4200–4207 *Nucleic Acids Research, 1999, Vol. 27, No. 21*

Isolation and characterization of the C-terminal nuclease domain from the RecB protein of *Escherichia coli*

Xue-Juan Julie Zhang and Douglas A. Julin*

Department of Chemistry and Biochemistry, University of Maryland, College Park, MD 20742, USA

Received July 12, 1999; Revised and Accepted September 16, 1999

ABSTRACT

The RecB subunit of the *Escherichia coli* RecBCD enzyme has been shown in previous work to have two domains: an N-terminal 100 kDa domain with ATP-dependent helicase activity, and a C-terminal 30 kDa domain. The 30 kDa domain had nuclease activity when linked to a heterologous DNA binding protein, but by itself it appeared unable to bind DNA and lacked detectable nuclease activity. We have expressed and isolated this 30 kDa domain, called $RecB^N$, and show that it does have nuclease activity detectable at high protein concentration in the presence of polyethylene glycol, added as a molecular crowding agent. The activity is undetectable in a mutant $RecB^N$ protein in which an aspartate residue has been changed to alanine. Structural analysis of the wild-type and mutant $RecB^N$ proteins by second derivative absorbance and circular dichroism spectroscopy indicates that both are folded proteins with very similar secondary and tertiary structures. The results show that the Asp→Ala mutation has not caused a significant structural change in the isolated domain and they support the conclusion that the C-terminal domain of RecB has the sole nuclease active site of RecBCD.

INTRODUCTION

The RecBCD enzyme from *Escherichia coli* and other bacteria, made up of the RecB, RecC, and RecD proteins, has multiple catalytic activities that enable it to perform two different functions in the cell (reviewed in 1,2). The enzyme is a powerful nuclease that can protect the cell by degrading foreign DNA such as that of some bacteriophages. RecBCD is also a DNA helicase that has an important role in the repair of double-strand DNA breaks by homologous recombination. Coordination of the two enzymatic activities in the working enzyme, and thus the two biological functions, is regulated by a specific DNA sequence called Chi (5′-GCTGGTGG).

The reaction catalyzed by RecBCD enzyme with a linear double-stranded DNA molecule is a remarkable process in which ATP-dependent DNA unwinding is linked to concomitant nuclease degradation. The specific details of the reaction depend on the conditions, particularly on the ATP and Mg^{2+} concentrations (3–5), but a large body of experimental work indicates the following mechanism. The enzyme travels along a linear double-stranded DNA molecule, unwinding the DNA in an ATP-dependent reaction (6–8). The single-stranded DNA produced by the helicase is the substrate for the nuclease activity of the enzyme (7,9,10). The enzyme will cut both strands of a DNA duplex (10), but the 3′-terminated strand from the entry point is degraded to small fragments, especially at relatively high Mg^{2+} concentration where the overall nuclease activity is greatest, while the opposite 5′-terminated strand is cut at lower frequency (11–13).

The nuclease activity of RecBCD is modulated when it encounters a specific sequence in double-stranded DNA, the Chi recombination hot-spot (5′-GCTGGTGG) (14). At low Mg^{2+}, where the nuclease activity is low, the enzyme makes a limited number of nicks in the Chi-containing strand near the Chi sequence (15,16). A more dramatic effect of Chi is observed at higher Mg^{2+}, where the enzyme prior to Chi has substantial non-specific nuclease activity. Degradation of the 3′-strand is arrested but the enzyme continues to unwind the DNA to produce a specific single-stranded fragment (12,13). Moreover, under some conditions, the weak 5′-strand cleavage is enhanced after Chi in an apparent reversal of the nuclease polarity (17). Events at Chi *in vivo* are critical for the repair of double-strand breaks in the *E.coli* chromosome, since the 3′-terminated single strand produced after Chi is bound by RecA protein which then initiates homologous recombination with an unbroken chromosome (18,19). The biochemical basis for the effect of Chi on RecBCD is not completely understood. Proposed mechanisms include abrogation of the RecD or RecB subunit's role in the nuclease activity (20,21), ejection of the RecD subunit (1), or complete dissociation of the enzyme into its component subunits (22).

Two of the three subunits, RecB and RecD, have been directly implicated in the nuclease activity of RecBCD. The *recD* gene and RecD protein were discovered through the isolation of mutations that altered the protective function of RecBCD (i.e. the nuclease activity) but not the recombination function (23,24). Experiments with the purified proteins confirmed that the RecBC enzyme is a helicase with very low but detectable nuclease activity (21,25,26). Recent experiments have shown that the RecB protein alone has two enzymatic activities: DNA unwinding (27) and nuclease (28). The nuclease activity is lost when a small 30 kDa C-terminal domain is removed by limited proteolysis, but this small protein fragment alone had no detectable nuclease activity (21). The inactivity is apparently due to its low binding affinity

*To whom correspondence should be addressed. Tel: +1 301 405 1821; Fax: +1 301 314 9121; Email: dj13@umail.umd.edu

for single-stranded DNA, since fusion of the C-terminal RecB domain to a heterologous single-stranded DNA binding protein [the gene 32 protein (gp32) of phage T4] produced a chimeric protein that bound and hydrolyzed circular single-stranded DNA (28).

The C-terminal domain of RecB is central to the overall nuclease activity of RecBCD, since an Asp→Ala mutation in this part of RecB abolished all the nuclease activities of RecBCD on both single- and double-stranded DNA substrates (28). The mutant enzyme (RecB^{D1080A}CD) unwinds double-stranded DNA at a rate similar to the wild-type enzyme but it exhibits no cleavage at Chi in the Chi-containing strand. Moreover, there was no evidence for activation of opposite strand cleavage after Chi. Thus this domain of RecB is apparently responsible for cleavage by RecBCD of both DNA strands in a duplex.

These findings raise some interesting and important questions regarding the mechanism of the RecBCD enzyme nuclease activity. In particular, it is not clear that a single nuclease active site should be able to catalyze the complex reaction described above. It is unlikely from stereochemical considerations that a single nuclease active site would hydrolyze DNA with both apparent polarities, as proposed for RecBCD enzyme. The possibility remains that there is a second, unidentified nuclease active site in the enzyme. However, if that were true then the *recB*^{D1080A} mutation would have to have had greater effects on the enzyme than simply altering a putative Mg^{2+} binding site in the nuclease active site of the 30 kDa domain. It is conceivable that this mutation has led to more extensive structural alteration of the domain or its interactions with other subunits, to have such far-reaching effects on the holoenzyme catalytic activity.

Further study of the function of the C-terminal domain of RecB in the RecBCD-catalyzed nuclease reaction, and characterization of the effects of the *recB*^{D1080A} mutation, will help to resolve these issues. To this end we have expressed and purified the C-terminal RecB domain and the Asp→Ala mutant domain, free of the gp32 moiety. We detect weak nuclease activity with the wild-type protein that is lacking in the mutant, confirming that this domain has a nuclease active site. The wild-type and mutant proteins are essentially identical in structure by two spectroscopic measures, indicating that there has been no drastic structural change as a result of the mutation. These data argue against the possibility that the D1080A mutation has caused any widespread disturbance in the structure of the RecBCD holoenzyme, and they emphasize the significance of the C-terminal nuclease domain of RecB in the overall activity of RecBCD.

MATERIALS AND METHODS

Materials

Single-stranded circular bacteriophage M13 DNA, a gift from Ms Jehanne Souaya, and RecB protein, a gift from Dr Misook Yu, were purified as described (28). Polyethylene glycol 8000 (PEG 8000, molecular weight 7–9 kDa) was purchased from Fisher Scientific. Egg white lysozyme, phenylmethylsulfonyl fluoride (PMSF), and imidazole were from Sigma. The *E.coli* GroEL and GroES proteins were a gift from Dr George Lorimer of this Department.

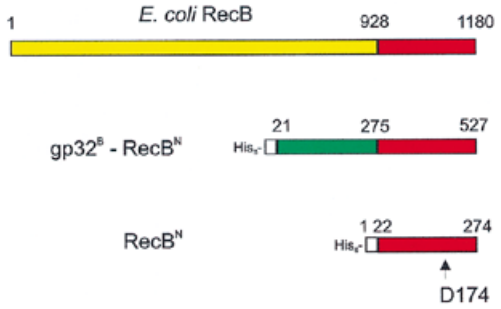

Figure 1. Structures of RecB, the chimeric gp32^{B}–RecB^{N} protein, and the RecB^{N} protein. In gp32^{B}–RecB^{N} (28), residues 928–1180 of RecB (red) were joined to residues 1–254 of the bacteriophage T4 gene 32 protein (green), a non-specific single-stranded DNA binding protein. RecB^{N} consists of residues 928–1180 of RecB with a 21 residue N-terminal peptide containing a hexahistidine tag. The position of the aspartate residue (D174) that has been mutated in RecB^{N} (D1080 in RecBCD; D427 in gp32^{B}–RecB^{N}) is also indicated.

Construction of a plasmid encoding the RecB^{N} protein

The 30 kDa C-terminal domain of RecB was expressed with an N-terminal hexahistidine tag using the vector pET-15b (Novagen Corp.). The DNA encoding the C-terminus of RecB in the plasmid pMY60 (28) was amplified by PCR using primers that introduced *Eco*RI and *Nde*I sites at the 5′-end of the gene and a *Bam*HI site just after the RecB stop codon. The amplified DNA product was first inserted into the vector pTZ19R cleaved with *Eco*RI and *Bam*HI. The DNA encoding the RecB^{N} protein was then excised from this plasmid by cutting with *Nde*I and *Bam*HI and ligated into pET-15b cut with the same enzymes to produce pET15-30. This plasmid encodes a 274 residue polypeptide called RecB^{N} (Fig. 1) consisting of a 21 amino acid N-terminal peptide (MGSSHH-HHHHSSGLVPRGSHM) followed by residues 928–1180 of the RecB protein. The DNA encoding the entire protein was sequenced and no mutations were found.

Asp174 in the RecB^{N} protein was changed to Ala using the QuickChange mutagenesis kit (Stratagene) as described previously for the D427A mutation in gp32^{B}–RecB^{N} (28; see Fig. 1). The DNA encoding the mutant protein was sequenced and no other mutations were found.

Expression and purification of the wild-type and D174A mutant RecB^{N} proteins

pET15-30 encoding either the wild-type RecB^{N} or RecB^{N}(D174A) was introduced into *E.coli* strain BL21(DE3) (Novagen Corp.) for protein expression. A 500 ml culture of BL21(DE3) (pET15-30) was grown at 37°C in LB medium containing ampicillin (25 µg/ml) to OD_{600} = 0.6. IPTG was added to 1 mM and growth was continued for 3 h. The culture was placed on ice for 5 min and the cells were harvested by centrifugation at 5000 *g* for 5 min. The cell pellet (2.5 g) was resuspended in 125 ml of ice-cold 50 mM Tris–HCl, pH 8, 2 mM EDTA and the cells were sedimented again. This cell pellet was resuspended in 100 ml of native binding buffer (20 mM sodium phosphate, pH 7.8, 0.5 M NaCl, 5 mM imidazole), PMSF was added to 1 mM, lysozyme was added to 100 µg/ml, and the mixture was placed on ice for 15 min. The cells were then lysed by sonication with cooling in a –5°C ice–salt bath.

The cell debris was removed by centrifugation at 12 000 *g* for 20 min at 4°C.

The $RecB^N$ protein was purified by chromatography on Ni^{2+} resin (25 ml bed volume, ProBond resin; Invitrogen Corp.) equilibrated with native binding buffer. The cleared lysate was applied to the column and weakly bound proteins were removed by washing with 100 ml of 20 mM sodium phosphate, pH 6.0, 60 mM imidazole, 0.5 M NaCl, 1 mM PMSF. The $RecB^N$ protein was eluted in a 300 ml linear gradient of 0.1–0.8 M imidazole in 20 mM sodium phosphate, pH 6.0, 0.5 M NaCl, 1 mM PMSF. EDTA was added to each empty fraction tube to give ~20 mM (final concentration) to inhibit proteases. Fractions containing the $RecB^N$ protein were pooled and concentrated by ultrafiltration. Larger protein impurities (including GroEL; see Results and Discussion) were removed by ultrafiltration through a Centricon-100 microconcentrator (Amicon). The $RecB^N$ protein that passed through the Centricon-100 filter was then concentrated using a Centricon-10 filter (Amicon). The purified protein was dialyzed into 10 mM potassium phosphate, pH 6.9, 2 mM EDTA, 10% (v/v) glycerol, 0.01 mM DTT, 0.1 mM PMSF. The protein concentration was determined from the absorbance at 280 nm using ε_{280} = 39400 M^{-1} cm^{-1}, calculated from the number of Trp (4) and Tyr (13) residues in the $RecB^N$ protein. The wild-type protein has a calculated molecular weight of 31 143 g/mol.

Samples for N-terminal sequencing were run on 12% SDS–PAGE and transferred to polyvinylidene difluoride membranes (Millipore Corp.) as described (21). N-terminal sequences were kindly determined by Dr Brian Martin (National Institutes of Health) by Edman degradation using an Applied Biosystems model 470 gas phase protein sequencer.

Single-stranded DNA endonuclease assay

The endonuclease activity of the wild-type and the mutant $RecB^N$ proteins was measured using circular single-stranded M13 phage DNA as the substrate. Most reaction mixtures contained 6 or 12 nM circular single-stranded M13 DNA with other conditions noted in the text and figure legends. The $RecB^N$ (1.8 μM) or RecB protein (20 nM) was added to initiate the reaction. Samples (3 μl) were removed at the specified time points and quenched with (final concentrations) 24 mM EDTA, 0.5% SDS, 8% glycerol, and 0.025% bromophenol blue. The quenched samples were loaded on 0.8% agarose gels prepared in 89 mM Tris–borate, pH 8.0, 2 mM EDTA and run at 120 V (4 V/cm) for 3 h. The gels were stained with ethidium bromide (5 μg/ml) for at least 1 h and photographed using Polaroid 667 film. The photographs were scanned using an Astra 1200S scanner (UMAX Technologies Inc.) and the fraction of DNA circles converted to linear was determined by analysis of the scanned image using the Scion Image program (Scion Corporation, Frederick, MD). Known amounts of circular DNA were loaded on some gels to serve as standards. The amount of linear DNA present initially in the circular DNA (3–8% of the total) was subtracted out.

UV absorbance spectroscopy and second derivative analysis

UV absorbance spectra of the mutant and wild-type $RecB^N$ proteins were taken in 10 mM potassium phosphate, pH 6.9, using a Hewlett Packard 8453 diode array UV-visible spectrophotometer. Fifty spectra were obtained at ambient temperature in a 1 cm path length cell at 1 nm wavelength intervals and then averaged.

The optical density (OD) data were corrected for light scattering with the HP 845X UV-visible system software as in Mach *et al.* (29) using the OD readings from 320 to 350 nm. (The calculated OD due to light scattering at 280 nm was 4 and 6% of the total OD for the wild-type and mutant proteins, respectively.) The absorbance data were smoothed by application of a polynomial of degree 3 and the second derivative was calculated with a filter length of 5 using the HP system software as in Savitzky and Golay (30).

Circular dichroism (CD) spectroscopy

The CD spectra of the mutant and wild-type proteins were measured using a Jasco J-500C spectropolarimeter equipped with a Jasco DP-500N data processor. The instrument was calibrated using (1*S*)-(+)-10-camphorsulfonic acid. Far-UV spectra (260–190 nm) were taken in a 0.1 cm path length cell with a 2 s time constant and a 10 ms sampling time. Four spectra were taken, averaged, and displayed on a chart recorder in analog form. The protein concentrations were 2.9 (wild-type) and 4.5 μM (mutant) in 10 mM potassium phosphate, pH 6.9. Near-UV spectra (320–250 nm) were taken in a 1.0 cm path length cell with a 4 s time constant and a 10 ms sampling time. The protein concentrations were 42 μM for both proteins in 10 mM potassium phosphate, pH 6.9, 2 mM EDTA, 10% (v/v) glycerol, 0.01 mM DTT. Eight spectra were taken and averaged. The observed ellipticities (θ) on the chart recordings (millidegrees/cm on the chart) were converted to molar ellipticities ([θ] in (degrees·cm^2/dmol)/cm) for the *y*-axes in Figures 6 and 7 from $[\theta] = [100 \times (\theta/1000)]/cl$, where *c* is the molar residue concentration (Fig. 6) or molar protein concentration (Fig. 7), and *l* is the cell path length in cm (31).

The CD spectrum of the wild-type $RecB^N$ protein was also obtained in digital form using a Jasco J-720 spectropolarimeter. The protein (93 μM) was in 10 mM potassium phosphate buffer, pH 6.9, 10% glycerol, 0.01 mM DTT. The spectrum was taken from 185 to 250 nm at 0.5 nm intervals in a 0.01 mm path length cell. Three spectra were taken and averaged. The secondary structure of the wild-type protein was calculated from these data (see Results and Discussion). The sample was then diluted 50-fold with 0.1 M potassium phosphate, pH 6.9, and the CD spectrum as a function of temperature was monitored at 222 nm in a 1 mm path length cell. The temperature was varied from 20 to 90°C at 1°C/min and θ_{222} readings were taken every 10 s.

RESULTS AND DISCUSSION

Expression and purification of the $RecB^N$ protein

A significant fraction of the overexpressed $RecB^N$ protein was soluble and found in the lysate supernatants. This soluble protein was readily purified by metal chelate affinity chromatography under native conditions on Ni^{2+} resin. The protein was highly pure after this step, but it was accompanied by at least one other larger protein (~60 kDa) that was intriguingly close in size to the RecD protein. However, N-terminal sequencing showed that this protein is the *E.coli* GroEL chaperonin protein [the sequence obtained, AAKDVKFGNDARVKMLR, matches that of GroEL (GenBank accession no. AAC77103)

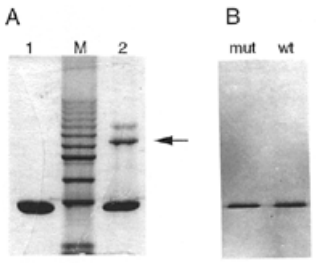

Figure 2. The purified wild-type $RecB^N$ protein and the $RecB^N$(D174A) mutant. Protein samples were analyzed by 12% SDS–PAGE and stained with Coomassie blue. (**A**) Lane 1, $RecB^N$ protein (2.9 μg) purified by chromatography on Ni^{2+} resin and filtration through a Centricon-100 filter. M, 10 kDa protein ladder (Life Technologies Corp.). The band in lane M next to the $RecB^N$ protein in lane 1 is the 30 kDa marker. Lane 2, protein retained by Centricon-100 and enriched for the larger impurities relative to the remaining $RecB^N$ protein. Arrow indicates GroEL protein (see Results and Discussion). (**B**) mut, Purified D174A mutant (1 μg loaded); wt, wild-type $RecB^N$ (0.4 μg).

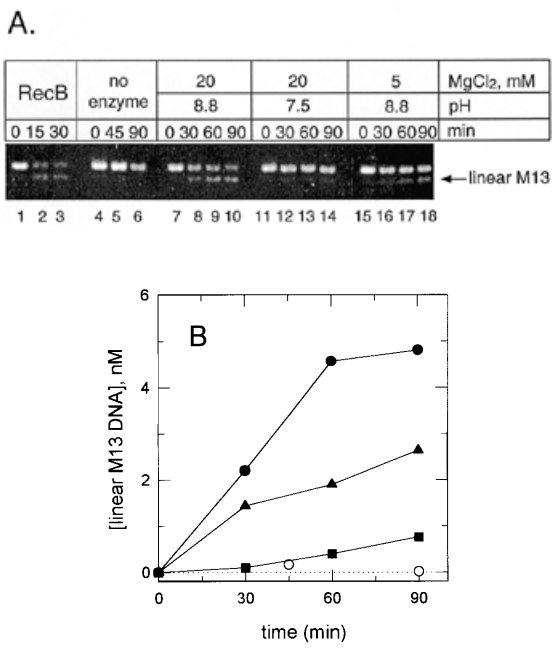

Figure 3. Single-strand endonuclease reactions. (**A**) Lanes 1–3, reaction mixture containing 50 mM MOPS–KOH, pH 7.0, 10 mM $MgCl_2$, 6 nM circular single-stranded M13 DNA, and 20 nM RecB protein. Lanes 4–6, 50 mM Tris–HCl, pH 8.8, 10 mM $MgCl_2$, 10% PEG 8000, 6 nM M13 DNA, with no added protein. Lanes 7–18, 12 nM M13 DNA, 10% PEG 8000, 1.8 μM $RecB^N$ protein, and buffer pH and $MgCl_2$ as indicated. Samples were removed at the indicated times and analyzed on a 0.8% agarose gel as described in Materials and Methods. (**B**) The reactions shown in (A) were quantitated using Scion Image (see Materials and Methods). Filled circles, lanes 7–10, $RecB^N$ protein, pH 8.8, 20 mM $MgCl_2$; squares, lanes 11–14, $RecB^N$ protein, pH 7.5 20mM $MgCl_2$; triangles, lanes 15–18, $RecB^N$ protein, pH 8.8, 5 mM $MgCl_2$; open circles, lanes 4–6, no added $RecB^N$ protein.

assuming that the N-terminal Met was removed]. The large impurities could be separated from the $RecB^N$ protein by ultrafiltration through a Centricon-100 microconcentrator (Fig. 2A).

The $RecB^N$ protein purified only through the Ni^{2+} column was susceptible to adventitious proteolysis to give a slightly smaller product that could be resolved from the remaining intact protein by 12% SDS–PAGE (not shown). Interestingly, the major proteolytic product had as its N-terminus RVVEEPTLTPHQF, which matches well to residues 27–39 of the His-tagged $RecB^N$ protein, (S)VVEEPTLTPHQF. Thus the 21 residue N-terminal His-tag peptide and five residues (AAGVA) of the RecB portion were removed. However, the resulting truncated protein appeared to be refractory to further degradation, reminiscent of the resistance against subtilisin that was originally used to produce the 30 kDa domain (21). The proteolysis was prevented by including PMSF throughout the purification procedure and EDTA after the Ni^{2+} column. The protein purified under these conditions migrated as a single band on SDS–PAGE (Fig. 2A).

Nuclease activity of the $RecB^N$ protein

In previous work, the 30 kDa proteolytic fragment had no detectable nuclease activity on circular single-stranded M13 DNA in reactions containing as much as 0.7 μM protein (21). The protein fragment had been purified away from the larger fragment because it flowed through a single-stranded DNA–agarose column, and its lack of activity was attributed to low affinity for the DNA substrate. However, the 30 kDa domain must be able to bind its DNA substrate, if only at the phosphodiester bond that is to be cleaved, if it has a nuclease active site. We reasoned that a slow nuclease reaction would be faster and more easily detected at high $RecB^N$ protein and/or high DNA concentrations. We thus sought to detect DNA cleavage using the highest possible $RecB^N$ protein concentration but with a DNA concentration suitable for detecting the cleaved DNA product by ethidium staining.

The purified $RecB^N$ protein gave essentially no cleavage of circular M13 DNA at 1.1 μM protein in 90 min at pH 7, 10 mM Mg^{2+} (~2% cleaved; data not shown), conditions under which the intact RecB protein has readily detectable endonuclease activity (Fig. 3A; 28). We did observe weak cleavage of the DNA by the $RecB^N$ protein at pH 8.8, the condition that was found to be optimal for the $gp32^B$–$RecB^N$ fusion protein (28). Having detected a low level of nuclease activity with the $RecB^N$ protein, we sought conditions that might enhance it. Somewhat greater cleavage was observed at higher Mg^{2+} concentration and if a neutral polymer was included in the assay mixture. Thus, addition of 5–10% PEG 8000 gave a clearly visible reaction at pH 8.8, 20 mM Mg^{2+}, while the reaction was slower but readily detectable at lower Mg^{2+} (5 mM), and still slower at pH 7.5 (Fig. 3A and B). A very high protein concentration (μM) was required under all conditions that we have tested. We presume that PEG enhances the reaction by acting as a 'molecular crowding agent' (32) that increases the effective concentrations of the protein and DNA in the assay. There was no detectable reaction in the absence of $RecB^N$ protein (Fig. 3A and B).

The slow rate of the nuclease reactions with the $RecB^N$ protein is not unexpected given the previous work cited above. However, the possibility that the $RecB^N$ protein preparation is contaminated with a small amount of another nuclease must also be addressed. We did so by preparing the same Asp→Ala

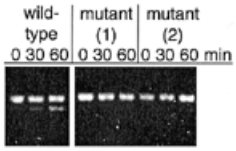

Figure 4. Single-strand endonuclease reactions with the wild-type and D174A mutant RecB^N proteins. Reaction mixtures contained 50 mM Tris–HCl, pH 8.8, 10% PEG 8000, 20 mM $MgCl_2$, 12 nM M13 DNA, and 1.8 μM of the indicated protein (two different preparations of the mutant protein were used). Samples were analyzed on a 0.8% agarose gel as described in Materials and Methods.

mutation that knocked out the nuclease activity of both the gp32^B–RecB^N fusion protein and of RecBCD (28). The RecB^N(D174A) mutant was purified by the same procedure as the wild-type protein (Fig. 2B). Two separate preparations of this mutant protein had no detectable nuclease activity (<2% circles cleaved) under conditions where the wild-type RecB^N protein is active (22% cleavage in 1 h; Fig. 4). This result, along with previous observations of the consequences for the nuclease activity of altering this aspartate residue (28), confirm that the observed nuclease cleavage reaction is promoted by the RecB^N protein.

We also noted that the nuclease activity of the RecB^N protein ultrafiltered through Centricon-100 (see Materials and Methods) was essentially identical to that of the material that had only been purified over the Ni^{2+} column (data not shown). This observation supports the conclusion that the activity is intrinsic to the RecB^N protein, and also shows that the filtration procedure is not detrimental to the protein.

The reaction rate with the RecB^N protein is exceedingly low, 0.0022 (± 0.0008) circles cleaved/h/RecB^N under the conditions of Figure 3A (lanes 7–10), and is much slower than that found previously with the intact RecB protein (~0.5 circles cleaved/h/RecB protein; 28). We did not compare the reaction with the RecB^N protein directly to that with the gp32^B–RecB^N fusion protein. However, previous experiments with that enzyme clearly showed greater cleavage of the circular substrate (~0.2 circles cleaved/h/gp32^B–RecB^N; estimated from data in 28), at pH 8.8 but in the absence of PEG.

The observation that a small amount of GroEL protein co-purified with the RecB^N protein suggested that at least some of the RecB^N sample is unfolded or misfolded, since the function of GroEL is to bind such proteins and facilitate their folding to the native state (33). GroEL would be retained on the Ni^{2+} column if the His-tag of RecB^N bound to GroEL were exposed enough to bind the column. (We know of no other published reports of GroEL co-purifying with histidine-tagged proteins. GroEL has been reported to co-purify with GST fusion proteins on glutathione columns; 34.) Moreover, misfolding of the RecB^N protein could conceivably account for its low enzymatic activity. Thus we tested whether the chaperonin activity of GroEL and GroES proteins could activate the RecB^N nuclease by remodeling its tertiary structure. However, there was no increase in the nuclease reaction rate when we incubated RecB^N (1 μM) in a reaction mixture that contained 12 nM M13 DNA and GroEL and GroES (1 μM each), KCl (50 mM), $MgCl_2$ (10 mM), and ATP (0.65 mM), at pH 7 (50 mM MOPS) or at pH 8.8 (50 mM Tris–HCl) (data not shown).

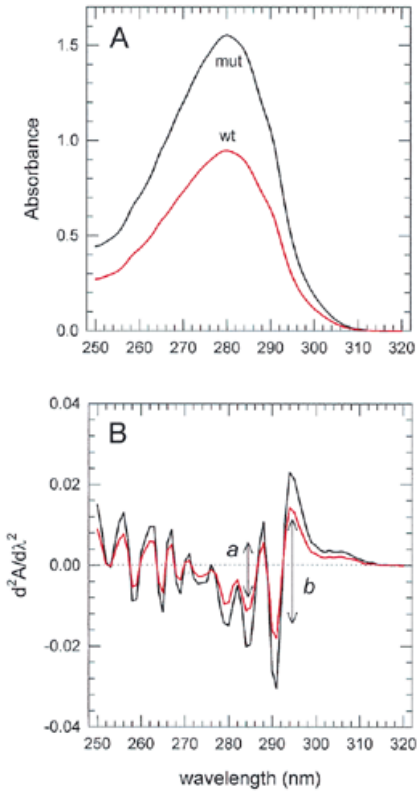

Figure 5. UV absorbance spectrum (A) and second derivative of the absorbance spectrum (B) for wild-type (red curves) and D174A mutant (black curves) RecB^N proteins. (**A**) Optical density was measured and corrected for light scattering as described in Materials and Methods. The protein concentrations were 24 (RecB^N) and 39 μM [RecB^N(D174A)], in 10 mM potassium phosphate, pH 6.9. (**B**) The second derivative of the absorbance in (A) was calculated as described in Materials and Methods. The peak-to-peak distances used to calculate the parameters *a* and *b* (Table 1) are indicated.

Structural comparison of the wild-type and D174A mutant RecB^N proteins

We wished to obtain some basic structural information on the RecB^N protein, to perhaps shed light on why its activity is so low. It could be that this isolated domain is unfolded or misfolded (the GroEL result mentioned above notwithstanding) and thus has low activity. Moreover, the D174A mutation might alter the local conformation near the active site or it could lead to substantial unfolding of the protein. Thus we carried out two spectroscopic analyses summarized below.

Second derivative of the absorbance spectrum. The second derivative of the absorbance spectrum ($d^2A/d\lambda^2$) of a protein in the 280 nm region is a sensitive indicator of the polarity of the microenvironments of aromatic residues, particularly tyrosine, in the protein (35). Thus this method is a good way to detect conformational differences between two related proteins. The near-UV absorbance spectra and the second derivatives of the absorbance are shown in Figure 5A and B for the wild-type and mutant proteins. The second derivatives are virtually identical for the two proteins and, with maxima at 294 and 288 nm and minima at 291 and 284, are typical of most proteins (35). The peak-to-trough differences *a* (288–284 nm) and *b* (294–291 nm),

and the *a*/*b* ratio, were calculated as in Ragone *et al*. (35) and are given in Table 1. Their magnitudes are in the range observed for native, folded proteins and the values for the wild-type and mutant are quite similar to each other. The 288–284 difference *a* in particular is sensitive to the polarity of the microenvironment of Tyr residues (35). The slight difference in the *a*/*b* ratio between the wild-type and D174A proteins could be because Asp174 is adjacent to a Tyr residue, whose environment would be slightly altered by loss of the carboxyl group of the Asp residue in the D174A mutant. This difference notwithstanding, it is clear that both proteins are folded and that their overall structures are very similar.

Table 1. Peak-to-peak distances in the second derivative spectra[a]

Protein	*a*	*b*	*a*/*b*
$RecB^N$	0.017	0.032	0.53
$RecB^N$(D174A)	0.031	0.053	0.58

[a]The parameters *a* and *b* are the absolute differences in the magnitude of the second derivative at 284 and 288 nm (*a*) and 291 and 294 nm (*b*) in Figure 5B.

Circular dichroism spectra. The far-UV CD spectra of the wild-type and mutant $RecB^N$ proteins (Fig. 6) are quite similar qualitatively and quantitatively (note the different *y*-axis scales for the two proteins in Fig. 6), indicating that their overall secondary structures are essentially identical. The large negative ellipticity in the 210–220 nm range and the maximum at 192 nm are typical of a protein with substantial secondary structure (31). The CD spectrum of the wild-type $RecB^N$ protein was also obtained in digital form from 185 to 250 nm (data not shown). Secondary structure calculations using these data suggest that the protein has substantial amounts of α-helix. The predicted structure calculated with the program CDNN (36) was 66% α-helix, 4% β-sheet, 12% β-turn, and 12% random coil, while the structure calculated using the program CDsstr (37) was predicted to be 78% α-helix, 4% extended β-sheet, 3% β-turn, 11% 3/10 helix, and 3% other structures.

The effect of temperature on the wild-type protein was tested by monitoring the CD signal at 222 nm from 20 to 90°C. There was little change in the ellipticity up to ~50°C, an abrupt increase at ~52°C, and then a gradual increase as the temperature was raised further (data not shown). This observation suggests unfolding of the protein followed by aggregation. Consistent with this finding was the fact that there was only partial recovery of the CD spectrum when the sample at 90°C was cooled quickly back to ~20°C (θ_{222} = –2.79 millidegrees before melting; –1.16 after melting and quick cooling). Thus while some of the heat-denatured protein could regain its secondary structure, some appeared to be irreversibly unfolded and/or aggregated.

The near-UV CD spectra of the wild-type and D174A mutant proteins are shown in Figure 7. The CD in this wavelength range is indicative of aromatic residues that are in an asymmetric environment due to the folded structure of the protein (38). The two spectra are virtually identical, indicating similar tertiary structures of these two proteins.

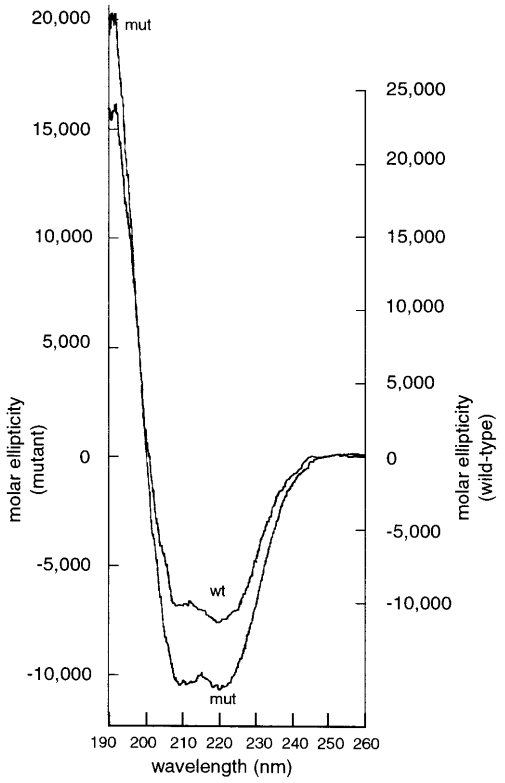

Figure 6. Far-UV CD spectra of wild-type and D174A mutant $RecB^N$ proteins. CD spectra were measured in a 0.1 cm path length cell with 2.9 (wild-type) or 4.5 μM (mutant) protein. The original *y*-axis scales in the analog spectra (in millidegrees) were converted to molar ellipticity (in degrees·cm^2/dmol residue) as described in Materials and Methods. Note that separate *y*-axis scales are given for the two proteins (mutant on the left, wild-type on the right) since their spectra were taken at different protein concentrations. The spectra of the wild-type and mutant proteins were overlaid for comparison.

Taken together, these two spectroscopic methods provide strong evidence that the wild-type and mutant proteins are folded polypeptides that have secondary and tertiary structure, and that their structures are very similar. Moreover, it is clear that the D174A mutation does not induce a drastic structural change in the protein, although we cannot rule out small structural differences between the mutant and wild-type proteins.

Overall summary

Previous work indicated that the RecB protein is organized into at least two domains: a C-terminal domain containing a nuclease active site with very low affinity for single-stranded DNA, joined to a larger N-terminal domain that binds the DNA and holds it near this active site for cleavage (21,28). We have shown here that the isolated C-terminal domain is indeed able to cleave DNA, as it must be if it in fact has a nuclease active site. The requirement for a high protein concentration and the effect of the molecular crowding agent PEG is consistent with its low DNA binding affinity. We presume that it has catalytic residues that interact with the phosphodiester bond that is cleaved, but that it otherwise has little interaction with the DNA chain adjacent to the cleavage site.

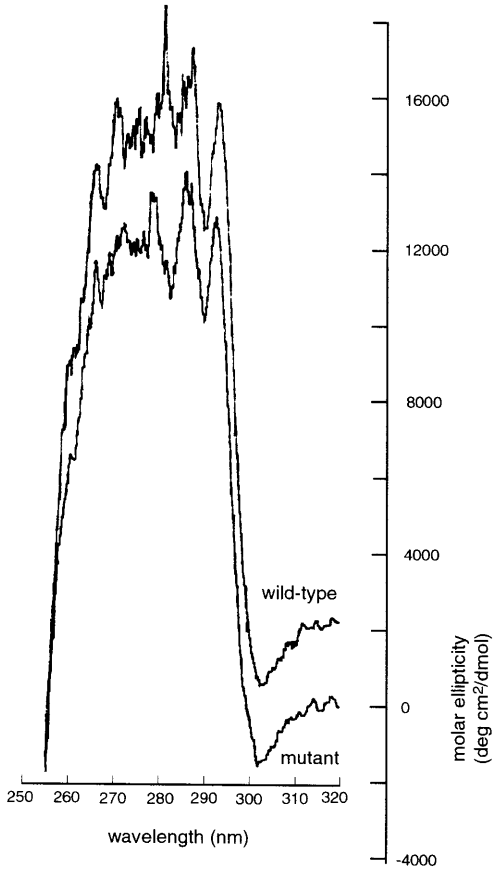

Figure 7. Near-UV CD spectra of wild-type and D174A mutant $RecB^N$ proteins. The near-UV CD was measured in a 1 cm path length cell with 42 μM protein (wild-type and mutant). The molar ellipticity for the *y*-axis scale (in degrees·cm^2/dmol protein) was calculated as described in Materials and Methods. The spectra of the wild-type and mutant proteins were overlaid and the wild-type spectrum was offset vertically to facilitate direct comparison of the two.

The effect of the D174A mutation on $RecB^N$, and by extension the D1080A mutation on the nuclease activity of RecBCD, is most likely due to a localized effect on the nuclease active site and not on the overall structure of the enzyme. That this single amino acid change abolishes all nuclease activity of this complex enzyme, with little if any change in structure of the C-terminal domain, therefore strengthens the argument that the C-terminus of RecB contains the nuclease active site of RecBCD, and there remains no evidence for a second nuclease site in this enzyme.

The linkage of both nuclease activity and a second catalytic activity as separable domains in a single polypeptide is an arrangement found in a number of other enzymes, the classic example being bacterial DNA polymerase I (39). Several proteins that have RecB-like nuclease domains have been identified recently by sequence analysis (40), including the DNA2 protein of *Saccharomyces cerevisiae*, a protein that like RecB also has a helicase domain (41), and the RecE protein of *E.coli*, a 5′→3′ exonuclease on double-stranded DNA (2). The Werner syndrome helicase has also been found to have separable DNA helicase and nuclease domains (42), although neither is homologous to the corresponding domain in RecB (40). In addition, there are a number of examples of nucleases in which the catalytic domain is separable from a sequence-specific DNA binding domain, including several restriction endonucleases (43,44) and homing endonucleases of some mobile introns and inteins (45). For some of these proteins, the isolated nuclease domain had detectable activity (42,43,46), while it was undetectable for others (44,47).

It remains to be determined why the very weak nuclease activity of RecB is so greatly increased when all three subunits are present in RecBCD. The very low activity of the isolated $RecB^N$ protein is undoubtedly in part an artefact of its separation from its natural milieu in the RecBCD holoenzyme. At the very least the other regions of the enzyme will maintain a high local DNA concentration near the active site in RecB, due to DNA–protein interactions by the N-terminal domain of RecB and by the other subunits. There are also likely to be protein–protein interactions of the 30 kDa domain with the rest of the enzyme that contribute in unknown ways. In this regard it should be noted that we cannot be certain that the structures of the wild-type and mutant 30 kDa domains, although similar to each other, are identical to their structures in their normal surroundings as part of the RecBCD holoenzyme.

It is interesting to note that having the nuclease domain well separated functionally from the major DNA binding domain(s) of RecBCD would have some benefits for the function of this enzyme. First, this arrangement might be key to explaining how a single nuclease active site could hydrolyze either strand of a DNA duplex. We presume by analogy to the catalytic mechanisms proposed for other nucleases that there must be a strict stereochemical orientation of the scissile phosphodiester bond relative to the array of catalytic residues on the enzyme (48,49). Thus one nuclease active site would be unable to hydrolyze DNA with both 3′→5′ and 5′→3′ polarity in that active site. However, if the unwound DNA that is the substrate for the RecBCD nuclease is flexible on the enzyme, then the fact that the DNA is not extensively bound right at the RecB nuclease active site might allow either strand to enter the active site with the orientation necessary for the hydrolysis reaction to be catalyzed. Thus, the single nuclease active site in the 30 kDa domain would be able to cleave whatever DNA strand is presented to it by the other subunits, as proposed previously (50). One possible effect of Chi may then be to control which DNA strand reaches the nuclease active site (51).

A second enzymatic advantage of low binding affinity for the DNA substrate, and more specifically for the cleaved product, in the immediate vicinity of the nuclease active site is to allow RecBCD to be both rapid and also highly processive in its double-strand nuclease reaction. High processivity requires that the enzyme remain bound to the substrate while a rapid reaction requires that product release not be excessively slow. Processive DNA polymerases such as bacterial DNA polymerase III and eukaryotic polymerase δ solve this problem by having a polymerase subunit with relatively low activity and processivity tethered to the DNA by other subunits of the holoenzyme complex (52). Processive nucleases such as exonuclease I and others were proposed some time ago to have an anchor domain to hold onto the DNA that is separate from a catalytic domain where cleavage occurs (53,54). RecBCD seems to have this type of structure, with the ‘anchor’ sites situated in the large N-terminal domain of RecB (21) and in

RecD and/or RecC (55). Finally, the kinetic consequences of tight binding to the substrate and product have been observed with a catalytic RNA derived from the *Tetrahymena* group I intron. Mutations that weakened binding of the ribozyme to its RNA substrate were found to increase the overall RNA endonuclease turnover rate by increasing the rate of product dissociation (56) and they endowed the ribozyme with modest processivity (57).

ACKNOWLEDGEMENTS

We thank Prof. David Jollie and Dr Phil Bryan for their assistance with the CD and absorbance spectrum analysis, Dr Brian Martin for N-terminal sequencing, and Prof. George Lorimer for helpful discussions and the gift of GroEL and GroES proteins. This work was supported by grant no. GM 39777 from the NIH. Julie Zhang was supported by funds from the Center for Biomolecular Structure and Organization, University of Maryland, College Park, MD.

REFERENCES

1. Myers,R.S. and Stahl,F.W. (1994) *Annu. Rev. Genet.*, **28**, 49–70.
2. Kowalczykowski,S.C., Dixon,D.A., Eggleston,A.K., Lauder,S.D. and Rehrauer,W.M. (1994) *Microbiol. Rev.*, **58**, 401–465.
3. Rosamond,J., Telander,K.M. and Linn,S. (1979) *J. Biol. Chem.*, **254**, 8646–8652.
4. Eggleston,A.K. and Kowalczykowski,S.C. (1993) *J. Mol. Biol.*, **231**, 605–620.
5. Dixon,D.A. and Kowalczykowski,S.C. (1995) *J. Biol. Chem.*, **270**, 16360–16370.
6. Taylor,A. and Smith,G.R. (1980) *Cell*, **22**, 447–457.
7. Muskavitch,K.M.T. and Linn,S. (1982) *J. Biol. Chem.*, **257**, 2641–2648.
8. Roman,L.J. and Kowalczykowski,S.C. (1989) *Biochemistry*, **28**, 2863–2873.
9. Goldmark,P.J. and Linn,S. (1972) *J. Biol. Chem.*, **247**, 1849–1860.
10. MacKay,V. and Linn,S. (1974) *J. Biol. Chem.*, **249**, 4286–4294.
11. Dixon,D.A. and Kowalczykowski,S.C. (1991) *Cell*, **66**, 361–371.
12. Dixon,D.A. and Kowalczykowski,S.C. (1993) *Cell*, **73**, 87–96.
13. Taylor,A.F. and Smith,G.R. (1995) *J. Biol. Chem.*, **270**, 24459–24467.
14. Smith,G.R. and Stahl,F.W. (1985) *Bioessays*, **2**, 244–249.
15. Ponticelli,A.S., Schultz,D.W., Taylor,A.F. and Smith,G.R. (1985) *Cell*, **41**, 145–151.
16. Taylor,A.F., Schultz,D.W., Ponticelli,A.S. and Smith,G.R. (1985) *Cell*, **41**, 153–163.
17. Anderson,D.G. and Kowalczykowski,S.C. (1997) *Genes Dev.*, **11**, 571–581.
18. Anderson,D.G. and Kowalczykowski,S.C. (1997) *Cell*, **90**, 77–86.
19. Kuzminov,A. (1995) *Mol. Microbiol.*, **16**, 373–384.
20. Dixon,D.A., Churchill,J.J. and Kowalczykowski,S.C. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 2980–2984.
21. Yu,M., Souaya,J. and Julin,D.A. (1998) *Proc. Natl Acad. Sci. USA*, **95**, 981–986.
22. Taylor,A.F. and Smith,G.R. (1999) *Genes Dev.*, **13**, 890–900.
23. Chaudhury,A.M. and Smith,G.R. (1984) *Proc. Natl Acad. Sci. USA*, **81**, 7850–7854.
24. Amundsen,S.K., Taylor,A.F., Chaudhury,A.M. and Smith,G.R. (1986) *Proc. Natl Acad. Sci. USA*, **83**, 5558–5562.
25. Korangy,F. and Julin,D.A. (1993) *Biochemistry*, **32**, 4873–4880.
26. Palas,K.M. and Kushner,S.R. (1990) *J. Biol. Chem.*, **265**, 3447–3454.
27. Boehmer,P.E. and Emmerson,P.T. (1992) *J. Biol. Chem.*, **267**, 4981–4987.
28. Yu,M., Souaya,J. and Julin,D.A. (1998) *J. Mol. Biol.*, **283**, 797–808.
29. Mach,H., Middaugh,C.R. and Denslow,N. (1995) In Coligan,J.E., Dunn,B.M., Ploegh,H.L., Speicher,D.W. and Wingfield,P.T. (eds), *Current Protocols in Protein Science*. John Wiley & Sons, New York, NY, pp. 7.2.1–7.2.21.
30. Savitzky,A. and Golay,J.E. (1964) *Anal. Chem.*, **36**, 1627–1639.
31. Woody,R.W. (1996) In Fasman,G.D. (ed.), *Circular Dichroism and the Conformational Analysis of Biomolecules*. Plenum Press, New York, NY, pp. 25–67.
32. Minton,A.P. (1997) *Curr. Opin. Biotechnol.*, **8**, 65–69.
33. Thomas,J.G., Ayling,A. and Baneyx,F. (1997) *Appl. Biochem. Biotechnol.*, **66**, 197–238.
34. Thain,A., Gaston,K., Jenkins,O. and Clarke,A.R. (1996) *Trends Genet.*, **12**, 209–210.
35. Ragone,R., Colonna,G., Balestrieri,C., Servillo,L. and Irace,G. (1984) *Biochemistry*, **23**, 1871–1875.
36. Bohm,G., Muhr,R. and Jaenicke,R. (1992) *Protein Eng.*, **5**, 191–195.
37. Johnson,W.C. (1999) *Proteins*, **35**, 301–312.
38. Woody,R.W. and Dunker,A.K. (1996) In Fasman,G.D. (ed.), *Circular Dichroism and the Conformational Analysis of Biomolecules*. Plenum Press, New York, NY, pp. 109–157.
39. Kornberg,A. and Baker,T.A. (1991) *DNA Replication*, 2nd Edn. W.H. Freeman and Co., New York, NY.
40. Aravind,L., Walker,D.R. and Koonin,E.V. (1999) *Nucleic Acids Res.*, **27**, 1223–1242.
41. Bae,S.H., Choi,E., Lee,K.H., Park,J.S., Lee,S.H. and Seo,Y.S. (1998) *J. Biol. Chem.*, **273**, 26880–26890.
42. Shen,J.C., Gray,M.D., Oshima,J., Kamath-Loeb,A.S., Fry,M. and Loeb,L.A. (1998) *J. Biol. Chem.*, **273**, 34139–34144.
43. Li,L., Wu,L.P. and Chandrasegaran,S. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 4275–4279.
44. Colandene,J.D. and Topal,M.D. (1998) *Proc. Natl Acad. Sci. USA*, **95**, 3531–3536.
45. Belfort,M. and Roberts,R.J. (1997) *Nucleic Acids Res.*, **25**, 3379–3388.
46. Derbyshire,V., Kowalski,J.C., Dansereau,J.T., Hauer,C.R. and Belfort,M. (1997) *J. Mol. Biol.*, **265**, 494–506.
47. Pingoud,V., Grindl,W., Wende,W., Thole,H. and Pingoud,A. (1998) *Biochemistry*, **37**, 8233–8243.
48. Gerlt,J.A. (1993) In Linn,S.M., Lloyd,R.S. and Roberts,R.J. (eds), *Nucleases*. 2nd Edn. Cold Spring Harbor Laboratory Press, Plainview, NY, pp. 1–34.
49. Beese,L.S. and Steitz,T.A. (1991) *EMBO J.*, **10**, 25–33.
50. Chen,H.-W., Randle,D.E., Gabbidon,M. and Julin,D.A. (1998) *J. Mol. Biol.*, **278**, 89–104.
51. Anderson,D.G. and Kowalczykowski,S.C. (1998) *J. Mol. Biol.*, **282**, 275–285.
52. Kelman,Z. and O'Donnell,M. (1995) *Nucleic Acids Res.*, **23**, 3613–3620.
53. Thomas,K.R. and Olivera,B.M. (1978) *J. Biol. Chem.*, **253**, 424–429.
54. Brody,R.S., Doherty,K.G. and Zimmerman,P.D. (1986) *J. Biol. Chem.*, **261**, 7136–7143.
55. Ganesan,S. and Smith,G.R. (1993) *J. Mol. Biol.*, **229**, 67–78.
56. Young,B., Herschlag,D. and Cech,T.R. (1991) *Cell*, **67**, 1007–1019.
57. Herschlag,D. (1992) *Biochemistry*, **31**, 1386–1399.