

# Catalytic and binding mutants of the junction-resolving enzyme endonuclease I of bacteriophage T7: role of acidic residues

M. Janine Parkinson, J. Richard G. Pöhler and David M. J. Lilley\*

CRC Nucleic Acid Structure Research Group, Department of Biochemistry, The University of Dundee, Dundee DD1 4HN, UK

Received October 30, 1998; Revised and Accepted November 13, 1998

## ABSTRACT

**Endonuclease I is a 149 amino acid protein of bacteriophage T7 that is a Holliday junction-resolving enzyme, i.e. a four-way junction-selective nuclease. We have performed a systematic mutagenesis study of this protein, whereby all acidic amino acids have been individually replaced by other residues, mainly alanine. Out of 21 acidic residues, five (Glu20, Glu35, Glu65, Asp55 and Asp74) are essential. Replacement of these residues by other amino acids leads to a protein that is inactive in the cleavage of DNA junctions, but which nevertheless binds selectively to DNA junctions. The remaining 16 acidic residues can be replaced without loss of activity. The five critical amino acids are located within one section of the primary sequence. It is rather likely that their function is to bind one or more metal ions that coordinate the water molecule that brings about hydrolysis of the phosphodiester bond. We have also constructed a mutant of endonuclease I that lacks nine amino acids (six of which are arginine or lysine) at the C-terminus. Unlike the acidic point mutants, the C-terminal truncation is unable to bind to DNA junctions. It is therefore likely that the basic C-terminus is an important element in binding to the DNA junction.**

## INTRODUCTION

Four-way (Holliday) DNA junctions are central intermediates in recombination processes (1–10). These branched DNA intermediates are ultimately resolved back to duplex DNA species, requiring a class of junction-resolving enzymes (reviewed in 11). They are important enzymes in processes of recombination and repair of DNA and have been isolated from bacteriophage-infected eubacteria (12,13), *Escherichia coli* (14–17), yeast (18–21) and mammalian cells (22,23) and their viruses (24). They are probably ubiquitous cellular enzymes. The junction-resolving enzyme of bacteriophage T7 is encoded by gene 3. Mutants in this gene are deficient in recombination (25,26) and accumulate branched DNA intermediates (27). The product of gene 3 was found to be a 149 amino acid endonuclease (28–30) that

introduces cleavages into branched DNA species including four-way junctions (13,31,32). A functionally related enzyme called endonuclease VII is encoded by phage T4 (12,33).

Junction-resolving enzymes are nucleases with a strong substrate specificity for branched DNA. There are thus two aspects to their properties. First, they exhibit a marked specificity for binding DNA junctions; complexes with four-way DNA junctions are typically not displaced by a 1000-fold excess of linear DNA of the same sequence (34–36). Second, they are nucleases that catalyse the hydrolysis of the phosphodiester bond. In order to bring about a resolution of a four-way junction, these enzymes invariably bind in dimeric form (35–39) and introduce paired cleavages in the DNA. The two cleavages occur within the lifetime of the enzyme–junction complex in the known cases of T7 endonuclease I (37) and T4 endonuclease VII (40). The binding event positions the active sites of the two subunits correctly on the junction such that the required cleavage is possible. This step also provides the structural selectivity of the process, as well as distorting the global structure of junctions in most cases (34,35,39,41–43) for reasons that are not yet apparent. The binding event can occur normally in the absence of cleavage. Thus inactive mutants of these enzymes normally bind to DNA junctions with almost unaltered affinity, as exemplified by T4 endonuclease VII (44). Furthermore, incubation of DNA junctions with heterodimeric phage resolving enzymes consisting of active and inactive (mutant) subunits leads to unilateral cleavage (37,40), showing that one subunit can cleave normally while the other is non-functional. T7 endonuclease I has been shown to cleave linear DNA when delivered to the substrate as a fusion with the *lac* repressor (45). The cleavage reaction can add a further level of specificity to the process, because some enzymes, exemplified by RuvC (46), CCE1 (36,47) and RusA (39,48), exhibit very significant sequence preferences. However, the differences in cleavage rates between junction sequences are not reflected in variation in the affinity of binding, indicating that they result from altered interactions with the transition state, rather than the ground state of the reaction. This again points to a separation between the basic binding of the protein to the junction and the cleavage of the DNA once bound, suggesting that the parts of the proteins involved in DNA cleavage are typical deoxynucleases.

\*To whom correspondence should be addressed. Tel: +44 1382 344243; Fax: +44 1382 201063; Email: dmjlilley@bad.dundee.ac.uk

One or more metal ions are frequently involved in phosphodiester bond hydrolysis by nucleases (49). In staphylococcal nuclease for example, a metal ion provides the hydrolytic OH<sup>-</sup> nucleophile in its coordination sphere, as well as stabilising developing negative charge in the transition state (49). In other nucleases, such as *E. coli* DNA polymerase I exonuclease (50), two separate metal ions carry out these functions. The binding sites for such catalytic metal ions frequently comprise a number of aspartate and/or glutamate residues (51–55) and acidic amino acids essential for nuclease activity have been identified in a number of junction-resolving enzymes, including T4 endonuclease VII (35,44), RuvC (56) and RusA (39). Random chemical mutagenesis of the gene encoding T7 endonuclease I led to the identification of a number of residues required for cleavage activity (34), including three acidic amino acids, namely Glu35, Glu65 and Asp55. To get a clearer position of the role of acidic residues in the phosphodiester cleavage mechanism we decided to change all the glutamate and aspartate residues systematically by converting them individually to alanine. This analysis has revealed two further acidic amino acids essential for DNA cleavage, a total of five altogether, and 16 non-essential residues. In the course of these studies we have also made an additional mutant form of endonuclease I in which the basic C-terminal nine amino acid peptide has been deleted. This mutant protein can be overexpressed to high levels, but can neither bind nor cleave DNA junctions. It is therefore probable that this positively charged region of the protein, in contrast to the acidic residues, is involved directly in binding the substrate DNA junctions.

## MATERIALS AND METHODS

### Preparation of single-stranded DNA for mutagenesis

The gene encoding T7 endonuclease I was excised from the plasmid pK19*endoI* by restriction cleavage using *EcoRI* and *BamHI* and ligated into pUC119. The resulting plasmid was transformed into *E. coli* strain CJ236 (*dut ung*) cells. Single-stranded DNA was prepared by transfection of a 100 ml culture of CJ236 containing pUC119*endoI* in LB supplemented with antibiotics after transfection by helper phage M13K07 at a multiplicity of infection of 10. This was incubated at 37°C overnight with vigorous shaking to ensure good aeration. The culture was harvested by centrifugation at 12 000 g for 15 min at 4°C. Phage were precipitated by addition of 0.25 vol. of 3.75 M ammonium acetate (pH 8.0), 20% polyethylene glycol (mol. wt 8000) and left on ice for 60 min. The pellet was resuspended in 400 µl of 10 mM Tris–HCl (pH 8.0), 1 mM EDTA and the phage lysed by the addition of 400 µl of chloroform:isoamyl alcohol (24:1). After vortexing for 1 min the solution was centrifuged for 5 min in an Eppendorf centrifuge and the upper aqueous phase was extracted with phenol:chloroform until there was no material visible at the interface. Finally, the aqueous phase was extracted with 400 µl of chloroform:isoamyl alcohol and the DNA was recovered by precipitation with ethanol.

### Site-directed mutagenesis

Mutagenic oligonucleotides were designed to effect the change of selected amino acids to alanine residues or a termination codon. Aliquots of 10 pmol of the oligonucleotides were phosphorylated and hybridised to ~1 µg of single-stranded pUC119*endoI* in 20 mM Tris–HCl (pH 7.5), 10 mM MgCl<sub>2</sub>, 1 mM DTT, 50 mM

NaCl by heating to 90°C followed by slow cooling to room temperature. Repair synthesis and ligation were performed in 10 µl of 20 mM Tris–HCl (pH 7.5), 10 mM MgCl<sub>2</sub>, 5 mM DTT, 5 mM dNTPs, 5 U T4 DNA ligase and 3 U T4 DNA polymerase for 5 min on ice, 5 min at room temperature, 2 h at 37°C and finally overnight at room temperature. The mixture was transformed (57) into *E. coli* DH5αF' (*dut<sup>+</sup> ung<sup>+</sup>*) and grown overnight in LB at 37°C. All the genes were sequenced to confirm the presence of the required mutations, using primer extension–dideoxy sequencing (58). Mutant genes were excised from pUC119 and religated into pK19.

### Examination of junction-cleavage activity in cell lysates

Mutated endonuclease I was expressed as N-terminal protein A fusions from the vector pK19 transformed in JM101. Cultures (5 ml) in LB were grown at 37°C to A<sub>600</sub> = 0.6 and then induced by addition of 0.5 mM IPTG and growth continued for 2 h. Cells were harvested by centrifugation in an Eppendorf centrifuge for 2 min. Cells were resuspended in 120 µl 50 mM Tris–HCl (pH 7.6), 150 mM NaCl, 1 µl/ml pepstatin and 2 mM pefabloc and sonicated for 3 bursts of 5 s (with resuspension on ice between bursts). After sonication the suspension was centrifuged in an Eppendorf centrifuge for 2 min to remove cell debris and the supernatant was used in the cleavage assay. Aliquots of 2 and 4 µl of each lysate were incubated with radioactive DNA junction in 50 mM Tris–HCl (pH 7.6), 100 mM NaCl, 10 mM MgCl<sub>2</sub>, 1 mM DTT for 10 min at room temperature and the reaction stopped by addition of formamide, heating to 90°C and rapid cooling. The samples were loaded onto a 10% sequencing gel in 90 mM Tris–borate (pH 8.3), 10 mM EDTA, 7 M urea and electrophoresed at a power warming the plates to be hot to the touch. The radioactive DNA was visualised by autoradiography at –70°C using Fuji RX X-ray film with Ilford fast tungstate intensifier screens.

### Examination of protein expression in induced cells

Cultures (5 ml) of cells transformed with plasmids expressing endonuclease I mutants as N-terminal protein A fusions were grown in LB, induced with 0.5 mM IPTG and grown for 2 h. Cells were harvested by centrifugation in an Eppendorf centrifuge for 2 min and the cells resuspended in 80 µl of 100 mM Tris–HCl (pH 7), 1% (w/v) SDS, 5% (v/v) glycerol, 10 mM DTT, 0.01% Bromophenol blue. The samples were boiled for 5 min and 4 µl were loaded onto a 12.5% polyacrylamide gel containing SDS. After electrophoresis proteins were visualised by staining in 0.25% (w/v) Coomassie blue R-250, 50% (v/v) methanol, 10% (v/v) acetic acid and the gel air dried.

### Recloning of endonuclease I E20A as an oligohistidine fusion

Plasmid pK19*endoIE20A* was used as a template for PCR, using primers 5'-CTAGTTATGCTGGGATCCATGGCAGGTTAC-3' and 5'-ACGCTCGTCGCTAAGCTTTTATTTCTTTCC-3' in order to incorporate *BamHI* and *HindIII* restriction sites at the 5'- and 3'-ends respectively. PCR was performed using the proofreading thermostable Pfu DNA polymerase. After digestion, the amplified fragment was ligated into the *BamHI* and *HindIII* sites of pQE30 (Qiagen) and transformed into *E. coli* strain DH5αF'. Transformants were characterised by restriction digestion.

## Overexpression and preparation of N-hexahistidine-endonuclease I E20A

Expression plasmids prepared as above were retransformed into *E. coli* strain M15 (pREP4). One litre cultures were grown to an  $A_{660}$  of 0.6, whereupon expression was induced by the addition of IPTG to a final concentration of 0.5 mM and growth for a further 2 h. Cells were harvested by centrifugation at 4000 g for 10 min at 4°C and the pellet resuspended in 10 ml of 50 mM sodium phosphate (pH 7.0), 200 mM NaCl. Cells were lysed by sonication and the lysate clarified by centrifugation at 15 000 g for 15 min at 4°C. The cleared lysate was applied to a Ni-NTA-agarose column (Qiagen) and washed extensively with 50 mM sodium phosphate (pH 7.0), 200 mM NaCl, 10 mM imidazole (wash buffer). Bound protein was eluted stepwise from 50 to 250 mM imidazole in wash buffer, in 100 mM steps. The N-hexahistidine-endonuclease I E20A eluted in the last step. Peak fractions were identified using a qualitative Bradford protein assay (BioRad), pooled and concentrated using a 10 kDa cut-off concentrator (Millipore). Protein concentration was calculated by a quantitative Bradford assay, using BSA as standard.

### Assembly of DNA junctions

All junctions used were based on the sequence of junction 3 (59).

*Electrophoretic retardation analysis.* Junctions were assembled from four strands, each of 30 nt, to generate a junction of 15 bp/arm.

*Comparative gel electrophoresis.* Each of the six species shortened in two arms were assembled from four oligonucleotides (of 80, 55 or 30 bp in length). For example, species BH (comprising long B and H arms and short R and X arms) was assembled from a b strand of 80 nt, h and x strands of 55 nt and an r strand of 30 nt.

*Analysis of cleavage by endonuclease I.* Junctions were assembled from four strands, each of 50 nt, to generate a junction of 25 bp/arm. The h strand was 5'-<sup>32</sup>P-labelled in these experiments.

Junctions were assembled by mixing stoichiometric amounts of each strand, one or more of which were radioactively 5'-<sup>32</sup>P-labelled. This was incubated at 70°C for 3 min followed by slow cooling to room temperature. Junction DNA was purified by electrophoresis in a 5% polyacrylamide gel, followed by excision of bands and electroelution.

### Electrophoretic mobility retardation analysis

Junction DNA (1 nM) was incubated with various amounts of protein in 50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1 mM DTT, 100 µg/ml BSA and either 10 mM MgCl<sub>2</sub> or 5 mM EDTA in a 10 µl reaction volume. After incubation at 25°C, loading dye (final concentration 5% w/v Ficoll 400, 0.25% bromophenol blue, 0.25% xylene cyanol) was added and the entire sample loaded onto a pre-running 6% polyacrylamide gel in 90 mM Tris-borate (pH 8.3), 10 mM EDTA. Electrophoresis was performed at 90 V for 16 h. The gel was dried onto Whatman 3MM paper followed by autoradiography as above.

### Comparative gel electrophoresis of junction conformation

The six species used for comparative gel analysis were each used at 1 nM concentration. N-hexahistidine-endonuclease I E20A was added to a final concentration of 500 nM, to ensure the

presence of both free and protein-bound species. The binding conditions used were essentially the same as above, except that 500 nM (calculated for nucleotide concentration) calf thymus competitor DNA was included in the binding reaction to minimise non-specific complex formation. Samples were applied to 9% polyacrylamide gels and electrophoresed in 90 mM Tris-borate (pH 8.3), 5 mM EDTA at 110 V for 16 h. Gels were dried onto Whatman 3MM paper followed by autoradiography as above.

## RESULTS

### Systematic alanine scanning mutagenesis of the acidic amino acid residues of T7 endonuclease I

In order to assess the role of acidic amino acids in the DNA cleavage reaction of endonuclease I we have systematically mutated each aspartate and glutamate (in addition to those previously identified; 34) in turn, replacing them with alanine. Site-directed mutagenesis was carried out using the Kunkel method (60). The gene encoding endonuclease I (T7 gene 3) was cloned into the *Bam*HI and *Eco*RI sites of pUC119 (61) to facilitate the mutagenesis. The resulting plasmid pUC119-endonuclease I was transformed into *E. coli* CJ236 (*dut ung*) and single-stranded, uracil-laden plasmid DNA produced. Primers containing the relevant mutation were used to prime second strand synthesis using the normal four deoxynucleotide triphosphates. The resulting plasmid was transformed into *E. coli* strain JM101 (*dut<sup>+</sup> ung<sup>+</sup>*), with consequent degradation of the parental strand and repair DNA synthesis from the mutated strand giving a substantial enrichment of the mutation. Transformants were screened for the presence of the relevant mutation by sequencing. Mutated endonuclease I genes were subcloned into pK19PRA (62) and transformed in *E. coli* JM101 for expression as an N-terminal fusion with protein A and again verified by sequencing. The mutations that were introduced into endonuclease I are listed in Table 1.

### Analysis of the nucleolytic activity of endonuclease I mutants

Each mutant of endonuclease I was expressed as a fusion with protein A in *E. coli* strain JM101. We have shown previously that N-terminal fusions with endonuclease I do not interfere with junction-cleavage activity of the enzyme (34). The activity of the mutant proteins was initially assayed directly from the induced cell lysate. Each strain was grown to an  $A_{600}$  of 0.6, whereupon expression was induced from the *lac* promoter by the addition of IPTG. After 2 h induction, the culture was harvested and lysed by sonication. The lysate was cleared by centrifugation and the supernatant used to cleave radioactive four-way junction DNA.

Figure 1 shows the results of the analysis of endonuclease I activity from the cell lysates. Where active endonuclease I is present in the extract, a clear product band is evident. Incubation with the extract from cells expressing the previously characterised endonuclease I E65K (34) gives no such product, confirming that this mutation leads to inactivity of junction resolution. From these results it is apparent that two further endonuclease I mutants, E20A and D74A, have lost the ability to cleave junction DNA in this assay. However, selective binding to DNA four-way junctions is retained by both proteins. The remaining 16 acidic amino acids have each been mutated to alanine with no loss of cleavage activity or specificity. These data are summarised in Table 1.



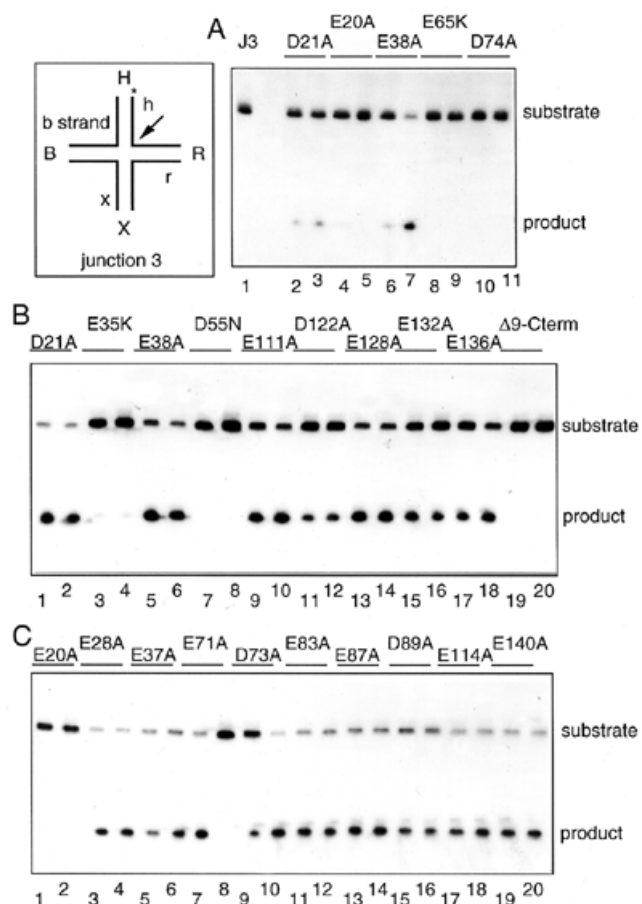
**Table 1.** Summary of the properties of the endonuclease I mutants constructed

Mutation	Cleavage of junction 3	Binding to junction
E20A	Inactive	Yes
D21A	Active	
E28A	Active	
E35K <sup>a</sup>	Inactive	Yes
E35Q	Inactive	
E37A	Active	
E38A	Active	
D55N <sup>a</sup>	Inactive	Yes
E65K <sup>a</sup>	Inactive	Yes
E65Q	Very low activity	
E71A	Active	
D73A	Active	
D74A	Inactive	Yes
E83A	Active	
E87A	Active	
D89A	Active	
E111A	Active	
E114A	Active	
D122A	Active	
E128A	Active	
E132A	Active	
E136A	Active	
D140A	Active	
Δ9C-term	Inactive	No

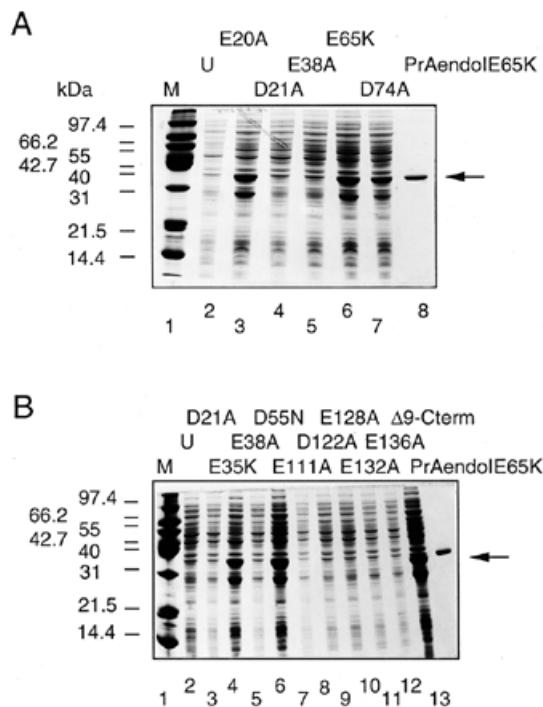
All are point mutants except for the final C-terminal 9 amino acid deletion mutant (Δ9C-term). The activity is taken from the assay of junction 3 cleavage using the cell lysates containing the N-terminal protein A fusion proteins and no attempt to quantify the level of activity has been made. The binding to four-way junctions has only been assessed for the inactive mutants and the absence of an entry in the third column only signifies that the experiment has not been performed.

<sup>a</sup>Mutants were isolated previously by random chemical mutagenesis with selection for loss of toxicity when expressed in *E.coli* (34).

Of the mutants isolated in the original random chemical mutagenesis (34), two involved changes from acidic to basic residues, i.e. E35K and E65K. This is an extreme change of character and we therefore asked whether smaller changes would also lead to inactivation of junction-cleavage. We therefore made the mutant endonuclease I E35Q, where the carboxylic acid side chain of the glutamate has been changed to the corresponding amide, glutamine. This mutant was found to be completely inactive in the cleavage of four-way junctions and thus we conclude that Glu35 provides an essential carboxylate moiety. Some cleavage of DNA junctions was observed with endonuclease I E65Q, but this level was very much lower than that found for other active forms of the enzyme.



**Figure 1.** Analysis of the junction-cleavage activity of T7 endonuclease I mutants in cell lysates. *Escherichia coli* cells expressing endonuclease I mutants as N-terminal protein A fusions were induced by addition of IPTG and cell lysates prepared by sonication. In each case 2 and 4 μl aliquots were individually incubated with four-way DNA junction 3 radioactively 5'-<sup>32</sup>P-labelled on the h strand (as shown in the insert) and the products analysed by separation on a polyacrylamide gel and autoradiography. The band arising from cleavage of the h strand by endonuclease I junction-resolving activity is indicated right as product. (A) Analysis of lysates from strains carrying five different mutants of endonuclease I. This includes the previously characterised endonuclease I E65K. Track 1, junction 3 not incubated with protein, showing the full-length h strand; tracks 2 and 3, 4 and 5, 6 and 7, 8 and 9, 10 and 11, incubation with D21A, E20A, E38A, E65K and D74A, respectively; even numbered tracks correspond to 2 μl aliquots and odd numbered tracks correspond to 4 μl aliquots of cell lysates. The cleavage by endonuclease I D21A is rather faint on this autoradiograph, but this activity is reproducible (tracks 1 and 2 in part B). (B) Analysis of lysates from strains carrying further mutants of endonuclease I, including deletion of the C-terminal nine amino acids (Δ9C-term). This includes the previously characterised endonuclease I E35K and D55N. Tracks 1 and 2, 3 and 4, 5 and 6, 7 and 8, 9 and 10, 11 and 12, 13 and 14, 15 and 16, 17 and 18, 19 and 20, incubation with D21A, E35K, E38A, D55N, E111A, D122A, E128A, E132A, E136A and Δ9C-term, respectively; odd numbered tracks correspond to 2 μl aliquots and even numbered tracks correspond to 4 μl aliquots of cell lysates. (C) Analysis of lysates from strains carrying further mutants of endonuclease I. Tracks 1 and 2, 3 and 4, 5 and 6, 7 and 8, 9 and 10, 11 and 12, 13 and 14, 15 and 16, 17 and 18, 19 and 20, incubation with E20A, E28A, E37A, E71A, D73A, E83A, E87A, D89A, E114A and D140A, respectively; odd numbered tracks correspond to 2 μl aliquots and even numbered tracks correspond to 4 μl aliquots of cell lysates. For an unknown reason, digestion with 4 μl of the extract containing endonuclease I E71A (track 8) failed to generate junction cleavage, however, the 2 μl incubation clearly shows good activity.



**Figure 2.** Overexpression of inactive endonuclease I mutants. *Escherichia coli* cells expressing endonuclease I mutants as N-terminal protein A fusions were induced and cells were harvested, resuspended and boiled. The proteins present in the extract were examined by electrophoresis in polyacrylamide gels containing SDS. The position of the full-length protein A–endonuclease I mutants is indicated by the arrows (right). In the overexpressing strains some proteolytic degradation products can be seen in addition to the full-length 302 amino acid fusion. (A) Track 1, mixture of proteins to serve as molecular weight markers (molecular weights shown on left); track 2, uninduced cells transformed with endonuclease I E20A; tracks 3–7, induced cells transformed with endonuclease I E20A, D21A, E38A, E65K and D74A, respectively; track 8, purified protein A–endonuclease I E65K. (B) Track 1, mixture of proteins to serve as molecular weight markers (molecular weights shown on left); track 2, uninduced cells transformed with endonuclease I E35K; tracks 3–12, induced cells transformed with endonuclease I D21A, E35K, E38A, D55N, E111A, D122A, E128A, E132A, E136A and Δ9C-term, respectively; track 13, purified protein A–endonuclease I E65K.

### Inactive endonuclease I mutants enhance the rate of cell growth and the yield of protein

In the course of preparing the cells for the assays of junction-cleavage activity, it was observed that the strains expressing the two new inactive mutants displayed enhanced growth when compared with those expressing active endonuclease I. This is in good agreement with the previously reported toxicity of bacteriophage junction-resolving enzymes when overexpressed in *E. coli* (34,44). The reduction in toxicity of the inactivating mutants is also reflected in higher levels of expression of endonuclease I for these mutants, shown by SDS–PAGE of the proteins present in the cell lysates (Fig. 2). While the cells expressing inactive mutant endonuclease I (E20A, E65K and D74) give bands corresponding to significantly overexpressed protein, this is absent in the lysates from cells expressing active endonuclease I (D21A and D38A). We also found similar overexpression from strains expressing the other inactive mutants, E35K and D55N, while none of the strains expressing the active mutants gave overexpression detectable by gel electrophoresis (data not shown, in some cases).

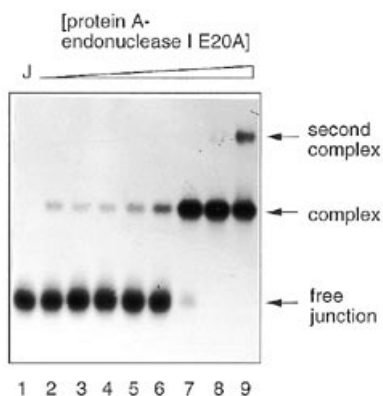
### Analysis of the DNA-binding activity of endonuclease I E20A

The systematic mutagenesis has led to the isolation of two new inactive mutants and we have characterised the DNA binding properties of one of these in greater depth. The inactive endonuclease I mutant proteins were initially produced as N-terminal fusions with protein A of *Staphylococcus aureus*. However, purification of the fusion proteins was hampered by proteolysis during the purification procedure. The presence of fusion-released mutant endonuclease I would hamper further analysis of DNA binding as the protein forms a strongly associated dimer (37), resulting in a heterogeneous mixture of fusion–non-fusion protein dimers in the reaction. Further characterisation of one mutant of endonuclease I, E20A, was therefore aided by expressing it as an N-terminal oligohistidine fusion. The endonuclease I E20A gene was amplified by PCR using primers that would incorporate unique *Bam*HI and *Hind*III restriction sites. After restriction digestion with those enzymes, the DNA fragment including the endonuclease I E20A gene was ligated into pQE30 (Qiagen), allowing subsequent expression and purification as an oligohistidine-fused protein.

N-hexahistidine–endonuclease I E20A was analysed for its ability to bind a four-way DNA junction using gel electrophoretic retardation analysis. For this purpose we used a non-migratory four-way DNA junction with arms each of 25 bp with the sequence of junction 3 (59). Radioactively  $5'$ - $^{32}$ P-labelled junction was incubated with increasing amounts of N-hexahistidine–endonuclease I E20A and analysed by non-denaturing polyacrylamide gel electrophoresis in the absence of added metal ions (Fig. 3). At moderate protein concentrations, a single discrete DNA–protein complex is formed, with a second complex being formed at higher protein concentrations. Previous studies indicated that this second complex is less specific and can be readily competed by duplex DNA (data not shown), in contrast to the first complex formed with endonuclease I, which is resistant to competition by a 1000-fold molar excess of duplex DNA (34). In general the binding of endonuclease I to junctions is cooperative and it is not easy to measure an apparent dissociation constant ( $K_d$ ) with accuracy. From these data, an apparent  $K_d$  in the region of 500 nM has been estimated, indicating that the affinity has been lowered compared with the value for endonuclease I E65K (34). However, the fusion adduct used in this case is different, which may well affect the binding affinity.

### Global structure of the four-way DNA junction bound by endonuclease I E20A

It has been previously determined that the four-way DNA junction undergoes a folding transition in the presence of micromolar concentrations of divalent metal ions such as magnesium (59,63). The global structure changes from an extended, square structure in the absence of magnesium to a right-handed, antiparallel, crossed structure in the presence of  $>80$   $\mu$ M magnesium ions, called the stacked X-structure. Upon binding of proteins, however, the junction can be unfolded and manipulated into novel structures determined by the protein. The global structures of the junction have been determined for complexes with all of the known junction-resolving enzymes, including T4 endonuclease VII (35), RuvC (41), RusA (39,43), CCE1 (42) and YDC2 (64), and other proteins, including the junction-binding protein RuvA (65), site-specific recombinases of the integrase family such as Cre (66) and HMG box proteins (67). Disruption of the folded state of the four-way junction

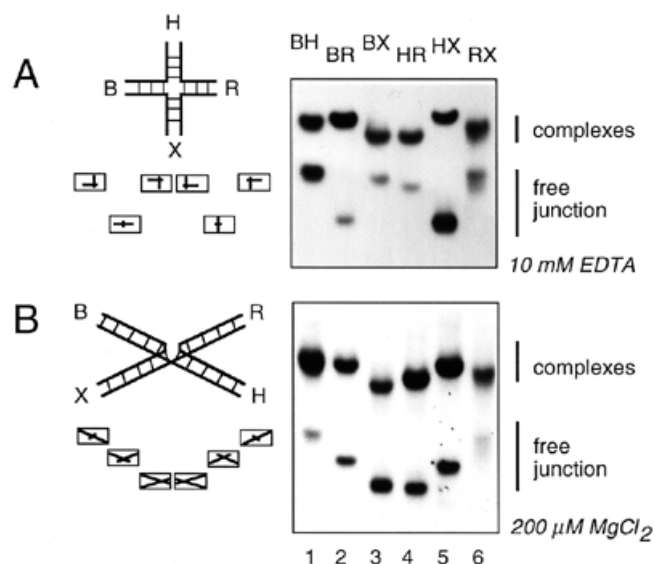


**Figure 3.** Binding of N-hexahistidine-endonuclease IE20A to a four-way DNA junction. Radioactively  $5'$ - $^{32}\text{P}$ -labelled junction (1 nM) was incubated with increasing amounts of N-hexahistidine-endonuclease IE20A and analysed by non-denaturing polyacrylamide gel electrophoresis in the absence of added metal ions. The complex migrates as a discrete retarded species. At high protein concentrations a second, more retarded complex is observed. Track 1, no added protein; track 2, 29.4 nM protein; track 3, 58.8 nM protein; track 4, 117.5 nM protein; track 5, 235 nM protein; track 6, 470 nM protein; track 7, 940 nM protein; track 8, 1.38  $\mu\text{M}$  protein; track 9, 3.76  $\mu\text{M}$  protein.

appears to be common to all of these proteins, even though the final global conformations differ in most cases.

The first example of alteration of global structure in the four-way junction by resolving enzymes was in fact found with T7 endonuclease I (59), using both wild-type enzyme and the inactive E65K mutant. It was found that the protein binding changed the global folding of the DNA in both the presence (E65K mutant only) and absence (both proteins) of magnesium ions. We therefore decided to examine the effect of binding the new mutant endonuclease VII E20A. The six possible subspecies of junction 3 with two long (40 bp) and two short (15 bp) arms were radioactively  $5'$ - $^{32}\text{P}$ -labelled. Each was incubated with 500 nM N-hexahistidine-endonuclease I E20A and analysed by polyacrylamide gel electrophoresis under non-denaturing conditions. Figure 4A shows the pattern of the protein-bound and free junction in the absence of added metal ions. The pattern of electrophoretic mobilities is very different from the free junction, which gives the slow, fast, slow, slow, fast, slow pattern indicative of the extended square structure under these conditions. Thus the protein has clearly imposed a new global structure on the junction. In the presence of 200  $\mu\text{M}$  magnesium ions the free junction gives the slow, intermediate, fast, fast, intermediate, slow pattern showing that it has folded into the stacked X-structure in the *isoII* conformer (based on B on X arm stacking), however, the electrophoretic mobility pattern exhibited by the complex is different and is the same as that found in the absence of metal ions. Comparison with the patterns obtained with complexes containing wild-type endonuclease I and endonuclease I E65K (59) show these to be identical. Thus endonuclease VII E20A imposes the same structural distortion on the DNA under the same conditions, despite the different mutation present. This indicates that the structural distortion on binding may occur in the absence of phosphodiester cleavage.

One further conclusion may be deduced from this experiment. It should be noted that the incubations of N-hexahistidine-endonuclease I E20A with the various long-short arm DNA junctions contained a large excess of calf thymus competitor DNA. That



**Figure 4.** Manipulation of the global structure of a DNA junction by N-hexahistidine-endonuclease IE20A.  $5'$ - $^{32}\text{P}$ -radioactively labelled junction 3 was assembled from four oligonucleotides to generate the six species with two long and two short arms. The four arms are named B, H, R and X as used in the original studies of this junction (59). The six two-long, two-short arm species were each incubated with N-hexahistidine-endonuclease IE20A and analysed by electrophoresis on an 8% polyacrylamide gel and by autoradiography. (A) The complex of junction 3 with N-hexahistidine-endonuclease IE20A in the absence of added metal ions. The six two-long, two-short arm species generated from junction 3 were incubated with N-hexahistidine-endonuclease IE20A in the presence of 10 mM EDTA. For each of the long-short arm species, two radioactive bands are evident, corresponding to free DNA and complex. The free junction migrates in a slow, fast, slow, slow, fast, slow pattern indicative of the extended, square structure (shown right). The RX species is partially diffuse in the free junction in this experiment. The junction-protein complex migrates in a different pattern, closely similar to that observed previously for wild-type endonuclease I and the E65K mutant (34). (B) The complex of junction 3 with N-hexahistidine-endonuclease I E20A in the presence of magnesium ions. This experiment was performed analogously, except that binding and electrophoresis were performed in the presence of 200  $\mu\text{M}$  magnesium ions. As before, bands corresponding to free and bound junctions are observed for each long-short arm species. The free DNA exhibits the slow, intermediate, fast, fast, intermediate, slow pattern indicative of the stacked X-structure with B on X coaxial stacking (shown right) (59). The structure of this stacking conformer gives rise to the three pairs of long-short arm species in which the included angles between the long arms are acute (BH and RX are slow species), obtuse (BR and HX are intermediate species) or linear (BX and HR are fast species). The pattern of mobilities for the long-short species of the junction-protein complex is clearly different from that of the free DNA, indicating a protein-induced change in the global conformation of the junction. The pattern is closely similar to that observed in the absence of magnesium ions (B), showing that the same global structure is induced by protein binding, irrespective of prevailing ion concentrations.

this excess of mainly linear DNA failed to displace the protein from the junction demonstrates the great structural selectivity of the endonuclease I mutant for the four-way DNA junction. This is typical of the junction-resolving enzymes in general.

#### Truncation of the C-terminal peptide of endonuclease I

We have noted that the C-terminal nine amino acids form a particularly basic peptide, with the sequence RLKRKGGKK, which might conceivably have an important role in DNA binding. A further mutation was therefore introduced in which the codon for Arg141 was replaced by a termination codon. This mutant



(endonuclease I  $\Delta$ 9C-term) was expressed as an N-terminal protein A fusion analogously to the point mutants above. Like the inactive point mutants it was overexpressed on induction with IPTG (Fig. 2B) and showed no detectable cleavage activity with junction 3 (Fig. 1B). Unlike all the point mutants analysed, however, the C-terminal deletion mutant (as an N-terminal protein A fusion) exhibited no binding to DNA junctions that was detectable by electrophoretic retardation (data not shown).

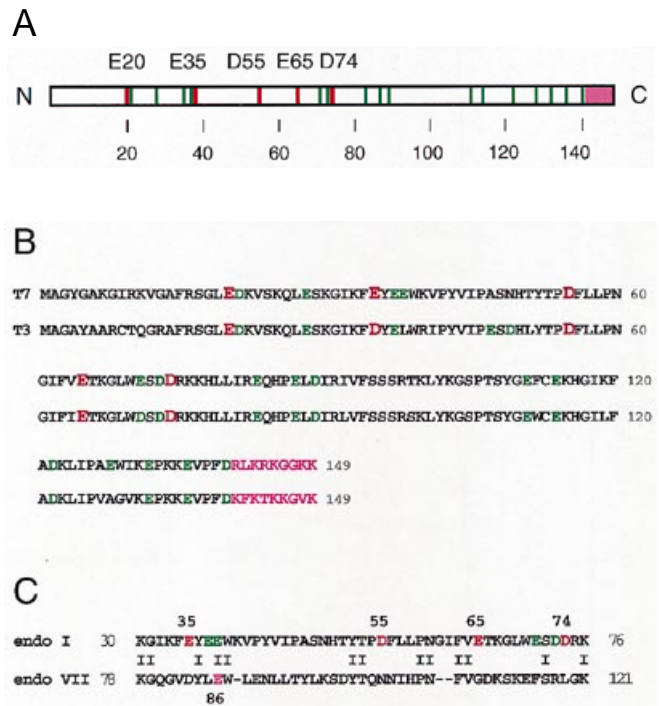
## DISCUSSION

Of the 21 acidic amino acids of T7 endonuclease I we have identified five that are essential for the cleavage of four-way DNA junctions. The remaining 16 residues can be changed to alanine without detectable loss of activity in cell lysates. It is likely that at least some of the five essential acidic amino acid side chains are involved in the coordination of a metal ion(s), providing a hydrolytic water molecule in the coordination sphere that is required for breakage of the phosphodiester bond. This is probably a common mechanism in the junction-resolving enzymes, since essential acidic amino acids have been found in T4 endonuclease VII (35,44), RuvC (56), RusA (39) and the yeast resolving enzymes CCE1 and YDC2 (M.F.White and D.M.J.Lilley, unpublished data).

Despite the complete inability of these mutant proteins to cleave DNA junctions, all the mutants tested retain the selective binding of four-way junctions and manipulate the conformation of the junction in the same manner. Retention of binding by catalytically incompetent mutants again seems to be a common property of the junction-resolving enzymes (35,39,44) and shows that binding can proceed in the absence of an ability to cleave junctions, i.e. that the two events are to some extent separable.

Both T7 endonuclease I and T4 endonuclease VII are highly toxic when overexpressed in *E.coli*, in contrast to cellular enzymes such as RuvC, CCE1 and YDC2. The toxicity appears to arise from the activity of these enzymes and mutations that lead to loss of junction-cleavage activity result in loss of toxicity. In the case of the mutants isolated in this study, the correlation between the loss of activity and overexpression in *E.coli* is total. This leaves the question of why some junction-resolving enzymes may be readily overexpressed, while others (the phage T7 and T4 enzymes especially) cannot. What distinguishes the enzymes of phage T7 and T4 is the low degree of sequence specificity in their cleavage reactions, compared with other junction-resolving enzymes such as RusA (39,48), RuvC (68) and CCE1 (36,47). A high degree of sequence selectivity for the cleavage process has the effect of restricting the cleavage to junctions that can undergo branch migration and thereby test many different sequences. In contrast, the lower sequence selectivity of the phage enzymes releases them to be much more promiscuous in their range of substrates, in keeping with their role in the repair of replicated phage DNA, and the resulting assault on the cellular DNA is likely to prove toxic to the bacterial cell when overexpressed. Indeed, this is part of their function in late phage infection.

The location of the five essential acidic residues in the primary sequence of endonuclease I is shown in Figure 5A. It is apparent that all five are located in the N-terminal half of the protein, suggesting that the folding of this section generates the active site for DNA cleavage. Phage T3 encodes a protein that is 79% identical (69) and is almost certainly the homologue of endonuclease I. Alignment shows that four of the five essential residues are identical in the two



**Figure 5.** The location of catalytically essential acidic residues in T7 endonuclease I. (A) Linear map of endonuclease I showing the location of essential acidic amino acids (red), inessential acidic amino acids (green) and the basic C-terminal peptide (magenta). (B) Sequence alignment between endonuclease I of phages T7 and T3. The essential acidic amino acids are highlighted in red, while the inessential ones are indicated green. The basic C-terminal peptide is shown magenta. (C) The region of sequence similarity between T7 endonuclease I and T4 endonuclease VII. The identities are marked by the I symbol. This region contains four of the five essential acidic amino acids of T7 endonuclease I and essential and inessential acidic residues of endonuclease I are highlighted in red and green, respectively. Note that none aligns with an identical residue in endonuclease VII. The essential Glu86 residue of T4 endonuclease VII is highlighted in magenta; this does not align with an essential acidic residue in T7 endonuclease I.

proteins, while Glu35 of the phage T7 enzyme is conservatively replaced by aspartate in the phage T3 protein (Fig. 5B). While this is clearly expected if these residues play an important role in catalysis, it must be noted that the overall sequence identity is so high that most of the non-essential acidic residues are also conserved. T7 endonuclease I shares some sequence similarity with T4 endonuclease VII, although the degree of similarity is very much lower in this case. The region of endonuclease I between amino acids 30 and 76 is 32% identical to the central region of endonuclease VII between amino acids 78 and 121. Four of the five essential acidic residues of endonuclease I are found in the 30–76 region. However, none of these aligns to identical residues in endonuclease VII (Fig. 5C) and in only one case is there a conservative alignment (Glu35 in endonuclease I aligns with Asp83 in endonuclease VII). The previously identified essential Glu86 of endonuclease VII was originally investigated because of its alignment with Glu38 in endonuclease I (35), however, this residue is non-essential in the T7 enzyme and is replaced non-conservatively by leucine in T3 endonuclease I. Overall, while both regions contain acidic amino acids required for catalysis, mutation analysis does not lend support to a common structure.

While removal of individual carboxylate functions from the five essential acidic residues leads to a catalytically incompetent protein that retains the ability to bind DNA junctions, removal of the basic C-terminus of endonuclease I results in a protein that can neither bind nor cleave junctions. This C-terminal peptide contains six of nine basic amino acids (five of nine in T3 endonuclease I) and is reminiscent of histone sequences. The truncated protein is expressed to a high level in *E.coli* and is therefore unlikely to be grossly unfolded. It therefore seems likely that the basic peptide might be directly involved in binding DNA. In general the junction-resolving enzymes are characterised by high pI values (11) and binding is likely to be mediated by basic faces. Alternatively, it might be required for protein-protein interaction; endonuclease I binds to DNA junctions as a strongly associated dimer (37) and failure to undergo dimerisation would probably compromise binding.

In the course of these studies we have isolated two quite different kinds of mutated endonuclease I. We have identified five critical acidic amino acids, essential for catalytic activity of the enzyme. Mutation resulting in the removal of the carboxylate function leads to an inactive protein that can nevertheless bind to DNA junctions with structural selectivity and with unchanged manipulation of the global structure. In contrast, removal of the C-terminal basic peptide leads to a mutant of a different kind. This protein has lost its ability to bind to DNA junctions and the absence of nucleolytic activity on junctions is likely to be the result of the failure to bind its substrate. This approach provides a genetic analysis of the function of the junction-resolving enzyme and will be important when a structure of the enzyme becomes available.

## ACKNOWLEDGEMENTS

We thank our colleagues Drs Marie-Jo Giraud-Panis and Malcolm White for discussion and the Cancer Research Campaign for financial support.

## REFERENCES

- Holliday,R. (1964) *Genet. Res.*, **5**, 282–304.
- Broker,T.R. and Lehman,I.R. (1971) *J. Mol. Biol.*, **60**, 131–149.
- Orr-Weaver,T.L., Szostak,J.W. and Rothstein,R.J. (1981) *Proc. Natl Acad. Sci. USA*, **78**, 6354–6358.
- Potter,H. and Dressler,D. (1976) *Proc. Natl Acad. Sci. USA*, **73**, 3000–3004.
- Kitts,P.A. and Nash,H.A. (1987) *Nature*, **329**, 346–348.
- Nunes-Düby,S.E., Matsomoto,L. and Landy,A. (1987) *Cell*, **50**, 779–788.
- Hoess,R., Wierzbicki,A. and Abremski,K. (1987) *Proc. Natl Acad. Sci. USA*, **84**, 6840–6844.
- Jayaram,M., Crain,K.L., Parsons,R.L. and Harshey,R.M. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 7902–7906.
- McCulloch,R., Coggins,L.W., Colloms,S.D. and Sherratt,D.J. (1994) *EMBO J.*, **13**, 1844–1855.
- Schwacha,A. and Kleckner,N. (1995) *Cell*, **83**, 783–791.
- White,M.F., Giraud-Panis,M.-J.E., Pöhler,J.R.G. and Lilley,D.M.J. (1997) *J. Mol. Biol.*, **269**, 647–664.
- Kemper,B. and Garabett,M. (1981) *Eur. J. Biochem.*, **115**, 123–131.
- de Massey,B., Studier,F.W., Dorgai,L., Appelbaum,F. and Weisberg,R.A. (1984) *Cold Spring Harbor Symp. Quant. Biol.*, **49**, 715–726.
- Iwasaki,H., Takahagi,M., Shiba,T., Nakata,A. and Shinagawa,H. (1991) *EMBO J.*, **10**, 4381–4389.
- Connolly,B., Parsons,C.A., Benson,F.E., Dunderdale,H.J., Sharples,G.J., Lloyd,R.G. and West,S.C. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 6063–6067.
- Sharples,G.J. and Lloyd,R.G. (1991) *J. Bacteriol.*, **173**, 7711–7715.
- Sharples,G.J., Chan,S.N., Mahdi,A.A., Whitby,M.C. and Lloyd,R.G. (1994) *EMBO J.*, **13**, 6133–6142.
- West,S.C. and Komer,A. (1985) *Proc. Natl Acad. Sci. USA*, **82**, 6445–6449.
- Symington,L. and Kolodner,R. (1985) *Proc. Natl Acad. Sci. USA*, **82**, 7247–7251.
- White,M.F. and Lilley,D.M.J. (1997) *Mol. Cell. Biol.*, **17**, 6465–6471.
- Whitby,M.C. and Dixon,J. (1997) *J. Mol. Biol.*, **272**, 509–522.
- Elborough,K.M. and West,S.C. (1990) *EMBO J.*, **9**, 2931–2936.
- Hyde,H., Davies,A.A., Benson,F.E. and West,S.C. (1994) *J. Biol. Chem.*, **269**, 5202–5209.
- Stuart,D., Ellison,K., Graham,K. and McFadden,G. (1992) *J. Virol.*, **66**, 1551–1563.
- Powling,A. and Knippers,R. (1976) *Mol. Gen. Genet.*, **149**, 63–71.
- Kerr,C. and Sadowski,P.D. (1975) *Virology*, **65**, 281–285.
- Tsujimoto,Y. and Ogawa,H. (1978) *J. Mol. Biol.*, **125**, 255–273.
- Center,M.S. and Richardson,C.C. (1970) *J. Biol. Chem.*, **245**, 6285–6291.
- Center,M.S. and Richardson,C.C. (1970) *J. Biol. Chem.*, **245**, 6292–6299.
- Sadowski,P.D. (1971) *J. Biol. Chem.*, **246**, 209–216.
- de Massey,B., Weisberg,R.A. and Studier,F.W. (1987) *J. Mol. Biol.*, **193**, 359–376.
- Dickie,P., McFadden,G. and Morgan,A.R. (1987) *J. Biol. Chem.*, **262**, 14826–14836.
- Kemper,B. and Janz,E. (1976) *J. Virol.*, **18**, 992–999.
- Duckett,D.R., Giraud-Panis,M.-E. and Lilley,D.M.J. (1995) *J. Mol. Biol.*, **246**, 95–107.
- Pöhler,J.R.G., Giraud-Panis,M.-J.E. and Lilley,D.M.J. (1996) *J. Mol. Biol.*, **260**, 678–696.
- White,M.F. and Lilley,D.M.J. (1996) *J. Mol. Biol.*, **257**, 330–341.
- Parkinson,M.J. and Lilley,D.M.J. (1997) *J. Mol. Biol.*, **270**, 169–178.
- Shah,R., Cosstick,R. and West,S.C. (1997) *EMBO J.*, **16**, 1464–1472.
- Giraud-Panis,M.-J.E. and Lilley,D.M.J. (1998) *J. Mol. Biol.*, **278**, 117–133.
- Giraud-Panis,M.-J.E. and Lilley,D.M.J. (1997) *EMBO J.*, **16**, 2528–2534.
- Bennett,R.J. and West,S.C. (1995) *J. Mol. Biol.*, **252**, 213–226.
- White,M.F. and Lilley,D.M.J. (1997) *J. Mol. Biol.*, **266**, 122–134.
- Chan,S.N., Vincent,S.D. and Lloyd,R.G. (1998) *Nucleic Acids Res.*, **26**, 1560–1566.
- Giraud-Panis,M.-J.E. and Lilley,D.M.J. (1996) *J. Biol. Chem.*, **271**, 33148–33155.
- Panayotatos,N. and Backman,S. (1989) *J. Biol. Chem.*, **264**, 15070–15073.
- Shah,R., Bennett,R.J. and West,S.C. (1994) *Cell*, **79**, 853–864.
- Schofield,M.J., Lilley,D.M.J. and White,M.F. (1998) *Biochemistry*, **37**, 7733–7740.
- Chan,S.N., Harris,L., Bolt,E.L., Whitby,M.C. and Lloyd,R.G. (1997) *J. Biol. Chem.*, **272**, 14873–14882.
- Fothergill,M., Goodman,M.F., Petruska,J. and Warshel,A. (1995) *J. Am. Chem. Soc.*, **117**, 11619–11627.
- Beese,L.S. and Steitz,T.A. (1991) *EMBO J.*, **10**, 25–33.
- Katayanagi,K., Miyagawa,M., Matsuchima,M., Ishikawa,M., Kanaya,S., Nakamura,H., Ikehara,M., Matsuzaki,T. and Morikawa,K. (1992) *J. Mol. Biol.*, **223**, 1029–1052.
- Cheng,X.D., Balendiran,K., Schildkraut,I. and Anderson,J.E. (1994) *EMBO J.*, **13**, 3927–3935.
- Kostrewa,D. and Winkler,F.K. (1995) *Biochemistry*, **34**, 683–696.
- Kohlstaedt,L.A., Wang,J., Friedman,J.M., Rice,P.A. and Steitz,T.A. (1992) *Science*, **256**, 1783–1790.
- Beese,L.S., Derbyshire,V. and Steitz,T.A. (1993) *Science*, **260**, 352–355.
- Saito,A., Iwasaki,H., Ariyoshi,M., Morikawa,K. and Shinagawa,H. (1995) *Proc. Natl Acad. Sci. USA*, **92**, 7470–7474.
- Cohen,S., Chang,A. and Hsu,L. (1972) *Proc. Natl Acad. Sci. USA*, **69**, 2110–2114.
- Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl Acad. Sci. USA*, **74**, 5463–5467.
- Duckett,D.R., Murchie,A.I.H., Diekmann,S., von Kitzing,E., Kemper,B. and Lilley,D.M.J. (1988) *Cell*, **55**, 79–89.
- Kunkel,T.A. (1985) *Proc. Natl Acad. Sci. USA*, **82**, 488–492.
- Vieira,J. and Messing,J. (1987) *Methods Enzymol.*, **153**, 3–11.
- Zuoco,J. and Boyd,A. (1992) *Anal. Biochem.*, **207**, 348–349.
- Murchie,A.I.H., Clegg,R.M., von Kitzing,E., Duckett,D.R., Diekmann,S. and Lilley,D.M.J. (1989) *Nature*, **341**, 763–766.
- White,M.F. and Lilley,D.M.J. (1998) *Nucleic Acids Res.*, **26**, 5609–5616.
- Parsons,C.A., Stasiak,A., Bennett,R.J. and West,S.C. (1995) *Nature*, **374**, 375–378.
- Guo,F., Gopaul,D.N. and Van Duyne,G.D. (1997) *Nature*, **389**, 40–46.
- Pöhler,J.R.G., Norman,D.G., Bramham,J., Bianchi,M.E. and Lilley,D.M.J. (1998) *EMBO J.*, **17**, 817–826.
- Bennett,R.J., Dunderdale,H.J. and West,S.C. (1993) *Cell*, **74**, 1021–1031.
- Beck,P.J., Gonzalez,S., Ward,C.L. and Molineux,I.J. (1989) *J. Mol. Biol.*, **210**, 687–701.