 113.8°. There is one protein:DNA complex in the asymmetric unit. Data were collected from flash frozen crystals at 100 K on ESRF beamlines 14.1 and 14.3, and processed using the HKL programs (Otwinowski and Minor, 1997) (Table 1). The structure was solved using multiple isomorphous replacement with the derivatives described in Table 1. Unless otherwise stated, the CCP4 program suite was used for the structure determination and subsequent manipulations (CCP4, 1994). Initial phasing from the heavy atom derivatives was undertaken with SHARP (La Fortelle and Bricogne, 1997), followed by solvent flattening with DM or SOLOMON. Model building was undertaken using TurboFrodo (Roussel and Cambillau, 1989). The positions of 21 methionine residues and 2 cysteine residues were based upon the observed heavy atom binding sites and were of considerable help in determining the overall fold and register of the sequence of the protein. Model refinement was undertaken with a combination of CNS (Brunger et al., 1998) and REFMAC interspersed with rounds of model building. Statistics concerning the quality of the final model are presented in Table 1.

Acknowledgments

We thank N. Raven for providing *T. maritima* cells, S. Sandler for kindly providing plasmid pSJS1240, S. Gamblin, S. Smerdon, and P. Walker for advice on improved freezing of crystals, R. Lloyd, K. Marians, and P. McGlynn for communicating and discussing data in advance of publication, S. West for useful discussions and critical reading of the manuscript, D. Hall for assistance on synchrotron visits, N. Cook and K. Mouw for technical assistance, and the ESRF, Grenoble for data collection facilities. This work was supported initially by the Wellcome Trust and subsequently by the Imperial Cancer Research Fund.

Received June 21, 2001; revised August 22, 2001.

References

Ariyoshi, M., Nishino, T., Iwasaki, H., Shinagawa, H., and Morikawa, K. (2000). Crystal structure of the Holliday junction DNA in complex with a single RuvA tetramer. Proc. Natl. Acad. Sci. USA *97*, 8257–8262.

Asai, T., Sommer, S., Bailone, A., and Kogoma, T. (1993). Homologous recombination-dependent initiation of DNA replication from DNA damage-inducible origins in Escherichia coli. EMBO J. *12*, 3287–3295.

Bird, L.E., Brannigan, J.A., Subramanya, H.S., and Wigley, D.B. (1998b). Characterisation of *Bacillus stearothermophilus* PcrA helicase: Evidence against an active rolling model. Nucleic Acids Res. *26*, 2686–2693.

Bird, L.E., Subramanya, H.S., and Wigley, D.B. (1998a). Helicases: A unifying structural theme? Curr. Opin. Struct. Biol. *8*, 14–18.

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). Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. *D54*, 904–925.

Cairns, J. (1963). The bacterial chromosome and its manner of replication as seen by autoradiography. J. Mol. Biol. *6*, 208–213.

Carson, M. (1991). Ribbons 2.0. J. App. Crystallogr. *24*, 958–961.

Caruthers, J.M., Johnson, E.R., and McKay, D.B. (2000). Crystal structure of yeast initiation factor 4A, a DEAD-box RNA helicase. Proc. Natl. Acad. Sci. USA *97*, 13080–13085.

Chen, Y., Narendra, U., Iype, L.E., Cox, M.M., and Rice, P.A. (2000). Crystal structure of a Flp recombinase-Holliday junction complex: Assembly of an active oligomer by helix swapping. Mol. Cell. *6*, 885–897.

CCP4 (Collaborative Computing Project 4) (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D *50*, 760–763.

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.

Del Tito, B.J., Jr., Ward, J.M., Hodgson, J., Gershater, C.J., Edwards, H., Wysocki, L.A., Watson, F.A., Sathe, G., and Kane, J.F. (1995). Effects of a minor isoleucyl tRNA on heterologous protein translation in *Escherichia coli*. J. Bacteriol. *177*, 7086–7091.

Dillingham, M.S., Wigley, D.B., and Webb, M.R. (2000). Unidirectional single-stranded DNA translocation by PcrA helicase: Measurement of step size and translocation speed. Biochemistry *39*, 205–212.

Duckett, D.R., Smurchie, A.I., and Lilley, D.M.J. (1990). The role of metal ions in the conformation of the four-way DNA junction. EMBO J. *9*, 583–590.

Dudas, K.C., and Kreuzer, K.N. (2001). UvsW protein regulates bacteriophage T4 origin-dependent replication by unwinding R-loops. Mol. Cell. Biol. *21*, 2706–2715.

Frei, C., and Gasser, S.M. (2001). RecQ-like helicases: the DNA replication checkpoint connection. J. Cell Sci. *113*, 2641–2646.

Gorbalenya, A.E., and Koonin, E.V. (1993). Helicases: amino acid sequence comparisons and structure-function relationships. Curr. Opin. Struct. Biol. *3*, 419–429.

Guo, K., Gopal, D.N., and Van Duyne, G.D. (1997). Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse. Nature *389*, 40–47.

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

Higgins, N.P., Kato, K., and Strauss, B. (1976). A model for replication repair in mammalian cells. J. Mol. Biol. *101*, 417–425.

Holm, L., and Sander, C. (1993). Protein structure comparison by alignment of distance matrices. J. Mol. Biol. *233*, 123–138.

Hong, X., Cadwell, G.W., and Kogoma, T. (1995). Escherichia coli RecG and RecA proteins in R-loop formation. EMBO J. *14*, 2385–2392.

Kim, J.L., Morgenstern, K.A., Griffith, J.P., Dwyer, M.D., Thomson, J.A., Murcko, M.A., Lin, C., and Caron, P.R. (1998). Hepatitis C virus NS3 RNA helicase domain with a bound oligonucleotide: the crystal structure provides insights into the mode of unwinding. Structure *6*, 89–100.

Kornberg, A., and Baker, T.A. (1992). DNA Replication (New York: W.H. Freeman).

Korolev, S., Hsieh, J., Gauss, G.H., Lohman, T.M., and Waksman, G. (1997). Major domain swiveling revealed by the crystal structures of complexes of *E. coli* Rep helicase bound to single-stranded DNA and ADP. Cell *90*, 635–647.

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

Kuzminov, A. (1995). Collapse and repair of replication forks in E. coli. Mol. Micro. *16*, 373–384.

La Fortelle, E. de, and Bricogne, G. (1997). Maximum-likelihood heavy atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. Meth. Enzymol. *276*, 472–494.

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

Machius, M., Henry, L., Planitkar, M., and Deisenhofer, J. (1999). Crystal structure of the DNA nucleotide excision repair enzyme UvrB from *Thermus thermophilus*. Proc. Natl. Acad. Sci. USA *96*, 11717–11722.

Mahdi, A.A., McGlynn, P., Levett, S.D., and Lloyd, R.G. (1997). DNA binding and helicase domains of the *Escherichia coli* recombination protein RecG. Nucleic Acids Res. *25*, 3875–3880.

Marians, K.J. (2000). PriA-directed replication fork restart in Escherichia coli. Trends Biochem. Sci. *25*, 185–189.

McGlynn, P., and Lloyd, R.G. (2000). Modulation of RNA polymerase by (p)ppGpp reveals a RecG-dependent mechanism for replication fork progression. Cell *101*, 35–45.

McGlynn, P., and Lloyd, R.G. (2001). Rescue of stalled replication forks by RecG: simultaneous translocation on the leading and lagging strand templates supports an active DNA unwinding model of fork reversal and Holliday junction formation. Proc. Natl. Acad. Sci. USA *98*, 8227–8234.

McGlynn, P., Mahdi, A.A., and Lloyd, R.G. (2000). Characterisation of the catalytically active form of RecG helicase. Nucleic Acids Res. *28*, 2324–2332.

McGlynn, P., Lloyd, R.G., and Marians, K.J. (2001). Formation of Holliday junctions by regression of nascent DNA in intermediates containing stalled replication forks: RecG stimulates regression even when the DNA is negatively supercoiled. Proc. Natl. Acad. Sci. USA *98*, 8235–8240.

Mosig, G. (1987). The essential role of recombination in phage T4 growth. Annu. Rev. Genet. *21*, 347–371.

Neuwald, A.F., Aravind, L., Spouge, J.L., and Koonin, E.V. (1999). AAA$^+$: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. Genome Res. *9*, 27–43.

Nicholls, A., and Honig, B.J. (1991). A rapid finite-difference algorithm, utilizing successive over-relaxation to solve the Poisson-Boltzmann equation. J. Comput. Chem. *12*, 435–445.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307–326.

Patel, S.S., and Pichia, K.M. (2000). Structure and function of hexameric helicases. Annu. Rev. Biochem. *69*, 651–697.

Postow, L., Ullsperger, C., Keller, R.W., Bustamante, C., Vologodskii, A.V., and Cozzarelli, N.R. (2001). Positive torsional strain causes the formation of a four-way junction at replication forks. J. Biol. Chem. *276*, 2790–2796.

Preugschat, F., Averett, D.R., Clarke, B.E., and Porter, D.J. (1996). A steady-state and pre-steady-state kinetic analysis of the NTPase activity associated with the hepatitis C virus NS3 helicase domain. J. Biol. Chem. *271*, 24449–24457.

Putnam, C.D., Clancy, S.B., Tsuruta, H., Gonzalez, S., Wetmur, J.G., and Tainer, J.A. (2001). Structure and mechanism of the RuvB Holliday junction branch migration motor. J. Mol. Biol. *311*, 297–310.

Rafferty, J.B., Sedelnikova, S.E., Hargreaves, D., Artymuik, P.J., Baker, P.J., Sharples, G.A., 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*, 415–421.

Roe, S.M., Barlow, T., Brown, T., Oram, M., Keeley, A., Tsaneva, I., and Pearl, L.H. (1998). Crystal structure of an octameric RuvA-Holliday junction complex. Mol. Cell *2*, 361–372.

Roussel, A., and Cambillau, C. (1989). TURBO-FRODO. In Silicon Graphics Geometry Partner Directory, Silicon Graphics, ed. (Mountain View, CA: Silicon Graphics), pp. 77–78.

Singleton, M.R., Sawaya, M.R., Ellenberger, T., and Wigley, D.B. (2000). Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides. Cell *101*, 589–600.

Skalka, A. (1974). A replicator's view of recombination (and repair). In Mechanisms in Recombination, R.F. Grell, ed. (New York: Plenum Press), pp. 421–432.

Soultanas, P., Dillingham, M.S., Wiley, P., Webb, M.R., and Wigley, D.B. (2000). Uncoupling DNA translocation and helicase activity in PcrA: direct evidence for an active mechanism. EMBO J. *19*, 3799–3810.

Story, R.M., Li, H., and Abelson, J.N. (2001). Crystal structure of a DEAD box protein from the hyperthermophile *Methanococcus jannaschii*. Proc. Natl. Acad. Sci. USA *98*, 1465–1470.

Subramanya, H.S., Bird, L.E., Brannigan, J.A., and Wigley, D.B. (1996). Crystal structure of a DExx box helicase. Nature *384*, 379–383.

Theis, K., Chen, P.J., Skorvaga, M., Van Houten, B., and Kisker, C. (1999). Crystal structure of UvrB, a DNA helicase adapted for nucleotide excision repair. EMBO J. *18*, 6899–6907.

Velankar, S.S., Soultanas, P., Dillingham, M.S., Subramanya, H.S., and Wigley, D.B. (1999). Crystal structures of complexes of PcrA helicase with a DNA substrate indicate an inchworm mechanism. Cell *97*, 75–84.

Vincent, S.D., Mahdi, A.A., and Lloyd, R.G. (1996). The RecG branch migration protein of Escherichia coli dissociates R-loops. J. Mol. Biol. *264*, 713–721.

West, S.C. (1996). The RuvABC proteins and Holliday junction processing in E. coli. J. Bacteriol. *178*, 1237–1241.

Whitby, M.C., and Lloyd, R.G. (1998). Targeting Holliday junctions by the RecG branch migration protein of *Escherichia coli*. J. Biol. Chem. *31*, 19729–19739.

Whitby, M.C., Ryder, L., and Lloyd, R.G. (1993). Reverse branch migration of Holliday junctions by RecG protein: a new mechanism for resolution of intermediates in recombination and DNA repair. Cell *75*, 341–350.

Yamada, K., Kunishima, N., Mayanagi, K., Ohnishi, T., Nishino, T., Iwasaki, H., Shinagawa, H., and Morikawa, K. (2001). Crystal structure of the Holliday junction migration protein RuvB from *Thermus thermophilus* HB8. Proc. Natl. Acad. Sci. USA *98*, 1442–1447.

Yao, N.H., Hesson, T., Cable, M., Hong, Z., Kwong, A.D., Le, H.V., and Weber, P.C. (1997). Structure of the hepatitis C virus RNA helicase domain. Nat. Struct. Biol. *4*, 463–467.

Yarranton, G.T., and Gefter, M.L. (1979). Enzyme-catalyzed DNA unwinding: studies on *Escherichia coli* rep protein. Proc. Natl. Acad. Sci. USA *76*, 1658–1662.

Accession Numbers

Coordinates have been deposited at the RCSB database with accession code 1GM5.