

Conversion of bovine pancreatic DNase I to a repair endonuclease with a high selectivity for abasic sites

Santiago Cal¹, Kok Leong Tan,
Alistair McGregor and Bernard A.Connolly²

Department of Biochemistry and Genetics, The University of Newcastle, Newcastle upon Tyne NE2 4HH, UK

¹Present address: Departamento de Bioquímica y Biología Molecular, Facultad de Medicina, Universidad de Oviedo, Oviedo 33006, Spain

²Corresponding author
e-mail: B.A.Connolly@ncl.ac.uk

Bovine pancreatic deoxyribonuclease I (DNase I) is a nuclease of relatively low specificity which interacts with DNA in the minor groove. No contacts are made between the protein and the major groove of the nucleic acid. DNase I is structurally homologous to exonuclease III, a DNA-repair enzyme with multiple activities. One of the main differences between the two enzymes is the presence of an additional α -helix in exonuclease III, in a position suggestive of interaction with the major groove of DNA. Recombinant DNA techniques have been used to add 14 amino acids, comprising the 10 amino acids of the exonuclease III α -helix flanked by a glycine rich region, to DNase I. The polypeptide has been inserted after serine 174, an amino acid on the surface of DNase I corresponding to the location of the extra α -helix in exonuclease III. The recombinant protein, DNase-exohelix, has been purified and its catalytic activities towards DNA investigated. The recombinant protein demonstrated a high selectivity for endonucleolytic cleavage at abasic sites in DNA, a property of exonuclease III but not native DNase I. Thus the insertion of 14 amino acids at Ser174, converts DNase I to an exonuclease III-like enzyme with DNA-repair properties.

Keywords: DNA abasic sites/DNA repair/DNase I/exonuclease III/protein–DNA interactions

Introduction

Several enzymes which catalyze phosphodiester bond hydrolysis in nucleic acids have considerable structural homology. Particularly striking is the similarity between bovine pancreatic deoxyribonuclease I (DNase I) (Suck *et al.*, 1984, 1988; Oefner and Suck, 1986; Suck and Oefner, 1986; Lahm and Suck, 1991; Weston *et al.*, 1992), a low specificity endonuclease, and the two homologous DNA-repair enzymes *Escherichia coli* exonuclease III (Mol *et al.*, 1995) and human apurinic/apyrimidinic endonuclease (HAP 1) (Gorman *et al.*, 1997). DNase I consists of two central β -sheets, each comprising six β -strands, flanked with extensive α -helical and loop regions. The topology of the two β -sheets, which are packed closely together, are remarkably similar and related

by a quasi 2-fold axis. Exonuclease III and HAP 1 contain a core β -sheet structure which is almost identical to that found in DNase I and several of the surrounding α -helices are also conserved between the two classes of enzyme. Many individual amino acids, proposed to have a critical role in either catalysis or DNA binding, are retained by the three proteins. This includes an Asp–His pair (D212–H252 for DNase I; D229–H259 for exonuclease III; D283–H309 for HAP 1) suggested to abstract a proton from the attacking water molecule and a glutamate (E39 for DNase I; E34 for exonuclease III; E96 for HAP 1) that serves to co-ordinate the essential metal ion. Site-directed mutagenesis of these amino acids in DNase I (Jones *et al.*, 1996) and HAP 1 (Barzilay *et al.*, 1995) has demonstrated their importance in activity and a similar mechanism has been proposed for the three enzymes (Weston *et al.*, 1992; Mol *et al.*, 1995; Jones *et al.*, 1996; Gorman *et al.*, 1997). However the mechanisms cannot be completely identical as an additional Glu78–His134 pair found in DNase I and suggested to be involved in leaving-group protonation (Weston *et al.*, 1992; Jones *et al.*, 1996) has no counterpart in the repair enzymes. This family of structurally related proteins can be extended as *E.coli* RNase H (Katayanagi *et al.*, 1990) shows similarities with DNase I. RNase H contains five β -strands and is smaller (155 amino acids) than DNase I (260 amino acids). However, four of the five β -strands in RNase H have almost identical directions and twists to four of the six β -strands in the two halves of the DNase I molecule. Furthermore the metal-binding sites of both nucleases, which form the catalytic centre, lie in similar regions of the conserved β -strands. The DNase I/RNase H fold, particularly the location and orientation of the central β -strands has also been observed in HIV and *E.coli* integrase and *E.coli* RuvC (Yang and Steitz, 1995). RNase H, integrase and RuvC all contain a pair of conserved carboxylates, involved in binding of the catalytic metal ion, which are located in structurally equivalent positions on two of the β -strands. As all the enzymes discussed bind nucleic acids, and subsequently hydrolyze phosphodiester bonds, it is unsurprising that they have features in common. However, the enzymes differ in the particular nucleic acids that they recognize and also the context within which they are able to cut phosphodiester bonds. These subtleties in selectivity must be reflected in structural differences between the nucleases which have been little explored, for example by attempting to interconvert the specificities of the proteins. Experiments are most promising with DNase I and exonuclease III/HAP 1 due to their very high structural similarities.

DNase I catalyzes the endonucleolytic hydrolysis of double-stranded DNA with a low, but not completely random, selectivity (Lomonosoff *et al.*, 1981; Drew and Travers, 1984, 1985). The hydrolysis patterns which result

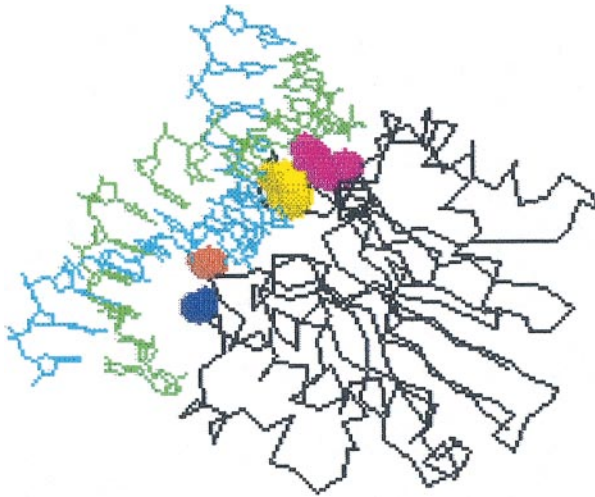


Fig. 1. DNase I–DNA structure. This structure contains the protein bound to a nicked quasi-continuous double helix of 14 base pairs which shows the locations of the major and the minor grooves of the DNA relative to the protein (Lahm and Suck, 1991; Suck *et al.*, 1998). The protein backbone is shown in black and the two DNA strands in green and cyan. The sidechain atoms of Arg41 (magenta) and Tyr76 (yellow), which bind to DNA in the minor groove and completely fill this groove, have been space-filled and highlighted. The positions of Ser138 (red) and Ser174 (blue) are indicated by representing their sidechain atoms in the space-filling mode. These two amino acids, which do not interact with the DNA, are located on the surface of DNase I in proximity to the DNA major groove.

from the digestion of a number of DNA duplexes can be reasonably well explained, particularly by reference to X-ray crystallographic data. Structures of DNase I, with bound oligonucleotides, show that the protein contacts the DNA via both the sugar-phosphate backbone and the minor groove (Suck *et al.*, 1988; Lahm and Suck, 1991; Weston *et al.*, 1992). Two amino acids, Arg41 and Tyr76, are inserted into the minor groove and this results in the widening of this groove by ~ 3 Å and its bending in the direction of the major groove. DNase I exhibits a preference for cutting DNA at flexible regions that can be easily bent towards the major groove. Rigid sequences, particularly long runs of A/T tracts, or those that bend away from the major groove, are poorly cut (Hogan *et al.*, 1989; Bruckner *et al.*, 1990; Mendoza *et al.*, 1990; Travers, 1993). DNase I does not make any interactions with the major groove of DNA. Most DNA-binding proteins, particularly those that show a high degree of specificity, make extensive contacts to the bases in the major groove, as this is where DNA sequences are most easily distinguished (Seeman *et al.*, 1976; Steitz, 1990; Harrison, 1991; Pabo and Sauer, 1992; Travers, 1993). The lack of direct contacts between DNase I and the major groove is one of the reasons for its low DNA hydrolysis selectivity. However, two regions of DNase I, centred on the surface amino acids Ser138 and Ser174 (Figure 1), approach the DNA major groove. Based on structural information and modelling it has been suggested that the insertion of additional amino acids, next to Ser138 or 174, has the potential to generate new contacts in the major groove (Wolf *et al.*, 1995). An Arg–Glu dipeptide sequence has previously been inserted into DNase I at Cys173 (converting the wild-type sequence C(173)SYV to

C(173)REGTV and the resulting protein showed interesting alterations to cleavage selectivity (Wolf *et al.*, 1995).

Exonuclease III and HAP 1 are multifunctional DNA-repair enzymes with the following activities: (i) AP endonuclease, which cleaves phosphodiester bonds 5' to abasic sites in DNA. (ii) Exonuclease (3'→5'), releasing 5'-mononucleotides from the 3'-ends of DNA. (iii) Phosphomonoesterase and repair-phosphodiesterase which removes phosphates and phosphoglycolates from the 3'-ends of DNA. (iv) Ribonuclease H, which degrades the RNA strand in an RNA–DNA duplex (Rogers and Weiss, 1980; Lloyd and Linn, 1993; Demple and Harrison, 1994; Barzilay and Hickson, 1995). Although exonuclease III and HAP 1 have limited amino acid sequence similarity with DNase I, X-ray crystallography (Mol *et al.*, 1995; Gorman *et al.*, 1997) has revealed considerable structural homology as detailed above. Unfortunately, structures of the two repair enzymes with DNA are not available, but modelling indicated that recognition of nucleic acids shows some similarity with DNase I. Thus equivalents of Arg41 and Tyr76, which DNase I uses to interact with the DNA minor groove, are present in the repair enzymes as are homologues of several of the amino acids that interact with phosphate groups. However, exonuclease III contains 19 extra amino acids, comprising an α -helix flanked by flexible loop regions, that have no equivalent in DNase I. This new element is inserted at a position equivalent to one between Ser174 and Tyr175 in DNase I. Comparison of the exonuclease III apo-enzyme with DNase I/DNA structures suggests that the additional α -helix is in the correct position to make contacts to the major groove of DNA. HAP1, a human homologue of exonuclease III, also contains this supplementary insert but the α -helix is shortened by two turns as compared with exonuclease III. The publication detailing the HAP 1 structure (Gorman *et al.*, 1997) also points out that both repair proteins contain two small surface loops, additional to the α -helix containing element discussed above, which are missing in DNase I. Modelling showed that one of these loops is probably positioned near the major and the other near the minor groove of DNA.

In summary, both repair proteins show many DNA-binding features that are common to DNase I but have unique structural elements that are postulated to produce additional contacts with the nucleic acid, particularly in the major groove. It is likely that these new protein–DNA contacts account for the different activities of exonuclease III/HAP 1 and DNase I, especially the recognition of abasic sites; a key activity of the repair enzymes absent in DNase I. The presence of extra amino acids following Ser174, found with the repair enzymes, has conceptual similarity with the Arg–Glu dipeptide previously incorporated into DNase I at Cys173. Therefore, the proposal (Wolf *et al.*, 1995) that insertions at this position of DNase I may lead to new interactions with DNA and alterations in activity seems to hold not only for engineered variants of DNase I but may have natural counterparts in exonuclease III and HAP 1. In this publication the role of the extra α -helix in the repair enzyme and the relationship between exonuclease III/HAP 1 and DNase I is explored by the preparation and investigation of a DNase I

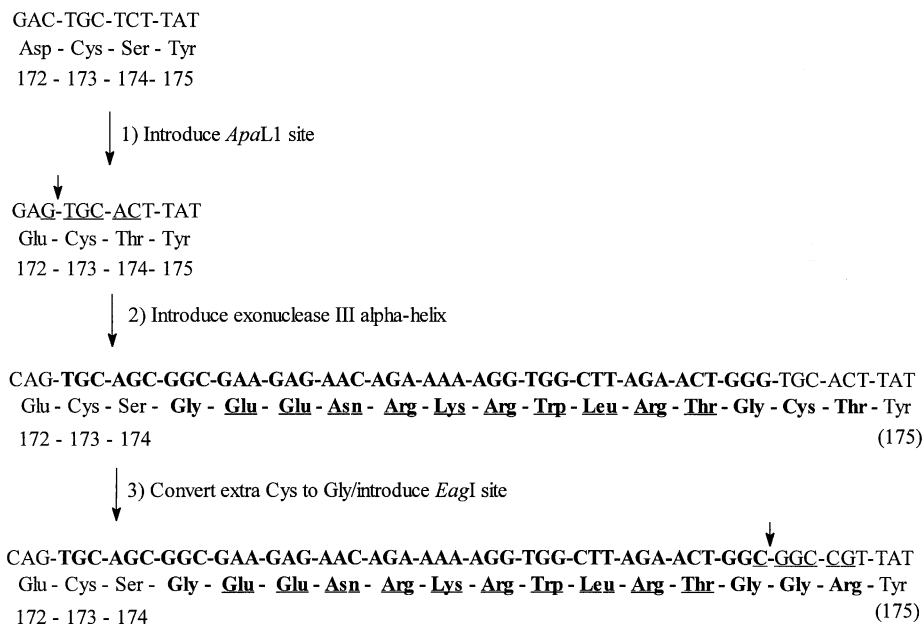


Fig. 2. Construction of a gene coding for DNase–exohelix. Step 1 consists of engineering a unique *Apa*LI restriction site (underlined, cleavage site shown with an arrow) into the DNase I gene. This also changes Asp172 to glutamic acid and Ser174 to threonine. In step 2 an oligonucleotide (bold) is inserted into the *Apa*LI site. This restores amino acid 174 to a serine and leads to the insertion of 14 amino acids (bold) between Ser174 and what was previously tyrosine 175, now shown as Tyr(175). The underlined amino acids comprise the exonuclease III α -helix. In step 3, the additional cysteine is removed by conversion to glycine and a unique *Eag*I restriction site (underlined, cleavage site shown with an arrow) introduced.

variant containing the exonuclease III α -helix inserted at Ser174.

Results

Construction of a gene for DNase I containing the exonuclease III α -helix inserted between Ser174 and Tyr175

The construction of the required DNase I hybrid (DNase–exohelix) originates from a gene for this enzyme previously prepared in our laboratory by chemical synthesis (Worrall and Connolly, 1990). Although a large number of restriction sites had been incorporated into the synthetic gene, there are none in the immediate region coding for Ser174. The DNase gene was, therefore, further engineered, using a PCR technique (Ho *et al.*, 1989), to incorporate a unique *Apa*LI site (GTGCAC) at Ser174 (Figure 2). This manipulation also converted Asp172 and Ser174 to Glu and Thr, respectively. An oligonucleotide coding for the 10 amino acids in the exonuclease III α -helix (EENRKRWLRT), flanked on either side by a single glycine residue was then introduced at the *Apa*LI site (Figure 2). Insertion of this oligonucleotide additionally destroyed the unique *Apa*LI site, converted Thr174 back to the original serine and placed two additional amino acids, Cys and Thr, in front of the original Y175 of DNase I. Wild-type DNase I contains four Cys residues, present as two disulfide bridges (Price *et al.*, 1969a; Moore, 1981; Suck *et al.*, 1984). We were concerned that the presence of a fifth Cys would interfere with disulfide-bond formation and hence protein folding. Therefore a final step was included to convert this Cys–Thr to Gly–Arg and at the same time introduce a new, unique, *Eag*I site, useful for confirmation of the presence of the insert. Figure 2 shows that the end result is the insertion of 14

amino acids between Ser174 and Tyr175, consisting of the ten amino acids comprising the exonuclease III α -helix surrounded by glycine rich regions. It was hoped that the glycine residues would confer flexibility to the DNase I-insert junction and so aid in the successful folding of both the DNase I to its tertiary structure and the inserted sequence to an α -helix. The actual, 19 amino acid, insert used in exonuclease III also has glycine residues flanking the α -helix, which are used to form a flexible linker that joins the helix to other secondary structural elements. The conservative change D172E, introduced during the DNase I gene manipulation would be expected to be inconsequential as this amino acid is on the surface of the protein and does not have a critical function.

Overexpression and purification of DNase–exohelix

The gene coding for DNase–exohelix was subcloned into the expression vector pKK223, which has a *tac*-promoter allowing controlled protein expression, by isopropyl β -D-thiogalactopyranoside (IPTG), in *lac* I^q hosts (Brosius and Holy, 1984). The *Eco*RI–*Hind*III restriction endonuclease sites, present in the polylinker of pKK223 and at the ends of the DNase I gene, were used for this purpose. This system has previously been used by us for the successful overexpression of inactive or very weakly active mutants of DNase I (Doherty *et al.*, 1995; Jones *et al.*, 1996) but cannot be used for wild-type DNase I (or active variants) because of toxicity problems (Worrall and Connolly, 1990; Doherty *et al.*, 1991, 1993). However, as detailed below, DNase–exohelix has relatively low activity and so is compatible with pKK223. Nevertheless, we observed that *E. coli* JM105 transformed with plasmids containing the gene for DNase–exohelix were rather unstable to storage and lost the ability to express these proteins within a

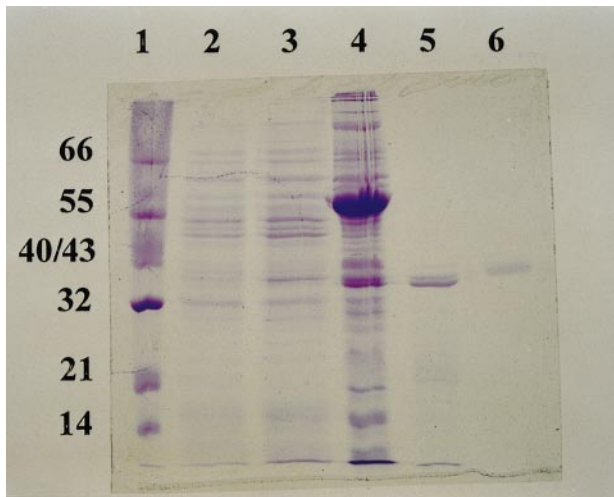


Fig. 3. Purification of DNase-exohelix as assessed by SDS-PAGE. Lane 1, mol. wt markers with masses (kDa) indicated; lane 2, crude soluble *E. coli* homogenate prior to DNase-exohelix induction; lane 3, crude soluble *E. coli* homogenate post-DNase-exohelix induction; lane 4, following ion-exchange chromatography on Q-Sepharose; lane 5, following gel-filtration chromatography on Superdex-G75; lane 6, following ion-exchange FPLC on DEAE-5PW.

week. Therefore, we always used freshly transformed JM105 cells to direct protein expression.

Previously, we have not experienced major problems with the purification of reasonable amounts of recombinant inactive DNase I variants from *E. coli* (Worrall and Connolly, 1990; Doherty *et al.*, 1993; Jones *et al.*, 1996). This was not the case with DNase-exohelix, which suffered from two problems: insolubility and susceptibility to proteolysis. This is perhaps not surprising as the presence of the extra polypeptide sequence may promote mis-folding and so lead to insolubility. Furthermore, the presence of flexible, glycine rich, regions surrounding the insert may confer protease sensitivity. The amount of soluble protein produced was maximized by: (i) inducing protein expression with IPTG for 3–4 h, rather than overnight; (ii) growing the *E. coli* at 30°C rather than 37°C; and (iii) including 0.5 M NaCl in the buffer used to lyse the *E. coli* cells by sonication. Using these precautions the majority of DNase-exohelix was soluble and appeared in the supernatant fraction following sonication. Proteolysis was minimized by: (i) including the protease inhibitors benzamidine and phenylmethylsulphonyl fluoride in the sonication buffer and the buffers used for the first purification column; and (ii) ensuring that the cell harvesting, sonication and first purification column took place in as short a time as possible.

A three-column purification procedure based on ion-exchange chromatography on Q-Sepharose, followed by FPLC gel filtration and a final FPLC ion-exchange chromatography using a Protein-Pak DEAE-5PW column was sufficient to purify the protein. Although the initial column did not lead to a great deal of protein purification it removed much of the nucleic acid present. Following each chromatographic step, fractions were tested for the presence of DNase-exohelix by SDS-PAGE, Western blotting and the ability to hydrolyse λ phage DNA (Sadowski and Hurwitz, 1969). In order to obtain pure protein, relatively shallow salt gradients were needed for

both ion-exchange columns and it was important to pool only the purest fractions. The purified protein appeared >95% pure by SDS-PAGE, stained with Coomassie Blue (Figure 3). Sequencing of the N-terminus gave MLKIAAFNIR, the expected sequence for recombinant DNase I (Moore, 1981; Worrall and Connolly, 1990). In addition, the purified protein showed reactivity with a polyclonal antibody prepared against commercially available bovine pancreatic DNase I (Jones, 1994) and was able to hydrolyse λ phage DNA (Sadowski and Hurwitz, 1969) (not shown). Typically 1–2 mg of purified protein was obtained from 8 g of *E. coli* cell paste. This yield is rather lower than previously observed with single amino acid mutations to DNase I (Doherty *et al.*, 1993, 1995; Jones *et al.*, 1996) and presumably arises from the insolubility and proteolytic sensitivity of the hybrid.

Hydrolysis of phosphodiester bonds by DNase-exohelix

When the small non-DNA chromophoric substrate thymidine-3',5'-di-(*p*-nitrophenyl)phosphate [T(npp)₂] (Liao, 1975) was hydrolyzed by DNase-exohelix, rates of between 10–15% of those seen with wild-type DNase I were obtained (Table I). As previously observed (Liao, 1975; Doherty *et al.*, 1995), the hydrolysis of this substrate did not follow Michaelis-Menten kinetics and this precluded detailed kinetic analysis. The observed reduction in rate to 10–15% of the wild-type value was obtained by averaging results obtained at 5, 10 and 20 mM levels of the chromophoric substrate. The ratios of activities seen with wild-type DNase I and DNase-exohelix were the same using either Mg²⁺ or Mn²⁺ as activating metal ion. Although T(npp)₂ has little structural similarity with DNA and DNase I cuts different bonds in these two substrates (Liao, 1975), the enzyme uses the same active-site amino acids and catalytic mechanism for the hydrolysis of both (Liao, 1975; Jones *et al.*, 1996). Therefore, the hydrolysis of this chromophore, by DNase I mutants, can be used to assess the integrity of the active site and to observe any indirect effects that result from mutations remote from the catalytic centre caused by, for example, conformational changes (Doherty *et al.*, 1995; Warren *et al.*, 1997). The insertion of the polypeptide after Ser174, which is remote from the active site, clearly causes some reduction in the hydrolysis of T(npp)₂. This probably results from slight disruptions to the overall conformation of the hybrid DNase I with a consequent sub-optimal alignment of key catalytic residues. Nevertheless, the rates obtained are still reasonable and indicate that the active site is functional, albeit at a reduced level.

The hydrolysis of DNA by DNase-exohelix was measured using the Kunitz assay (Kunitz, 1950; Doherty *et al.*, 1995). This assay involves measuring the hydrolysis of double-stranded DNA by observing the increase in absorbance at 260 nm. Previously, we obtained K_m and V_{max} values for wild-type DNase I and several mutants, using calf thymus DNA concentrations of between 0.001 and 0.1 mg/ml (Doherty *et al.*, 1995; Jones *et al.*, 1996; Warren *et al.*, 1997). However, DNase-exohelix was found to have a very low activity in the Kunitz assay and this made accurate determination of kinetic parameters impossible. The measurements at several concentrations of DNA between 0.001 and 0.1 mg/ml consistently gave

Table I. A summary of the activities of DNase I, DNase–exohelix and exonuclease III

Activity	DNase I	DNase–exohelix	Exonuclease III
Thymidine-3',5'-di-(<i>p</i> -nitrophenyl) phosphate hydrolysis	Yes	Yes, at 10–15% of DNase I rate	Not determined
DNA hydrolysis by endonucleolytic action	Yes	Yes, at 0.07% of DNase I rate	No
3'-exonuclease activity	No	Yes	Yes
3'-DNA phosphatase/ phosphodiesterase activity	No	Not determined	Yes
DNA hydrolysis selectivity	Low, but not completely random. Endonucleolysis only	Lower than DNase I due to combination of endo- and exonucleolysis	Low, but not completely random. Exonucleolysis only
RNase H activity	No	No	Yes
Hydrolysis at abasic sites in DNA	No	Yes	Yes

The properties of DNase I and exonuclease III are based on previous publications (Rogers and Weiss, 1980; Drew and Travers, 1984; Liao, 1975; Demple and Harrison, 1994; Doherty *et al.*, 1995). The properties of DNase–exohelix have been evaluated in this publication.

a rate for DNase–exohelix of ~0.07% of the wild-type value (Table I). This is a much greater reduction than observed with T(npp)₂ and needs to be accounted for by factors additional to the slight disruption to the active site measured with the chromophoric substrate. It is likely that the presence of the polypeptide insert either directly prevents or hinders DNA binding or, more probably, DNA binding occurs but in a non-productive manner. In the latter case DNA would bind to DNase–exohelix but the presence of the polypeptide insert would prevent the subsequent distortion of the DNA that is essential for the correct positioning of the scissile phosphate at the active site and efficient hydrolysis.

3'-exonuclease activity of DNase–exohelix

Oligonucleotides containing a ³²P-label at the phosphodiester group nearest to the 3'-terminal (prepared by a single addition of an [α-³²P]ddNTP with terminal deoxynucleotidyl transferase) were used to assess any 3'-exonuclease activity of DNase–exohelix, wild-type DNase I and, as a control, exonuclease III. The oligonucleotides were designed to form stem-loop structures with the 3'-end either recessed or protruding by two bases. As shown in Figure 4 and summarized in Table I, exonuclease III has a very strong 3'-exonuclease activity and released a 5'-phosphate mononucleotide from the 3'-end of the oligonucleotides. This activity was observed whether the 3'-terminal was recessed or protruding [although exonuclease III is usually considered to act on 3'-termini that are either blunt or recessed, it will also degrade 3'-overhangs providing they do not exceed four bases (Promega technical bulletin, 517)]. We also observed some endonucleolytic cleavage with exonuclease III (Figure 4), which may arise from impurities in the commercial preparation. Wild-type DNase I showed no measurable 3'-exonuclease activity (Figure 4). However, it is clear that DNase–exohelix does have a 3'-exonuclease activity, as shown by the presence of a band on the gel at the position of a mononucleotide (Figure 4). This activity was measured in the presence of 0, 0.5 or 2 mM Ca²⁺ as DNase I contains two structural Ca²⁺ ions that are involved in folding the protein and stabilizing its structure (Price *et al.*, 1969a,b; Price, 1972, 1975; Moore, 1981) but no such metal ions

are found with exonuclease III. Clearly, 3'-exonuclease activity is seen under all three conditions. The DNase–exohelix (and also DNase I) also shows several bands that arise from endonucleolytic cleavage at internal positions. As expected wild-type DNase I did not cut the oligonucleotides in the single-stranded loop region and hydrolysis was restricted to the double-stranded stem (Figure 4). A similar result was seen for DNase–exohelix which demonstrates that this variant is also unable to hydrolyze single-stranded DNA and that endonucleolytic cleavage is confined to double-stranded DNA. Finally, it is clear that the 3'-exonuclease activity of DNase–exohelix is relatively weak, when compared with the endonuclease activity. The sum of intensities of the bands arising from endonucleolysis being much greater than the single band due to 3'-exonucleolysis. The opposite pattern is observed with exonuclease III. None of the enzymes showed any 5'-exonuclease activity (not shown).

DNA hydrolysis selectivity of DNase–exohelix

A duplex formed by annealing AGATAGAGATATCTTAGGATCCGTTAAGG and AGACCTTAACGGATCCTAAGATATCTCTAT was used to investigate the specificity of DNA hydrolysis by DNase–exohelix and as controls, DNase I and exonuclease III. The hydrolysis patterns found are shown in Figure 5. DNase–exohelix gives a smoother pattern than both wild-type DNase I (which cuts in an endonucleolytic manner) and exonuclease III (which degrades DNA exonucleolytically from the 3'-end). Gaps are present in both the DNase I and exonuclease III digests and band intensity (which is proportional to cutting probability) varies greatly. In contrast, DNase–exohelix is able to cut at almost every phosphodiester, and the band intensities vary less than for the other two enzymes, even though not every bond is cut with the same frequency. This smooth hydrolysis pattern is most simply explained by a combination of endo- and exonuclease activities. Thus endonucleolytic cleavage by DNase–exohelix, the normal activity of DNase I, liberates 3'-OH termini which can then serve as a substrate for further degradation by the exonuclease activity present in the hybrid. Such dual activity would be expected to give rise to a more even hydrolysis pattern than seen with DNase I.

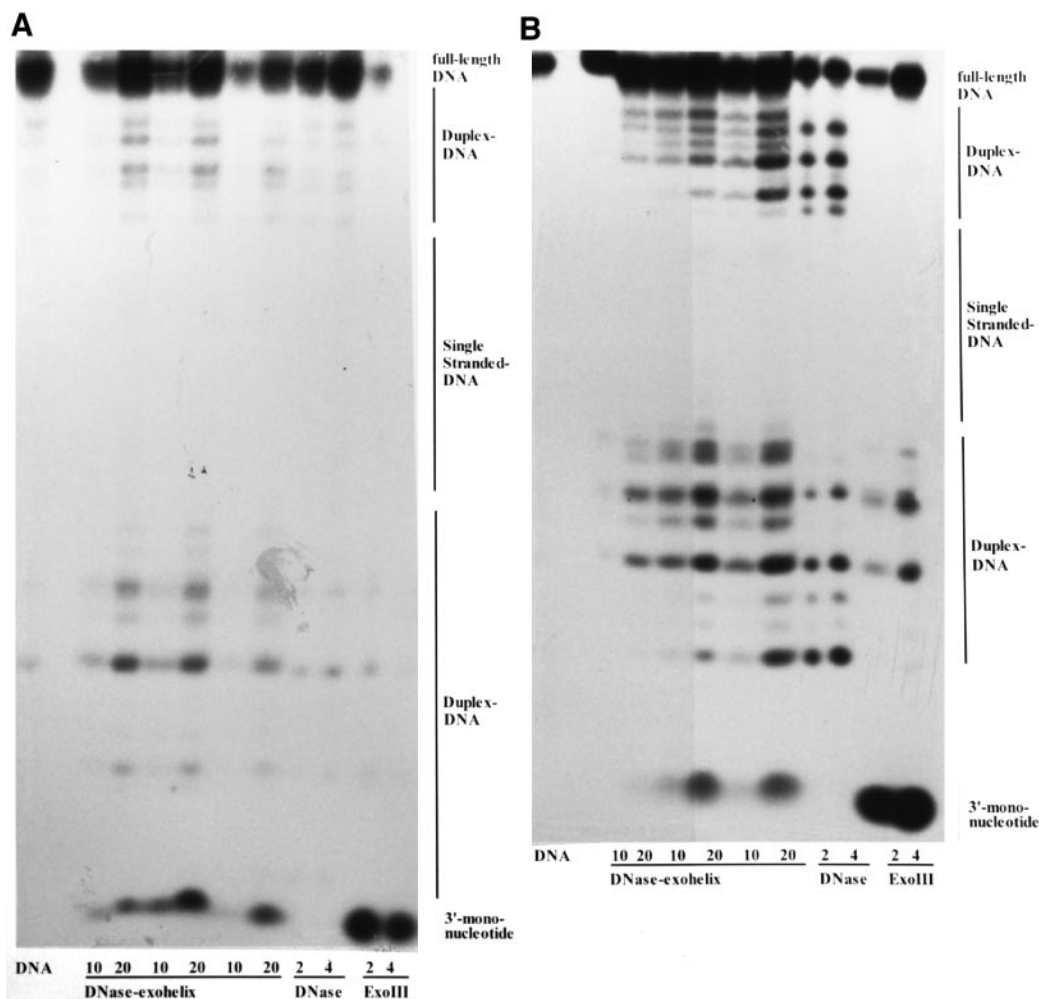


Fig. 4. Evaluation of 3'-exonuclease activity. The stem-loop structures (A) and (B):

(A): 5'-ATTGCATGATATCTAAGGGAAAAGATATCATGCA*-3'

(B): 5'-ATGATATCTTTCCCTTAGATATCATGCA*-3'

(*A = ddA and the phosphate between *A and the next base is ^{32}P -labelled) were incubated separately with DNase-exohelix (1 μg), DNase I (1.5 ng) or exonuclease III (120 units) for the times shown (in minutes) and the hydrolysis products analyzed by gel electrophoresis followed by autoradiography. The three sets of experiments (each with a 10 and 20 min incubation) with DNase-exohelix used 0 (left-hand lanes), 0.5 (middle lanes) and 2 (right-hand lanes) mM CaCl_2 . With DNase I only 2 mM CaCl_2 was used and with exonuclease III the manufacturer's buffer (see Materials and methods), which lacks Ca^{2+} . The lane marked 'DNA' contains the oligonucleotides prior to nuclease addition. The location of the single- and double-stranded regions are indicated on the figure as is the position of the mononucleotide released by any 3'-exonuclease activity. These were determined using the *EcoRV* restriction site present in both oligonucleotides and also a snake venom phosphodiesterase digest (not shown).

RNase H activity of DNase-exohelix

When a duplex consisting of an RNA strand hybridized to a DNA strand was treated with DNase-exohelix no cutting of either the RNA or the DNA was observed. An identical result was found with wild-type DNase I. Controls showed that exonuclease III was able to degrade the RNA strand of this duplex (data not shown; summarized in Table I).

AP endonuclease activity of DNase-exohelix

Two substrates were used to evaluate any AP-endonuclease activity of DNase-exohelix. The first contains a true abasic site, i.e., 2'-deoxyribose, and is generated from an oligonucleotide containing a single dU using the enzyme uracil-DNA glycosylase (Lindahl *et al.*, 1977; Gu *et al.*, 1994; Shida *et al.*, 1996). Alternatively an oligonucleotide containing a stable analogue of a true abasic site, 1',2'-dideoxyribose, was prepared by chemical synthesis. It is

known that oligonucleotides containing 1',2'-dideoxyribose are substrates for the AP-endonuclease activity of exonuclease III (Shida *et al.*, 1996). Double-stranded oligonucleotides containing either a single true abasic site or 1',2'-dideoxyribose in one of the strands were treated with DNase-exohelix, wild-type DNase I and exonuclease III and the products analyzed by gel electrophoresis as shown in Figure 6. For both oligonucleotides it is clear that DNase-exohelix shows an extremely high selectivity for cutting at the abasic site and the band on the gel representing cleavage at this position is by far the most prominent. The presence (0.5 or 2 mM) or absence of Ca^{2+} did not have any bearing on this activity. Examination of the gels shows that the prominent band occurs at the base immediately 5' to the abasic site (Figure 6). This indicates that cutting takes place at the 5'-N \downarrow pAb-3' to give new termini consisting of an abasic 5'-phosphate and a normal 3'-OH. This is exactly the same phosphodiester

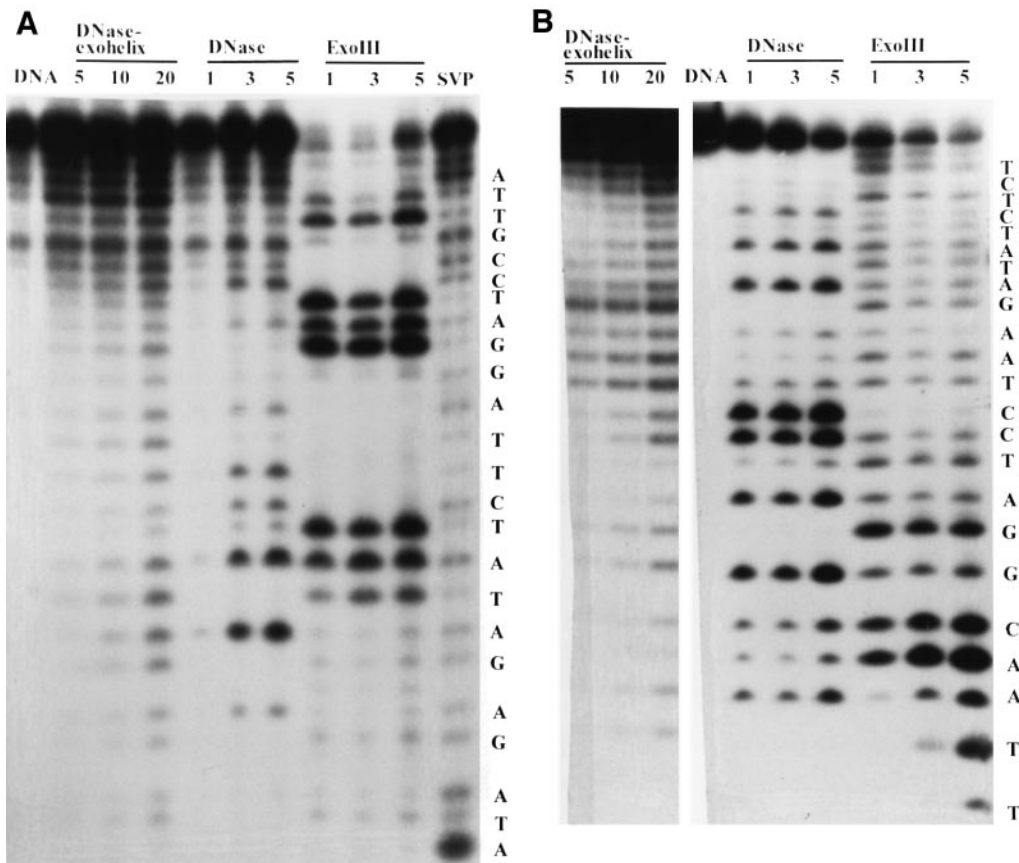


Fig. 5. Hydrolysis of the following duplex:

(A) 5'-pAGATAGAGATATCTTAGGATCCGTTAAGG-3'

(B) 3'-TATCTCTATAGAATCCTAGGCAATTCAGAp-5'

by DNase-exohelix (1 µg), DNase I (1.5 ng) or exonuclease III (120 units). The two strands (identified as A and B) were individually labelled with ^{32}P at their 5'-termini and digested with the nucleases for the times (in minutes) indicated. For DNase-exohelix and DNase I 0.5 mM CaCl_2 was present in the buffer. With exonuclease III the manufacturer's recommended buffer, lacking CaCl_2 was used. The hydrolysis patterns were evaluated, for each strand (A) and (B), by gel electrophoresis followed by autoradiography. The lanes marked 'DNA' represent the oligonucleotides prior to nuclease addition. Individual bands were identified using the *EcoRV* (GATATC) and *BamHI* (GGATCC) sites present in the oligonucleotides (not shown) and digestion with snake venom phosphodiesterase [shown for (A) only, SVP]. Band identity is given on the sides of the autoradiogram.

bond that exonuclease III cuts in response to abasic sites (Rogers and Weiss, 1980; Demple and Harrison, 1994). Controls, consisting of oligonucleotides in which the abasic site was replaced by the natural Watson-Crick hydrogen-bonding base, did not yield a hypersensitive site with DNase-exohelix. This indicates that the sensitivity of the two abasic sites towards DNase-exohelix is not caused by indirect effects e.g. the abasic sites having been placed in a sequence context that fortuitously generated a site that was well cut by DNase I. Furthermore, wild-type DNase I shows no enhanced specificity for the abasic sites. This indicates that these sites are not incidentally sensitive to DNase I and clearly demonstrates that the polypeptide insert is absolutely essential for effective cutting at the abasic location. Although DNase-exohelix strongly prefers abasic sites, cuts at other positions are seen, as would be expected from the results obtained with the short oligonucleotide mentioned above. This is most clearly visible with the gel obtained using the true abasic site, where weak cutting at other locations can be observed. In the case of 1',2'-dideoxyribose, where half the quantity of DNase-exohelix was used to clearly reveal the best-cut location, specificity for the abasic site appears near absolute. Some enhancement of cutting at the nucleotide

5' to the abasic site was also observed in the case of the true abasic site. We assume this arises from an initial endonucleolytic cut at the abasic site, followed by 3'-exonuclease activity at the newly generated 3'-OH terminus. Figure 6 also shows the results seen with exonuclease III. Cutting at the abasic site does not appear to be as prominent as with DNase-exohelix. However, many bands resulting from fragments of a smaller size than would be obtained from cleavage at the abasic site are prominent. This probably arises from a primary endonucleolytic cut at the abasic site followed by 3'-exonuclease activity. The DNA hydrolysis patterns seen with DNase-exohelix and exonuclease III show that each enzyme has a different activity ratio for endonucleolysis at abasic sites relative to 3'-exonucleolysis. Clearly this ratio is high for DNase-exohelix and this results in accumulation of products cut at the abasic site. Finally, we have examined cutting of the DNA strand complementary to that containing the abasic sites. No hypersensitive sites, e.g. at the base directly opposite or nearly opposite the abasic site, were observed with DNase-exohelix and the cutting patterns seen for a complementary strand hybridized to either the abasic or the control strand were almost identical (not shown).

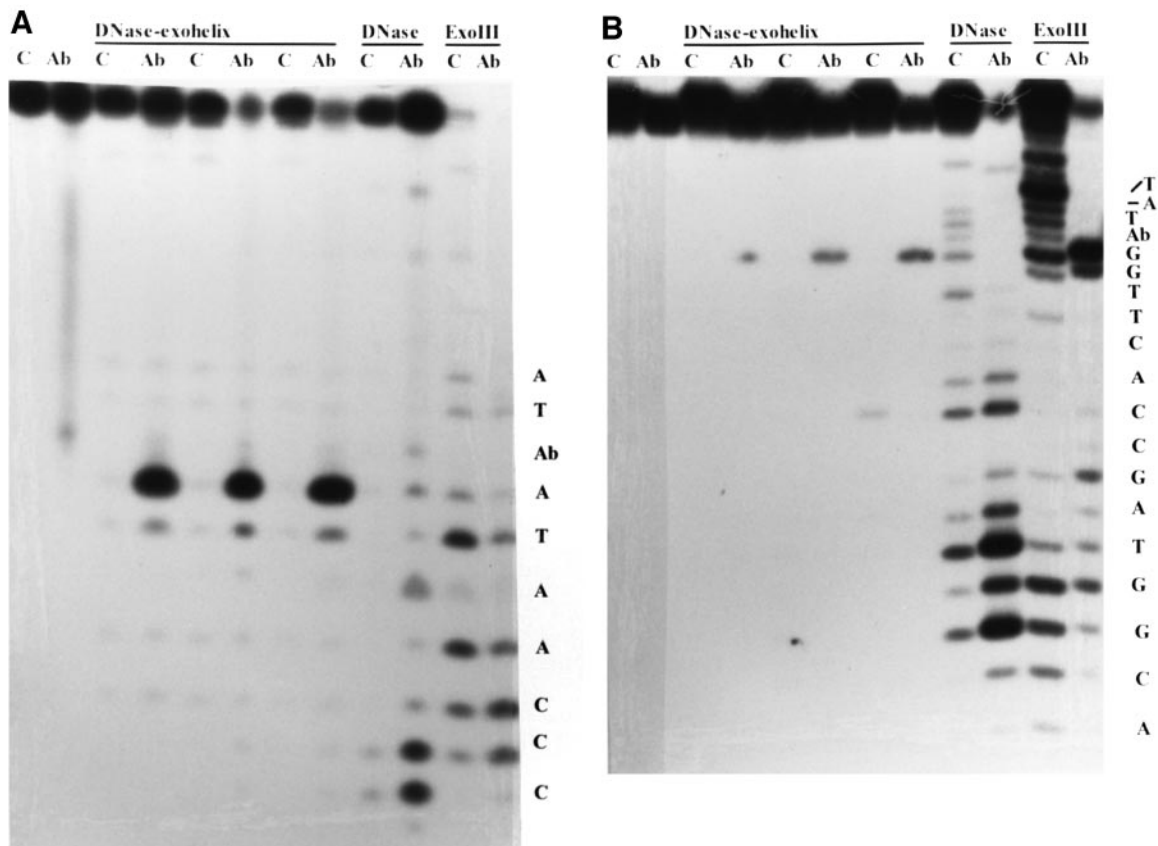


Fig. 6. Hydrolysis of oligonucleotides containing abasic sites. The following duplexes:

- (A) 5'-³²pTATCCCAATAAXTAAGCTTCCTATC-3'
 3'-ATAGGGTTATAATT CGAAGGATAG-5'
 (B) 5'-³²pCCCCACGGTAGCCACTTGGZTATCCACAGACT-3'
 3'-GGGGTGCCATCGGTGAACCTATAGGTGTCTGA-3'

(X = either T [control oligonucleotide] or an abasic site [2'-deoxyribose]; Z is either dA [control oligonucleotide] or a stable abasic site analogue [1',2'-dideoxyribose]) were treated with DNase-exohelix (1 µg for duplex A, 0.5 µg for duplex B; 10 min), DNase I (1.5 ng; 2 min) or exonuclease III (140 units; 2 min) and the products were analysed by gel electrophoresis followed by autoradiography. The lanes marked 'Ab' represent the results found with oligonucleotides containing an abasic site whereas lanes marked 'C' are the controls. Tracks corresponding to the oligonucleotides prior to nuclease addition (labelled 'C' or 'Ab' only) are also given. The three sets of experiments with DNase-exohelix used 0 (left-hand lanes), 0.5 (middle lanes) and 2 (right-hand lanes) mM CaCl₂. With DNase I only 2 mM CaCl₂ was used and this was omitted for exonuclease III. Band identification made use of the *Hind*III site present in duplex A and the *Eco*RV site present in duplex B, together with a snake venom phosphodiesterase ladder (not shown). Band identity is given on the sides of the autoradiogram with the abasic position represented by Ab.

Discussion

The key difference between exonuclease III/HAP 1 and DNase I lies in their endonucleolytic cleavage selectivity. DNase I is a low-selectivity endonuclease with no activity at abasic sites whereas the two repair enzymes only cut DNA in response to this lesion. This publication shows that insertion of an α -helix, derived from exonuclease III, into DNase I results in a hybrid with a very marked preference for abasic sites. This helix, which represents one of the main differences between the two classes of nuclease, clearly plays an important role in the discrimination between abasic and normal sites. We cannot be sure if the polypeptide actually forms an α -helix in the context of an insertion into DNase I but this seems a reasonable assumption, particularly as it is flanked by glycine-rich flexible linkers. Whatever the final structure adopted, polypeptide insertion clearly results in a preference for abasic over normal sites.

We believe that the helix does not contact the abasic site directly. Models for the interaction between exonuclease III (Mol *et al.*, 1995) and HAP 1 (Gorman *et al.*, 1997),

based on the DNase I-DNA interaction, show that the helix probably inserts into the DNA major groove, at some distance from the abasic site. Furthermore, the amino acid sequence (both of the α -helix and its flanking loop amino acids) and hence the structure and interaction with DNA of this element vary between the three proteins. This would argue against its use in a precise recognition of the abasic site. Rather it seems that several of the positively charged and polar amino acids present in the helix contact phosphate groups to stabilize an interaction in the major groove (Mol *et al.*, 1995; Gorman *et al.*, 1997). These additional contacts probably stabilize a non-productive protein-DNA complex and suppress a further distortion of the DNA, necessary to place the scissile phosphate group at the active site. This would be expected to lead to a reduction in cutting at normal phosphodiester sites. Our results, which show a reasonably competent active site [assessed by T(npp)₂ hydrolysis] but very poor cutting of DNA (measured in the Kunitz assay), are compatible with this idea. Cleavage at abasic sites may be promoted by two mechanisms. First, the loss of Watson-Crick

hydrogen bonds and stacking interactions should increase the flexibility of DNA at abasic sites. Therefore, the energy barrier of the final DNA deformation, required for the attainment of a catalytically active complex, should be reduced leading to reasonable hydrolysis. Secondly, as pointed out previously (Gorman *et al.*, 1997), NMR studies of oligonucleotides have shown that the sugar at an abasic site exists as an equilibrium between an 'intra' (normal) and 'extra' DNA-helical orientation (Goljer *et al.*, 1995). Catalysis may arise by the sugar flipping into the 'extra' helical conformation and thus positioning the scissile phosphate at the active site. In this scenario, a large global DNA conformational change, seen with DNase I and normal oligonucleotides is replaced by a much smaller local one. In a recent study with HAP 1 (Erzberger *et al.*, 1998) it was concluded that abasic DNA, but not unmodified DNA, undergoes a unique conformational change following protein binding which in the presence of Mg^{2+} leads to incision. We agree with this point of view and this publication shows that the α -helix is a very important factor in this key conformational shift, that involves both the DNA and the protein.

DNase I-exohelix also has a 3'-exonuclease activity, suggesting that the helix also contributes to this activity. With exonuclease III (Mol *et al.*, 1975) it was suggested that the extra α -helix could act as a pivot point to tilt the DNA and place the end of the helix at the active site. A similar mechanism could be in operation here. However, a DNase I variant, in which all contacts to the DNA minor groove were removed (Doherty *et al.*, 1995), also showed 3'-exonuclease activity. Thus, the inserted α -helix may also contribute indirectly to this activity by interfering with the normal protein-DNA minor groove contacts. Finally, although DNase-exohelix has a much stronger specificity for abasic sites than DNase I, its preference for abasic over normal sites, is still not as high as exonuclease III or HAP 1. Thus the endonucleolytic activity of the hybrid protein can still cut at normal bases, whereas endonucleolytic action of the two repair enzymes is largely limited to abasic sites alone. Two factors may account for this. First, the two loop regions, additional to the helix studied here, which are also present in the repair enzyme and absent from DNase I may play a role. Furthermore, stabilization of an 'extra-helical' sugar at an abasic site by hydrophobic stacking with Phe266 has been proposed for HAP 1 (Gorman *et al.*, 1995). An aromatic amino acid in this position is conserved in all AP nucleases, but is replaced by threonine in DNase I. However, it has recently been shown that mutagenesis of this amino acid leads to only a 6-fold reduction in specific activity and binding at abasic sites, suggesting it cannot play a critical role (Erzberger *et al.*, 1998). We are currently investigating the functions of these elements by further mutagenesis of DNase I. Secondly, if recognition of abasic sites is dominated by a DNA-protein conformational change, rather than specific recognition of the missing base site, one might expect protein structural integrity to be important. In this scenario the additional helix used might not be the optimal one in the DNase I context and may allow the key conformational change, albeit at a very poor rate, with normal oligonucleotides as well as those containing abasic sites. This may be addressed by variation in the inserted helix and its flexible flanking regions. Together,

the two approaches may produce a nuclease with a higher, and hopefully absolute, preference for abasic positions.

Materials and methods

Construction of a gene coding for DNase-exohelix

A gene coding for wild-type DNase I in M13mp19 (DNase I rec 5; Worrall and Connolly, 1990) was used as the starting material. A restriction endonuclease site for *Apa*LI (GTGCAC) was introduced near the region of the gene coding for Ser174 by the PCR-based overlap extension technique (Ho *et al.*, 1989). The standard M13 forward and backward primers were used as the outer primers, together with (as the internal primers): GGTGACTTCAACGCTGAGTGCACCTTATGTAACC and GGTTACATAGTGCACCTCAGCGTTGAAGTCACC. PCR amplification was carried out using Pfu DNA polymerase (Promega) under the following conditions: first cycle; denaturation, 94°C, 3 min; annealing, 55°C, 2 min; polymerization, 72°C, 1 min; subsequent 30 cycles; denaturation, 94°C, 1 min; annealing, 55°C, 3 min; polymerization, 72°C, 2 min. The two PCR products were purified by agarose gel electrophoresis and then digested for 3 h with 12 units of *Apa*LI restriction endonuclease. Following a further gel purification, the fragments containing *Apa*LI cohesive ends were joined with T4 DNA ligase (2 units, overnight reaction) and then digested with 20 units (overnight) of *Eco*RI and *Hind*III restriction endonucleases (these restriction sites flank the DNase I gene). The fragment was cloned into *Eco*RI-*Hind*III-digested M13mp19 and the resulting plasmids were introduced into *E. coli* JM105 and selected on Luria-Bertani (LB) agar plates containing ampicillin and X-gal. Several colonies were picked and used to inoculate 5 ml of liquid LB-ampicillin media. The cultures were grown overnight at 37°C, after which small scale plasmid preparations were carried out. The presence of a recombinant plasmid was initially checked by DNA digestion with *Eco*RI and *Hind*III and the release of a fragment about 816 base pairs in length (the size of the wild-type DNase I gene). The presence of an *Apa*LI site was also monitored by digestion with this enzyme. Several plasmids, having the necessary restriction endonuclease cleavage pattern, were completely sequenced, in both strands, and one having the desired sequence (i.e. wild-type DNase I plus a unique *Apa*LI site and no other changes) was used in the next step.

The M13mp19 recombinant was cut with *Apa*LI and the duplex produced by mixing pTGCAGCGGCAAGAGAACAGAAAAAGG-TGGCTTAGAAGTGGG together with pTGCAAAAAGTTCTAAGC-CACCTTTTCTGTTCTCTTCGCCGC (phosphorylation carried out with polynucleotide kinase) was inserted at this site using DNA ligase. This duplex has *Apa*LI cohesive ends and encodes the exonuclease III α -helix (EENRKRWLRT) flanked by glycine residues. The resulting plasmids were introduced into *E. coli* JM105 and further treated (i.e. tested for appropriate restriction endonuclease sites and a DNase I insert of the correct size, followed by complete sequencing of the modified DNase I gene) as described above. One plasmid, the integrity of which had been completely established by sequencing of the entire gene, was used for the next stage.

The unwanted, extra, cysteine was removed from the M13mp19 plasmid, containing the DNase-exohelix gene, by the PCR overlap extension method exactly as described above. The internal primers used were ATCCGGCCGTTATGTAACCTCTTC and GACTCGGCCGCCA-GTTCTTAAGCCACCTTTTGTG. This manipulation also introduces a unique *Eag*I site and converts a Cys-Thr dipeptide sequence to Gly-Arg (see Figure 2). An M13mp19 plasmid containing the required DNase-exohelix sequence, which had been verified by sequencing of the entire gene was used for protein production.

Overproduction and purification of DNase-exohelix

The gene coding for DNase-exohelix was excised from M13mp19 using *Eco*RI and *Hind*III restriction endonucleases, and purified by gel electrophoresis. The purified gene was inserted into the *Eco*RI and *Hind*III sites of the expression vector pKK223-3 (Pharmacia) using DNA ligase. The resulting plasmids were introduced into *E. coli* JM105 and selected on LB agar plates containing ampicillin. Several colonies were picked and used to inoculate 5 ml of liquid LB-ampicillin media. The cultures were grown for 4 h, at 30°C, and protein expression induced by the addition of 0.4 mM IPTG for three h. Cells were collected by centrifugation and lysed by boiling in SDS. A small amount of the centrifuged lysate was analyzed by SDS-PAGE and the amount of recombinant DNase I observed by Coomassie Blue staining. The colony giving the highest levels of protein was grown on a large scale

(8×500 ml), at 30°C until the cultures had an OD₅₅₀ of ~0.5 (~4 h). IPTG (0.4 mM) was added and the cultures grown for a further 3.5 h. The cells (8 g) were collected by centrifugation and disrupted, on ice, by suspension followed by sonication in 20 mM Tris pH 8.2, 2 mM CaCl₂, 0.1 mM benzamidine, 0.1 mM phenylmethylsulphonyl fluoride (buffer A) supplemented with 0.5 M NaCl. Following centrifugation, the supernatant was diluted with buffer A, to a NaCl concentration of 50 mM. The diluted supernatant was immediately applied (cold-room temperature) to a 15×3 cm Q-Sepharose Fast-Flow (Pharmacia) column equilibrated to buffer A containing 50 mM NaCl. The protein was eluted with a 50 mM to 1 M NaCl gradient, in buffer A, over 1 l. Fractions were monitored by: (i) SDS-PAGE with Coomassie Blue staining; (ii) Western blotting with a polyclonal antibody raised against commercially available DNase I purified from bovine pancreas (Jones, 1994). This used an enhanced chemiluminescence (ECL) kit supplied by Amersham. (iii) Digestion of phage λ DNA (Sadwoski and Hurwitz, 1969). The purest fractions were pooled and concentrated to 2 ml using Centriprep spin-concentrators (Amicon). The protein solution (as eight 0.25 ml aliquots) was applied to a Superdex-G75 gel filtration column (Pharmacia) connected to a Waters Advanced Protein Purification System. The column was equilibrated to, and eluted with, 20 mM Tris pH 8.2, 2 mM CaCl₂, 50 mM NaCl (buffer B) at a flow rate of 0.75 ml/min. The purest fractions (detected as above) were pooled and concentrated to 2 ml using a Centriprep. Final purification used a Protein-Pak DEAE-5PW (Waters) column connected to a Waters Advanced Protein Purification System. The column was equilibrated with buffer B and eluted with a gradient formed from this buffer and one containing 1 M NaCl. Gradient Profile: *t* = 0, 100% 50 mM NaCl, 0% 1 M NaCl; *t* = 5 min, 100% 50 mM NaCl, 0% 1 M NaCl; *t* = 10 min, 50% 50 mM NaCl, 50% 1 M NaCl; *t* = 30 min, 25% 50 mM NaCl, 75% 1 M NaCl. Fractions were analyzed as described above and only those that were pure were pooled and concentrated to a small volume (Centriprep). The protein was stored in 30% glycerol at -20°C. Purity was assessed by SDS-PAGE. The identity of the proteins was confirmed N-terminal sequencing, Western blotting and phage λ DNA digestion.

Hydrolysis of T(npp)₂ and DNA by DNase-exohelix

The hydrolysis of T(npp)₂, at concentrations of 5, 10 and 20 mM was carried out as described previously (Doherty *et al.*, 1995), in 0.5 ml of 25 mM Tris pH 7.2, 1 mM CaCl, and either 10 mM MgCl₂ or 10 mM MnCl₂. Wild-type DNase I (0.5 µg) and 4 µg of DNase-exohelix were used. Kunitz assays, which measure the hydrolysis of DNA by observing the increase in absorbance at 260 nm, were performed as described (Kunitz, 1950; Doherty *et al.*, 1995) in 1 ml of 10 mM Tris pH 8, 1 mM MgCl₂, 0.1 mM CaCl₂ and between 0.001 and 0.1 mg/ml sonicated calf thymus DNA as substrate. The course of the reaction was monitored at 30°C and initiated by adding a quantity of DNase-exohelix (typically 30 µg) sufficient to give a measurable rate over ~30 min. Controls with wild-type DNase I (typically 0.02 µg) were carried out at the same time.

Measurement of 3'-exonuclease activity

The activity of 3'-exonuclease was measured using two oligonucleotides: ATTGCATGATATCTAAGGGAAAAGATATCATGC and ATGATATCTTTTCCCTTAGATATCATGC, designed to form stem-loop structures. A single base was added to the 3'-end of both oligonucleotides using [α -³²P]ddATP and terminal deoxynucleotidyl transferase (Sambrook *et al.*, 1989). Any 3'-exonuclease activity of DNase-exohelix was measured in 20 µl of 20 mM Tris pH 7.2 containing 10 mM MgCl₂ and either 0, 0.5 or 2 mM CaCl₂. Oligonucleotide concentrations of between 5 and 10 nM (≈1–3 ng) were used and 1 µg of the enzyme. For wild-type DNase I [obtained from Boehringer Mannheim and further purified as described by Worrall and Connolly (1990)] 1.5 ng of protein was used and only the buffer containing 2 mM CaCl₂. In the case of exonuclease III (supplied by Promega and used without further purification) 120 units of enzyme (as defined by the supplier) and the buffer provided by Promega (66 mM Tris pH 8, 0.66 mM MgCl₂) were used. Hydrolysis products were examined by gel electrophoresis (Doherty *et al.*, 1995). The EcoRV restriction site present in both oligonucleotides and snake venom phosphodiesterase hydrolysis were used for band assignment. Identical buffer conditions, concentrations of enzyme and oligonucleotide and electrophoretic analysis method were used in the assays outlined below unless otherwise stated.

Evaluation of DNA hydrolysis selectivity of DNase-exohelix

A duplex formed by annealing AGATAGAGATATCTTAGATCCGTAAAGG and AGACCTTAACGGATCCTAAGATATCTCTAT was used to study DNA-cleavage selectivity. Strands were individually

labelled with ³²P at their 5' ends (as above), allowing the hydrolysis of both to be followed. With DNase-exohelix and DNase I only the 0.5 mM CaCl₂ buffer was used. The EcoRV and BamHI sites, together with snake venom phosphodiesterase digestion were used for band assignment.

Measurement of RNase activity

A duplex formed by annealing r(AGAUAGAGAUUCUUAGGA-UCCGUUAAGU) and AGAACTTAACGGATCCTAAGAGATCTCT was used for measuring RNase activity. For DNase-exohelix and DNase I only the 2 mM CaCl₂ buffer was used. The RNA strand was synthesized and purified as described (McGregor *et al.*, 1996). The RNA was labelled at its 5'-OH with ³²P (as above) and annealed with the complementary DNA strand. Hydrolysis of the DNA was also tested by 5'-labelling of this strand. In these experiments diethylpyrocarbonate-treated water and RNase-free reagents were used.

Measurement of AP endonuclease activity

Two duplexes were used to measure AP endonuclease activity. One was produced by the hybridization of 5'-[³²P]TATCCCAATAAXTAAGCTT-CCTATC with a complementary strand. In this case X represents a true abasic site (i.e. has 2'-deoxyribose at this position) or T in a control. This oligonucleotide was produced by incubating 5'-[³²P]TATCCCAATAUTAAGCTTCTATC with uracil DNA glycosylase (Gu *et al.*, 1994) and purifying the resulting abasic site-containing strand by gel electrophoresis. A second oligonucleotide had the sequence CCCACGGTAGCCAATTGGZTATCCACAGACT. Here Z is a stable analogue of an abasic site; 1',2'-dideoxyribose or dA in a control. This oligonucleotide was chemically synthesized using the commercially available phosphoramidite dSpacer (Glen Research). Following synthesis and purification, this oligonucleotide was labelled at its 5' end with ³²P and hybridized to a complementary sequence. Both duplexes (either containing 2'-deoxyribose/1',2'-dideoxyribose or the controls) were used to check AP endonuclease activity of DNase-exohelix, DNase I and exonuclease III. Cutting of the complementary strands were also checked by appropriate radioactive labelling. Band assignment made use of the HindIII and EcoRV sites in the true abasic- and stable abasic-site oligonucleotides together with snake venom phosphodiesterase digestion.

Acknowledgements

We thank Ms Pauline Heslop for expert technical assistance and Dr Dietrich Suck (EMBL, Heidelberg) for supplying the co-ordinates on which Figure 1 is based. This work was supported by the UK BBSRC, MRC and Wellcome Trust. S.C. was the recipient of a Fellowship from the Ministerio Español de Educacion y Ciencia, Spain.

References

- Barzily,G. and Hickson,I.D. (1995) Structure and function of apurinic/aprimidinic endonucleases. *BioEssays*, **17**, 713–719.
- Barzily,G., Mol,C.D., Robson,C.N., Walker,L.J., Cunningham,R.P., Tainer,J.A. and Hickson,I.D. (1995) Identification of critical active-site residues in the multifunctional human DNA repair enzyme HAP1. *Nature Struct. Biol.*, **2**, 561–567.
- Brosius,J. and Holy,A. (1984) Regulation of ribosomal RNA promoters with a synthetic lac operator. *Proc. Natl Acad. Sci. USA*, **81**, 6929–6933.
- Bruckner,I., Jurvkovski,A. and Savic,A. (1990) Sequence dependent structural variations of DNA revealed by DNase I. *Nucleic Acids Res.*, **18**, 891–894.
- Demple,B. and Harrison,L. (1994) Repair of oxidative damage to DNA: enzymology and biology. *Annu. Rev. Biochem.*, **63**, 915–948.
- Doherty,A.J., Worrall,A.F. and Connolly,B.A. (1991) Mutagenesis of the DNA binding residues in bovine pancreatic DNase I: an investigation into the mechanism of sequence discrimination by a sequence-selective nuclease. *Nucleic Acids Res.*, **19**, 6129–6132.
- Doherty,A.J., Connolly,B.A. and Worrall,A.F. (1993) Overproduction of the toxic protein, bovine pancreatic DNase I, in *Escherichia coli* using a tightly controlled T7-promoter based system. *Gene*, **136**, 337–340.
- Doherty,A.J., Worrall,A.F. and Connolly,B.A. (1995) The roles of arginine 41 and tyrosine 76 in the coupling of DNA recognition to phosphodiester bond cleavage by DNase I: a study using site-directed mutagenesis. *J. Mol. Biol.*, **251**, 366–377.
- Drew,H.R. and Travers,A.A. (1984) DNA structural variations in the *E.coli* Tyr T promoter. *Cell*, **37**, 491–502.

- Drew,H.R. and Travers,A.A. (1985) Structural junctions in DNA: the influence of flanking sequences on nuclease digestion specificities. *Nucleic Acids Res.*, **13**, 4445–4467.
- Erzberger,J.P., Barsky,D., Schärer,O.D., Colvin,M.E. and Wilson,D.M.,III (1998) Elements in abasic site recognition by the major *Escherichia coli* apurinic/apyrimidinic endonucleases. *Nucleic Acids Res.*, **26**, 2771–2778.
- Goljer,I., Kumar,S. and Bolton.P.H. (1995) Refined solution structure of a DNA heteroduplex containing an aldehydic abasic site. *J. Biol. Chem.*, **270**, 22980–22987.
- Gorman,M.A., Morera,S., Rothwell,D.G., de La Fortelle,E., Mol,C.D., Tainer,J.A., Hickson,I.D. and Freemont,P.S. (1997) The crystal structure of the human DNA repair endonuclease HAP1 suggests the recognition of extra-helical deoxyribose at DNA abasic sites. *EMBO J.*, **16**, 6548–6558.
- Gu,L., Huang,S.-M. and Sander,M. (1994) Single amino acid changes alter the repair specificity of *Drosophila* Rrp1. *J. Biol. Chem.*, **51**, 32685–32692.
- Harrison,S.C. (1991) A structural taxonomy of DNA-binding domains. *Nature*, **353**, 715–719.
- Ho,S.N., Horton,R.M. and Pease,I.R. (1989) Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene*, **77**, 61–68.
- Hogan,M.E., Robertson,M.W. and Austin,R.H. (1989) DNA flexibility may dominate DNase I cleavage. *Proc. Natl Acad. Sci. USA*, **86**, 9273–9277.
- Jones,S.J. (1994) Studies on the active site mutants of bovine pancreatic DNase I. PhD thesis, University of Newcastle, UK.
- Jones,S.J., Worrall,A.F. and Connolly,B.A. (1996) Site-directed mutagenesis of the catalytic residues of bovine pancreatic deoxyribonuclease I. *J. Mol. Biol.*, **264**, 1154–1163.
- Katayanagi,K., Miyagawa,M., Matsushima,M., Ishikawa,M., Kanaya,S., Ikehara,M., Matsuzaki,T. and Morikawa,K. (1990) Three-dimensional structure of RNase H from *E.coli*. *Nature*, **347**, 306–309.
- Kunitz,M. (1950) Crystalline deoxyribonuclease I. Isolation and general properties: spectrophotometric method for the measurement of deoxyribonuclease activity. *J. Gen. Physiol.*, **33**, 363–377.
- Lahm,A. and Suck,D. (1991) DNase I-induced DNA conformation: 2 Å structure of a DNase I-octamer complex. *J. Mol. Biol.*, **221**, 645–667.
- Liao,T.-H. (1975) Deoxythymidine-3',5'-di-*p*-nitrophenylphosphate as a synthetic substrate for bovine pancreatic deoxyribonuclease. *J. Biol. Chem.*, **250**, 3721–3724.
- Lindahl,T., Ljungquist,S., Siebert,W., Nyberg,B. and Sperens,B. (1977) DNA-N glycosidases: properties of uracil-DNA glycosidase from *Escherichia coli*. *J. Biol. Chem.*, **252**, 3286–3294.
- Lloyd,R.S. and Linn,S. (1993) Nucleases involved in DNA repair. In Linn,S.M., Lloyd,R.S. and Roberts,R.J. (eds), *Nucleases*. 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 263–316.
- Lomonosoff,G.P., Butler,P.J.G. and Klug,A. (1981) Sequence dependent structural variation in the conformation of DNA. *J. Mol. Biol.*, **149**, 745–760.
- McGregor,A., Vaman Rao,M., Duckworth,G., Stockley,P.G. and Connolly,B.A. (1996) Preparation of oligoribonucleotides containing 4-thiouridine using Fpmp chemistry. Photo-crosslinking to RNA binding proteins using 350 nm irradiation. *Nucleic Acids Res.*, **24**, 3173–3180.
- Mendoza,R., Markovits,J., Jaffrezou,J.P., Muzzard,G. and LePeq,J.B. (1990) DNase I susceptibility of bent DNA and its alteration by ditercalinium and distamycin. *Biochemistry*, **29**, 5035–5043.
- Mol,C.D., Kuo,C.-F., Thayer,M.M., Cunningham,R.P. and Tainer,J.A. (1995) Structure and function of the multifunctional DNA-repair enzyme exonuclease III. *Nature*, **374**, 381–386.
- Moore,S. (1981) Pancreatic DNase. In Boyer,P.D. (ed.), *The Enzymes*. 3rd edn, Vol. 14. Academic Press, New York, NY, pp. 281–296.
- Oefner,C. and Suck,D. (1986) Crystallographic refinement of DNase I at 2 Å resolution. *J. Mol. Biol.*, **192**, 605–632.
- Pabo,C.O. and Sauer,R.T. (1992) Transcription factors: structural families and principles of DNA recognition. *Annu. Rev. Biochem.*, **61**, 1053–1095.
- Price,P.A. (1972) Characterisation of Ca²⁺ and Mg²⁺ binding to bovine pancreatic deoxyribonuclease A. *J. Biol. Chem.*, **247**, 2895–2899.
- Price,P.A. (1975) The essential role of Ca²⁺ in the activity of bovine pancreatic deoxyribonuclease. *J. Biol. Chem.*, **250**, 1981–1982.
- Price,P.A., Liu, T.-Y., Stein,W.H. and Moore,S. (1969a) Properties of chromatographically purified bovine pancreatic deoxyribonuclease. *J. Biol. Chem.*, **244**, 917–923.
- Price,P.A., Stein,W.H. and Moore,S. (1969b) Effect of divalent cations on the reduction and re-formation of the disulfide bonds of deoxyribonuclease. *J. Biol. Chem.*, **244**, 929–932.
- Rogers,S.G. and Weiss,B. (1980) Exonuclease III of *Escherichia coli* K-12, an AP endonuclease. *Methods Enzymol.*, **65**, 201–211.
- Sadowski,P.D. and Hurwitz,J. (1969) Enzymatic breakage of deoxyribonucleic acid: purification and properties of endonuclease II from T4-phage infected *E.coli*. *J. Biol. Chem.*, **244**, 6192–6198.
- Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Seeman,N.C., Rosenberg,J.M. and Rich,A. (1976) Sequence specific recognition of double helical nucleic acids by proteins. *Proc. Natl Acad. Sci. USA*, **73**, 804–808.
- Shida,T., Noda,M. and Sekiguchi,J. (1996) Cleavage of single- and double-stranded DNAs containing an abasic residue by *Escherichia coli* exonuclease III (AP endonuclease VI). *Nucleic Acids Res.*, **24**, 4572–4576.
- Steitz,T.A. (1990) Structural studies of protein–nucleic acid interaction: the sources of sequence specific binding. *Q. Rev. Biophys.*, **23**, 205–280.
- Suck,D. and Oefner,C. (1986) Structure of DNase I at 2 Å resolution suggests a mechanism for binding to and cutting DNA. *Nature*, **321**, 620–625.
- Suck,D., Oefner,C. and Kabsch,W. (1984) Three-dimensional structure of bovine pancreatic DNase I at 2.5 Å resolution. *EMBO J.*, **3**, 2423–2430.
- Suck,D., Lahm,A. and Oefner,C. (1988) Structure refined to 2 Å of a nicked DNA oligonucleotide complex with DNase I. *Nature*, **332**, 465–468.
- Travers,A.A. (1993) *DNA–Protein Interactions*. Chapman and Hall, London, UK.
- Warren,M.A., Evans,S.J. and Connolly,B.A. (1997) Effects of non-conservative changes to tyrosine, 76, a key DNA binding residue of DNase I, on phosphodiester bond cleavage and DNA hydrolysis selectivity. *Protein Eng.*, **10**, 279–283.
- Weston,S., Lahm,A. and Suck,D. (1992) X-ray structure of the DNase I-d(GGTATACC)₂ complex at 2.3 Å resolution. *J. Mol. Biol.*, **226**, 1237–1256.
- Wolf,E., Bruckner,I. and Suck,D. (1995) Mutational analysis of DNase I–DNA interactions: design, expression and characterization of a DNase-I loop-insertion mutant with altered selectivity. *Protein Eng.*, **8**, 283–291.
- Worrall,A.F. and Connolly,B.A., (1990) The chemical synthesis of a gene coding for bovine pancreatic DNase I and its cloning and expression in *E.coli*. *J. Biol. Chem.*, **265**, 21889–21895.
- Yang,W. and Steitz,T.A. (1995) Recombining the structures of HIV integrase, RuvC and RNase H. *Structure*, **3**, 131–134.

Received August 18, 1998; revised September 28, 1998;
accepted October 15, 1998