

# DNA binding and helicase domains of the *Escherichia coli* recombination protein RecG

Akeel A. Mahdi, Peter McGlynn, Stuart D. Levett and Robert G. Lloyd\*

Department of Genetics, University of Nottingham, Queens Medical Centre, Nottingham NG7 2UH, UK

Received June 20, 1997; Revised and Accepted August 8, 1997

## ABSTRACT

The *Escherichia coli* RecG protein is a unique junction-specific helicase involved in DNA repair and recombination. The C-terminus of RecG contains motifs conserved throughout a wide range of DNA and RNA helicases and it is thought that this C-terminal half of RecG contains the helicase active site. However, the regions of RecG which confer junction DNA specificity are unknown. To begin to assign structure–function relationships within RecG, a series of N- and C-terminal deletions have been engineered into the protein, together with an N-terminal histidine tag fusion peptide for purification purposes. Junction DNA binding, unwinding and ATP hydrolysis were disrupted by mutagenesis of the N-terminus. In contrast, C-terminal deletions moderately reduced junction DNA binding but almost abolished unwinding. These data suggest that the C-terminus does contain the helicase active site whereas the N-terminus confers junction DNA specificity.

## INTRODUCTION

Studies of DNA repair and recombination in *Escherichia coli* have identified two systems, RecG and RuvABC, required for branch migration and resolution of Holliday structures formed between homologous duplexes. The RuvABC system has been extensively characterized and the structures of RuvA and C have been solved at atomic resolution (1,2). A heteromultimeric RuvAB complex acts as a specialized DNA helicase and catalyses ATP-dependent branch migration of the junction (3,4), whilst a RuvC dimer introduces two symmetrical nicks within two strands of the junction DNA, possibly within a RuvABC complex (5–7). This results in formation of recombinant duplex products, with the precise nature of the DNA exchange being dependent upon the extent of migration of the Holliday junction along the homologous duplexes and orientation of the cleaved DNA strands.

The RecG protein shares some functional overlap with RuvAB, since it is also a branch-specific DNA helicase (8,9). However, single mutations in the *recG* or *ruvABC* genes cause defects in DNA repair and recombination and therefore the systems are not redundant (10). *In vitro* analyses of the two systems have also demonstrated the very different nature of these specialized helicases. Whereas tetrameric RuvA binds junction DNA and

hexameric RuvB rings catalyse branch migration of the bound DNA (4,11), RecG can catalyse branch migration without any accessory proteins (8,12). Thus RecG fulfills the roles of both RuvA and RuvB. There is also evidence that RecG is targeted to early intermediates of recombination formed by RecA. RecG can counter the formation of such intermediates in the presence of RecA (12) and RecG can also specifically bind to and unwind D loops (13). *In vivo* RecG also functionally interacts with PriA, a protein essential for recombination primed by DNA replication from D loops (14). R loops, formed by assimilation of homologous RNA into duplex DNA, are also unwound by RecG, but not by RuvAB (15,16). RecG therefore possesses multiple catalytic functions in addition to branch migration of DNA. These studies have led to models of RuvAB and RecG action. RecG has been postulated to be targeted to the initial DNA intermediates formed by RecA, possibly driving branch migration of three strand DNA intermediates formed at D loops into regions of duplex–duplex pairing to form Holliday structures (14,17). These structures could then be branch migrated and resolved by RuvABC.

Analysis of the sequence of RecG has identified seven motifs within the C-terminal half of the protein that are conserved amongst many DNA and RNA helicases (18,19). These conserved regions have been the subject of extensive mutagenesis studies in a wide range of helicases. Motif I contains the ‘Walker A’ motif responsible for binding the triphosphate tail of ATP (20,21) and putative functions have been assigned to the other regions. The importance of these motifs in RecG helicase function was demonstrated by a mutation in motif III which, whilst not affecting junction DNA binding, abolished all helicase activity (22). The crystal structure of a DNA helicase has also helped to suggest functions for the conserved helicase sequences and how these sequences interact to promote DNA unwinding (23). All seven motifs are clustered around a cleft formed between two structurally similar domains. ADP has been shown to bind within this cleft and it has been postulated that duplex DNA also binds within this negatively charged groove. ATP binding or hydrolysis might therefore affect the conformation of both domains and this may catalyse unwinding of the bound duplex DNA. Although DNA and RNA helicases are a very diverse group of enzymes with regard to sequence, quaternary structure and substrate specificity, the seven conserved motifs suggest that such helicases may share the same basic fold of the two related domains. Other, diverse sequences attached to this conserved structure may give the individual enzymes different substrate specificities and activities (23).

\*To whom correspondence should be addressed. Tel: +44 115 9709406; Fax: +44 115 9709906; Email: bob.lloyd@nottingham.ac.uk

Despite RecG being one of only two junction-specific helicases identified so far, there is little information available on structure–function relationships within this multifunctional enzyme. The conserved helicase motifs located near the C-terminus are presumed to be essential for the unwinding reaction, but the role of the N-terminal region is unknown. To begin to address these issues a series of specific deletions have been engineered into RecG and their effects on the catalytic activity of the protein have been assayed.

## MATERIALS AND METHODS

### Introduction of restriction sites into *recG*

*NdeI* and *HindIII* sites had previously been introduced upstream and downstream respectively of *recG* to form pGS772 (8). This plasmid allowed overexpression of *recG* using IPTG-inducible expression of the T7  $\phi$ 10 promoter in pT7-7 (24). A further five restriction sites were engineered into *recG* by PCR amplification using primers whose sequence introduced silent mutations into the *recG* coding sequence (Fig. 1). The *SstII* restriction site was made by changing <sub>297</sub>AGCGCGG to TCCGCGG (the number refers to the nucleotide position of the initial base, with the first base of the coding sequence designated 1). The *SalI* site was introduced by altering <sub>667</sub>GTCGTC to GTCGAC. The *KpnI* site was made by changing <sub>1128</sub>GGTACA to GGTACC. The *PstI* site was made by changing <sub>1585</sub>CTGCCG to CTGCAG. The *EcoRI* site was made by changing <sub>1932</sub>GAATTT to GAATTC. Each engineered DNA fragment of *recG* was sequenced to confirm that only the desired changes were introduced. The *recG* clone containing all of the engineered restriction sites was subcloned into pT7-7 to form pAM210. This plasmid was identical to pGS772 except for the five restriction sites above. Copies of *recG* both with and without the above five restriction sites were also cloned into pET-14b (Novagen) to form pAM209 and pAM202 respectively. These allowed overexpression of RecG with an N-terminal histidine tag.

### Construction of truncated *recG* genes

Utilizing the restriction fragments generated above, deletions at either end of *recG* were generated by PCR using a primer specifying the deletion plus one of the primers used to introduce the above restriction sites. For RecG $\Delta$ N60 the oligonucleotide primer used was 5'-GCTCTAGACATATGTCACGGTG-GAAGCGGA-3', which encoded the *NdeI* site encompassing the initiator methionine of RecG followed by codon 61. RecG $\Delta$ N144 was generated with the primer 5'-GCTCTAGACATATGCAGG-AAACGCTCAGCCGGT-3', which encoded the *NdeI* site followed by codon 145. RecG $\Delta$ C32 was generated with 5'-GCG-AAGCTTAAACTTCCGGGATCATCGCCT-3', whose final *recG* codon was number 661. Each DNA fragment generated by PCR was sequenced to confirm that only the required changes were present. RecG $\Delta$ N60 $\Delta$ C32 was constructed by cloning fragments containing each of the individual deletions into the same construct. The truncated versions of *recG* were subcloned as *XbaI*–*HindIII* fragments into both pT7-7 and pET-14b. The pET-14b constructs were pAM231 ( $\Delta$ N60 mutation), pAM227 ( $\Delta$ N144), pAM228 ( $\Delta$ C32) and pAM241 ( $\Delta$ N60 $\Delta$ C32).

RecG $\Delta$ C47 was constructed by subcloning the *XbaI*–*HindIII* *recG* fragment from pAM210 into pGEM-7f(-) (Promega). This plasmid was then cut with *EcoRI* and *HindIII*, the cleaved

restriction sites filled in with Klenow fragment and then ligated. This removed the final 47 codons from *recG*. This truncated *recG* was then cloned into pET-14b as a *XbaI*–*BamHI* fragment to form pPM112. The result was a *recG* gene with the final 47 codons deleted but with 23 codons derived from the pET-14b sequence added to the 3'-end.

### Determination of sensitivity to UV light

Sensitivity to UV light was determined as described previously (25). AB1157 (26) and N3793 ( $\Delta$ *recG263::kan*) (14) were used as wild-type and *recG* strains respectively.

### Purification of RecG proteins

Wild-type RecG was purified as described previously (15). AM1125, an *E. coli* BL21(DE3) *plysS* strain carrying  $\Delta$ *recG263::kan* (equivalent to GS1451 in 22), was transformed with each plasmid encoding a histidine tagged RecG protein. Batches of 500 ml of cells (100 ml for full-length HisRecG) were grown in LB broth containing 10 g/l NaCl and induced with 0.4 mM IPTG when the OD<sub>650</sub> reached 0.4–0.6. After 3 h induction the cells were harvested and resuspended in 4 ml 5 mM imidazole, 0.5 M NaCl, 20 mM Tris–HCl, pH 8. The cells were then sonicated on ice three times for 30 s and the lysate cleared by centrifugation at 28 000 *g* for 20 min. The supernatant was then loaded under gravity onto a 1 ml column of iminodiacetic acid insolubilized onto Sepharose 6B Fast Flow (Sigma), previously charged with NiSO<sub>4</sub> and equilibrated in the above lysis buffer. After washing the column with lysis buffer, bound proteins were eluted with 1 M imidazole, 0.5 M NaCl and 20 mM Tris–HCl, pH 8. The eluted proteins were then dialysed against 20 mM Tris–HCl, pH 8, 100 mM EDTA to remove any nickel ions present and then dialysed against 200 mM KCl, 1 mM EDTA, 50 mM Tris–HCl, pH 8. Contaminant proteins from the nickel column were then removed by gel filtration through a Superose 12 column (Pharmacia) with 200 mM KCl, 1 mM EDTA, 50 mM Tris–HCl, pH 8, as running buffer. The RecG proteins were then dialysed against 1 mM EDTA, 1 mM dithiothreitol, 50% (v/v) glycerol, 50 mM Tris–HCl, pH 8, and stored at –80°C. All chromatography and dialysis steps were performed at 4°C.

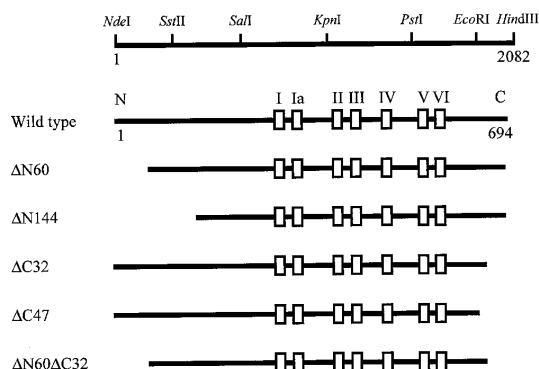
Protein concentrations were estimated using the Bradford assay with bovine serum albumin as standard (27) and are expressed as moles of monomeric protein.

### Construction of DNA substrates

A four way duplex junction designed to mimic a Holliday structure was made by annealing four 50mer oligonucleotides whose sequences have been described previously (28). This junction (J2) contained a 2 bp homologous core within which the junction point was free to branch migrate.

### DNA binding, helicase and ATPase assays

DNA binding and helicase assays were performed as described previously (13). Each assay was performed at least twice. The conversion of [ $\alpha$ -<sup>32</sup>P]ATP to [ $\alpha$ -<sup>32</sup>P]ADP was measured by thin layer chromatography as described (29). The 10  $\mu$ l reaction mixtures contained 20 mM Tris–HCl, pH 7.5, 2 mM dithiothreitol, 5 mM MgCl<sub>2</sub>, 5 mM ATP, 0.01–0.05  $\mu$ Ci [ $\alpha$ -<sup>32</sup>P]ATP, 100 mg/ml bovine serum albumin and 50 nM of the appropriate RecG protein. This buffer system was that used for measurement of



**Figure 1.** Schematic diagram of the *recG* gene with engineered *NdeI* and *HindIII* restriction sites, plus the five restriction sites introduced for the construction of mutations. The seven conserved helicase motifs, I–VI, within the *RecG* protein (18) are marked and the extent of the truncation within each mutant protein is indicated. The number of nucleotides and amino acids are also indicated.

helicase activity. An aliquot of 1 mg single-stranded 60mer oligo-nucleotide or  $\Phi X174$  RFI double-stranded DNA was added to reactions as required. Reactions were incubated at 37°C for 30 min and then quenched by addition of 2.5  $\mu$ l 0.5 M EDTA.

## RESULTS

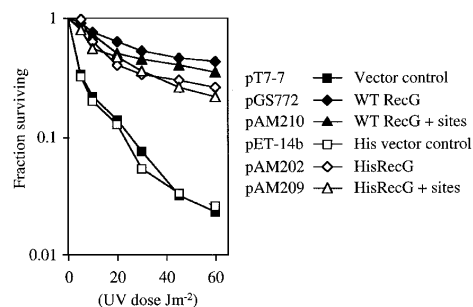
### Construction of truncated versions of *RecG*

To facilitate construction of deletion mutants five restriction sites were introduced within the coding sequence of the *RecG* overexpression plasmid pGS772 to form pAM210, as described in Materials and Methods. pAM210 was then used to construct five deletions of *recG* (Fig. 1).

The wild-type *RecG* protein was successfully expressed in a soluble form from both pGS772 and pAM210. However, the mutant forms of *RecG* were very poorly expressed or expressed in a predominantly insoluble form as judged from SDS-polyacrylamide gels (data not shown). To try and improve the recovery of soluble forms of the mutant proteins the mutant and wild-type versions of *recG* were subcloned into pET-14b. This vector allows expression of a fusion protein consisting of a 20 amino acid N-terminal peptide containing six tandem histidine residues plus the protein sequence of interest. Overexpression of these histidine tag fusion proteins in a *recG* strain did not generally improve the levels of expression or solubility (data not shown). However, the nickel binding property of the histidine tag was utilized to recover both wild-type and mutant *RecG* proteins in a soluble form as described in Materials and Methods.

### Comparison of wild-type and histidine tagged *RecG*

To ascertain the *in vivo* DNA repair activity of *RecG* fused to a histidine tag at the N-terminus the UV resistance of a  $\Delta recG263$  strain (N3793) carrying plasmid-encoded full-length *recG* with and without a histidine tag and with and without the engineered restriction sites was assayed (Fig. 2). The introduction of restriction sites had no observable effect upon the ability of these plasmids to promote UV survival (compare pGS772 with pAM210). The DNA repair defect in N3793 was also complemented by *RecG* proteins with a histidine tag and so the *RecG* fusion proteins were



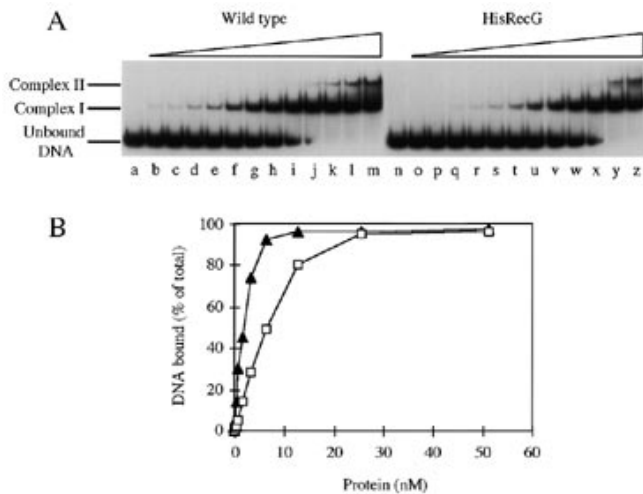
**Figure 2.** The effect of plasmids encoding wild-type and His*RecG* proteins upon the UV survival of strain N3793 ( $\Delta recG263::kan$ ). The data are the means of two experiments.

still competent in DNA repair. However, both strains containing plasmids encoding histidine tagged *RecG* (pAM202 with no engineered restriction sites and pAM209 with the engineered sites) were reproducibly more sensitive to UV radiation than the strains harbouring plasmids which encoded *RecG* without a histidine tag. This suggested that the N-terminal fusion caused some decrease in the ability of *RecG* to promote DNA repair. The effect of each of the above plasmids upon UV resistance in the wild-type strain AB1157 was also assayed, but there was no observable effect (data not shown). This demonstrated that the histidine tagged *RecG* did not inhibit the DNA repair functions of wild-type *RecG*.

*RecG* has been demonstrated to bind and unwind junction DNA and to hydrolyse ATP in a DNA-dependent reaction (8,9). The purified histidine tagged *RecG* fusion protein expressed from pAM209 (His*RecG*) was compared with wild-type *RecG* to see if these activities were impaired by the N-terminal fusion. Junction-specific DNA binding was assayed in band shifts with a synthetic Holliday structure having a 2 bp homologous core (J2). Wild-type *RecG* and His*RecG* gave the same pattern of DNA binding, forming an initial protein–DNA complex (complex I) at lower protein concentrations and a second complex (complex II) at higher concentrations (Fig. 3A). However, upon titration of a range of protein concentrations it was observed that His*RecG* was less efficient at junction DNA binding than the wild-type (Fig. 3B). Two- to 4-fold higher concentrations of His*RecG* were required in order to match the band shift activity of wild-type *RecG*. The ability to catalyse unwinding of the same synthetic Holliday junction was also analysed (Fig. 4). Approximately 10–20-fold more His*RecG* was required to match the dissociation activity of wild-type *RecG* (Fig. 4B and data not shown). Similar relative levels of dissociation were observed using four strand junctions containing homologous cores of 0 and 11 bp (data not shown). These data demonstrated that the reduction in the *in vitro* helicase activity of His*RecG* was not restricted to a junction with a 2 bp homologous core.

The reduction in junction-specific helicase activity for His*RecG* could have been due to a decrease in the rate of DNA unwinding or to a reduction in protein stability. To ascertain which of these factors contributed to the decreased levels of junction unwinding by His*RecG*, dissociation reactions were monitored over a period of 20 min (Fig. 4C). The results revealed that both wild-type *RecG* and His*RecG* had very similar patterns of activity, with all catalysis finished within 5–10 min of adding the protein. This could have been due to a reduction in ATP concentration or an





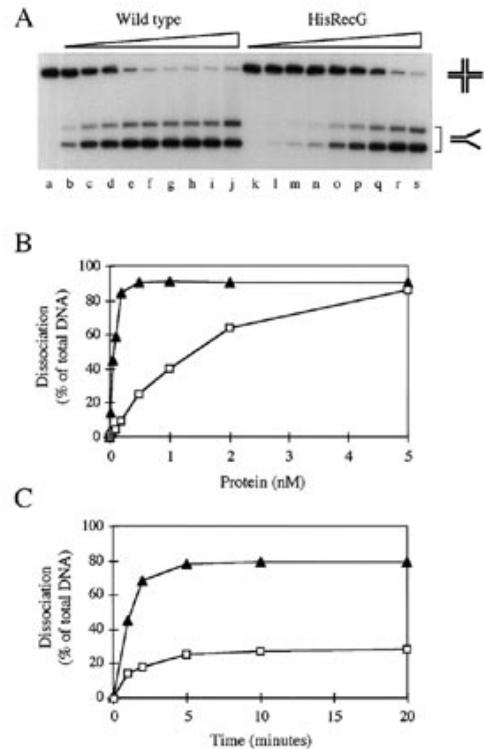
**Figure 3.** Binding affinity of wild-type and histidine tagged RecG for synthetic four strand junction DNA. (A) Autoradiograph of a band shift gel in which 0.3 nM J2 was mixed with 0.025, 0.05, 0.1, 0.2, 0.4, 0.8, 1.6, 3.2, 6.4, 12.8, 25.6 and 51.2 nM wild-type RecG (lanes b–m) or HisRecG (o–z). Protein was omitted from lanes a and n. (B) Quantification of the percentage of DNA bound in (A) by wild-type (closed triangles) and HisRecG (open boxes).

increase in ADP concentration as the reaction proceeded or it could have been caused by protein instability. However, the data do indicate that the stability of HisRecG does not differ substantially from wild-type RecG within the limits of the assay. We conclude that the reduced junction dissociation activity of HisRecG must be due to a reduction in the rate of DNA unwinding. HisRecG had an ~2–3-fold lower initial rate as compared with wild-type RecG (Fig. 4C and data not shown).

The junction-specific DNA helicase activity of RecG is coupled to hydrolysis of ATP (8; R.G.Lloyd and M.C.Whitby, unpublished observations). Therefore, wild-type RecG and HisRecG ATPase activities were compared in the absence of DNA and with single-stranded or double-stranded DNA in the reaction mix. No ATPase activity was detected for either wild-type or HisRecG in the absence of DNA or with single-stranded DNA present, but both proteins possessed double-stranded DNA-stimulated ATPase activity (data not shown). However, the level of ATP hydrolysis by HisRecG was reduced 10-fold as compared with wild-type RecG (2 and 21 nmol ATP hydrolysed respectively in 30 min). It should also be noted that, contrary to previously published data, the ATPase activity of wild-type RecG was specifically stimulated by double-stranded DNA and not by single-stranded DNA (R.G.Lloyd and M.C.Whitby, unpublished observations). Previous data (8,22) used single-stranded  $\Phi$ X174 DNA which could contain regions of duplex DNA as a result of the formation of secondary structures by intramolecular base pairing.

#### Deletion of N- and C-terminal regions of RecG

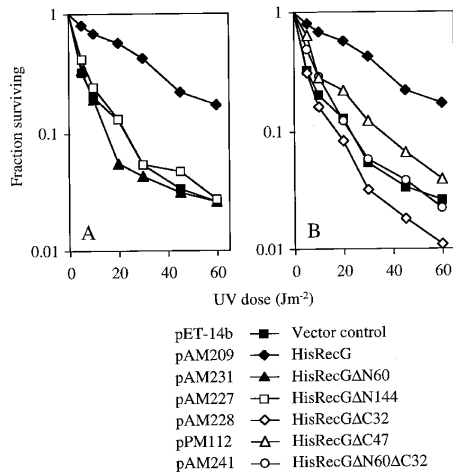
Deletions were chosen on the basis of conservation of protein sequence between the 10 putative RecG proteins available to date (data not shown). Nine truncated versions of *recG* were constructed but soluble protein could not be recovered from four of the deletions (unpublished data). Therefore, only five histidine tagged truncated RecG proteins were analysed in detail.



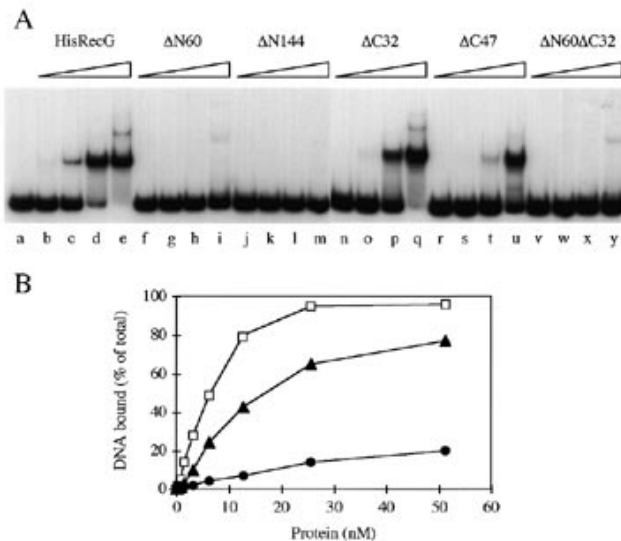
**Figure 4.** Dissociation of junction DNA by wild-type and histidine tagged RecG. (A) Autoradiograph of gel showing of four strand junction DNA (J2) to form flayed duplex products. 0.3 nM J2 was mixed with 0, 0.01, 0.05, 0.1, 0.2, 0.5, 1, 2, 5 or 50 nM wild-type RecG (lanes b–j) or HisRecG (lanes k–s). Protein was omitted from the reaction in lane a. (B) Comparison of the percentage of dissociation catalysed by wild-type RecG (closed triangles) and HisRecG (open boxes) from panel (A). (C) Time course of the unwinding of 0.3 nM J2 by 0.5 nM wild-type RecG (closed triangles) and HisRecG (open boxes).

The ability of the truncated versions of HisRecG to promote DNA repair was assessed by examining the effect of the deletion plasmids upon sensitivity to UV exposure. The *recG*<sup>+</sup> strain AB1157 and the  $\Delta$ *recG* mutant N3793 were transformed with plasmids encoding each of the five histidine tagged deletion mutants and the full-length HisRecG control. In the *recG*<sup>+</sup> background none of the plasmids had any effect upon UV sensitivity (data not shown). This demonstrated that the truncated RecG proteins had no effect upon the wild-type RecG encoded by the chromosome. In strain N3793 the plasmid encoding HisRecG restored resistance to UV light (Figs 2 and 5). In contrast, the truncated versions of HisRecG were unable to do so (Fig. 5). Plasmids encoding the wild-type and truncated versions of RecG lacking the histidine fusion peptide were also tested. The truncated forms of RecG again failed to restore resistance to UV light in the *recG* strain N3793 (data not shown).

The failure of the truncated HisRecG proteins to complement a *recG* mutation was investigated by analysing the activities of the proteins. DNA binding was assayed using band shift gels (Fig. 6A). Full-length HisRecG bound  $82 \pm 6\%$  of the total DNA at 100 nM protein. However, HisRecG $\Delta$ N60 gave only  $5 \pm 1\%$  binding at 100 nM and HisRecG $\Delta$ N144 gave no detectable DNA binding. The two band shift complexes obtained with HisRecG $\Delta$ N60 migrated more slowly than those of HisRecG, which indicated



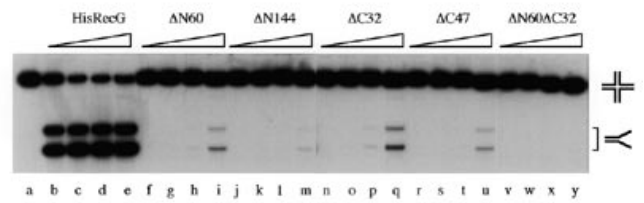
**Figure 5.** UV survival of strain N3793 ( $\Delta recG263::kan$ ) harbouring plasmids encoding full-length and truncated forms of HisRecG.



**Figure 6.** Binding of junction DNA by full-length and truncated forms of HisRecG. (A) Autoradiograph of a band shift gel in which 0.1, 1, 10 and 100 nM each protein was mixed with 0.3 nM J2. Protein was omitted from lane a. (B) Comparison of the percentage of J2 DNA bound by HisRecG (open boxes), HisRecG $\Delta$ C32 (closed triangles) and HisRecG $\Delta$ C47 (closed circles) in band shift gels.

either a different DNA:protein ratio within the complexes or a different conformation of the protein–DNA complex. In contrast, both HisRecG $\Delta$ C32 and HisRecG $\Delta$ C47 bound DNA to give similar band shift patterns to full-length HisRecG. HisRecG $\Delta$ C32 bound 85  $\pm$  9% and HisRecG $\Delta$ C47 bound 61  $\pm$  13% of the DNA at 100 nM protein. HisRecG $\Delta$ N60 $\Delta$ C32 gave a band shift identical to that of the single HisRecG $\Delta$ N60 mutant.

The DNA binding of HisRecG $\Delta$ C32 and HisRecG $\Delta$ C47 was analysed in more detail (Fig. 6B). HisRecG $\Delta$ C32 bound junction DNA with some reduction in affinity as compared with HisRecG. An  $\sim$ 2.5-fold higher concentration of HisRecG $\Delta$ C32 was required to achieve the same level of band shift. The affinity of



**Figure 7.** Autoradiograph of the products of dissociation reactions between junction DNA and full-length and truncated forms of HisRecG. 0.3 nM J2 DNA was mixed with 0.1, 1, 10 and 100 nM protein. Lane a lacked protein.

HisRecG $\Delta$ C47 for junction DNA was more severely reduced with a 15-fold higher concentration of the truncated protein required to match HisRecG. The length of the C-terminal region of RecG therefore correlated with the affinity for junction DNA.

Analysis of junction unwinding indicated that 100 nM HisRecG- $\Delta$ N60 catalysed dissociation of 1% of junction DNA as compared with 80  $\pm$  4% for HisRecG (Fig. 7). This correlated with the low level of DNA binding seen with this mutant (see above). HisRecG $\Delta$ N144 gave a trace amount of unwinding (<1%) despite there being no detectable band shift with this mutant. In contrast, both HisRecG $\Delta$ C32 and HisRecG $\Delta$ C47 gave only 3 and 1% dissociation respectively, despite the efficiency of DNA binding demonstrated above. The double mutant HisRecG $\Delta$ N60 $\Delta$ C32 had no detectable unwinding activity.

The *in vivo* studies described above indicate that there were no dominant negative effects upon wild-type RecG by any of the mutant RecG proteins analysed. The absence of any dominant negative effect was also observed *in vitro*. HisRecG $\Delta$ N60 or HisRecG $\Delta$ C47 were mixed with wild-type RecG and the mixtures then used to unwind junction DNA. Junction dissociation catalysed by wild-type RecG was not inhibited even when the mutant protein was present in 100-fold excess (data not shown).

Despite the lack of helicase activity and, for the N-terminal deletions, a lack of junction DNA binding, it was possible that the mutant proteins still retained ATPase activity. However, no ATP hydrolysis was detected for any of the mutants either in the presence or absence of DNA (data not shown).

## DISCUSSION

The properties of the RecG helicase which confer junction specificity and which define its role in DNA repair and recombination are unknown. Comparison of the 10 RecG sequences available indicated that all of the helicase motifs were clustered within the C-terminal portion of the protein (Fig. 1 and data not shown). However, despite having many conserved residues within the RecG proteins from a variety of prokaryotic species, the N-terminal 289 residues have no known homology with other proteins which could give an indication of the function of this region (data not shown).

The aim of this study was to begin to identify the features crucial to the multiple functions of RecG by creating a series of deletions of both the N- and C-termini. However, expression of truncated versions of RecG resulted in problems both with the levels of expression and with protein solubility. These difficulties were overcome by expression of the truncated forms of RecG with an N-terminal histidine tag. The wild-type RecG and full-length HisRecG were compared both *in vivo* and *in vitro*. The presence of the histidine tag reduced the *in vivo* efficiency of DNA repair

by RecG. This correlated with the *in vitro* data which demonstrated that HisRecG bound and unwound junction DNA with less efficiency than the wild-type and also exhibited a reduction in double-stranded DNA-stimulated ATPase activity. DNA binding is obviously a prerequisite for both junction unwinding and DNA-stimulated ATPase activity. It is therefore possible that the N-terminal fusion inhibited DNA–RecG interactions, although a more global structural rearrangement causing disruption of multiple functions of RecG cannot be excluded.

The five truncated versions of RecG, expressed with a histidine tag, were compared with full-length HisRecG. All five failed to complement the DNA repair defect seen in a  $\Delta recG$  strain, indicating that they had lost some activity essential for repair. However, they had different properties. The two N-terminal deletions had severely reduced DNA binding affinities. In both cases this feature was associated with very low levels of junction unwinding. Therefore junction DNA binding appears to have been disrupted by deletion of the N-terminal region, which correlates with the reduction in junction DNA affinity displayed by full-length HisRecG as compared with wild-type RecG.

The two C-terminal deletion mutants had significant but reduced levels of junction DNA binding. However, the amount of junction unwinding was as low as that displayed by the N-terminal deletion mutants. Therefore, the C-terminal 47 amino acids of RecG are important for helicase activity. Although this region contains none of the helicase motifs, it does contain 16 residues that are conserved between the 10 known versions of RecG (data not shown). This section of the protein may have a specific function that has yet to be recognized. However, this region may also be directly involved in helicase activity, which might explain the moderate reduction in DNA binding as well as junction unwinding in these mutants. The conserved helicase motifs are thought to interact with bound DNA in order to catalyse melting of the duplex (19,23). Therefore, mutations near the presumed helicase active site which reduced unwinding may have also attenuated protein–DNA interactions at the active site. ATP hydrolysis is also known to be linked to DNA unwinding (30). The severe reduction in ATPase activity seen with all the mutants may be explained by the inhibition of unwinding activity.

Deletion of both N- and C-termini in HisRecG $\Delta$ N60 $\Delta$ C32 resulted in binding levels similar to HisRecG $\Delta$ N60, but no detectable unwinding. This suggests that the DNA binding deficiency in the  $\Delta$ N60 mutation cannot support detectable levels of unwinding in the helicase-deficient  $\Delta$ C32 mutation.

Our findings suggest that RecG has at least two functional domains. The conserved helicase motifs form a C-terminal region that is essential for DNA unwinding. Not only are the recognized motifs crucial for this activity (22), but the C-terminal 47 residues are also required. The N-terminus may form a second functional domain that is involved in DNA binding. The helicase motifs are also thought to have a DNA binding role (23) and so it is tempting to speculate that the N-terminal region targets RecG specifically to junction DNA. The only known structure of a helicase, PcrA from *Bacillus stearothermophilus*, has a complex domain structure in which the helicase motifs are clustered around the interface formed between two domains (23). It has been postulated that all helicases share the same structure around the conserved motifs (23,31,32). The distinguishing properties of individual helicases may therefore be conferred by other, variable regions attached to the common helicase structure. The N-terminal region of RecG

may be such a variable domain, conferring junction specificity upon RecG. Other RecG-specific functions, such as interactions with other proteins involved in DNA repair and recombination, may also be specified by such domains.

## ACKNOWLEDGEMENTS

We thank Lynda Harris for technical support and Gary Sharples for drawing our attention to the conservation of sequences in RecG. This work was supported by the Medical Research Council, the Biotechnology and Biological Sciences Research Council and the Royal Society.

## REFERENCES

- 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) *Science*, **274**, 415–421.
- Ariyoshi, M., Vassilyev, D.G., Iwasaki, H., Nakamura, H., Shinagawa, H. and Morikawa, K. (1994) *Cell*, **78**, 1063–1072.
- Tsaneva, I.R., Müller, B. and West, S.C. (1992) *Cell*, **69**, 1171–1180.
- Parsons, C.A., Stasiak, A., Bennett, R.J. and West, S.C. (1995) *Nature*, **374**, 375–378.
- Dunderdale, H.J., Benson, F.E., Parsons, C.A., Sharples, G.J., Lloyd, R.G. and West, S.C. (1991) *Nature*, **354**, 506–510.
- Iwasaki, H., Takahagi, M., Shiba, T., Nakata, A. and Shinagawa, H. (1991) *EMBO J.*, **10**, 4381–4389.
- Whitby, M.C., Bolt, E.L., Chan, S.N. and Lloyd, R.G. (1996) *J. Mol. Biol.*, **264**, 878–890.
- Lloyd, R.G. and Sharples, G.J. (1993) *EMBO J.*, **12**, 17–22.
- Whitby, M.C., Vincent, S. and Lloyd, R.G. (1994) *EMBO J.*, **13**, 5220–5228.
- Lloyd, R.G. (1991) *J. Bacteriol.*, **173**, 5414–5418.
- Mitchell, A.H. and West, S.C. (1994) *J. Mol. Biol.*, **243**, 208–215.
- Whitby, M.C., Ryder, L. and Lloyd, R.G. (1993) *Cell*, **75**, 341–350.
- McGlynn, P., Al-Deib, A.A., Liu, J., Mariani, K.J. and Lloyd, R.G. (1997) *J. Mol. Biol.*, **270**, 212–221.
- Al-Deib, A.A., Mahdi, A.A. and Lloyd, R.G. (1996) *J. Bacteriol.*, **178**, 6782–6789.
- Vincent, S.D., Mahdi, A.A. and Lloyd, R.G. (1996) *J. Mol. Biol.*, **264**, 713–721.
- Fukuoh, A., Iwasaki, H., Ishioka, K. and Shinagawa, H. (1997) *EMBO J.*, **16**, 203–209.
- Whitby, M.C. and Lloyd, R.G. (1995) *EMBO J.*, **14**, 3302–3310.
- Lloyd, R.G. and Sharples, G.J. (1991) *J. Bacteriol.*, **173**, 6837–6843.
- Gorbalenya, A.E. and Koonin, E.V. (1993) *Curr. Opin. Struct. Biol.*, **3**, 419–429.
- Walker, J.E., Saraste, M., Runswick, M.J. and Gay, N.J. (1982) *EMBO J.*, **1**, 945–951.
- Fry, D.C., Kuby, S.A. and Mildvan, A.S. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 907–911.
- Sharples, G.J., Whitby, M.C., Ryder, L. and Lloyd, R.G. (1994) *Nucleic Acids Res.*, **22**, 308–313.
- Subramanya, H.S., Bird, L.E., Brannigan, J.A. and Wigley, D.B. (1996) *Nature*, **384**, 379–383.
- Tabor, S. and Richardson, C.C. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 1074–1078.
- Lloyd, R.G. and Buckman, C. (1991) *J. Bacteriol.*, **173**, 1004–1011.
- Bachmann, B.J. (1996) In Neidhardt, F.C., Curtiss, R., Ingraham, J.L., Lin, E.C.C., Low, K.B., Magasanik, B., Reznikoff, W.S., Riley, M., Schaechter, M. and Umberger, H.E. (eds), *Escherichia coli and Salmonella Cellular and Molecular Biology*, 2nd Edn. ASM Press, Washington, DC, pp. 2460–2488.
- Bradford, M.M. (1976) *Anal. Biochem.*, **72**, 248.
- Saito, A., Iwasaki, H., Ariyoshi, M., Morikawa, K. and Shinagawa, H. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 7470–7474.
- Kornberg, A., Scott, J.F. and Bertsch, L.L. (1978) *J. Biol. Chem.*, **253**, 3298–3304.
- Pause, A. and Sonenberg, N. (1992) *EMBO J.*, **11**, 2643–2654.
- Gorbalenya, A.E., Koonin, E.V., Donchenko, A.P. and Blinov, V.M. (1988) *FEBS Lett.*, **235**, 16–24.
- Hodgman, T.C. (1988) *Nature*, **333**, 22–23.