Nucleic Acids Research, 1998, Vol. 26, No. 11 **2771–2778**

Elements in abasic site recognition by the major human and *Escherichia coli* apurinic/apyrimidinic endonucleases

Jan P. Erzberger, Daniel Barsky, Orlando D. Schärer[1,+], Michael E. Colvin and David M. Wilson, III*

Biology and Biotechnology Research Program, L-452, Lawrence Livermore National Laboratory, Livermore, CA 94551, USA and [1]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA

Received December 22, 1997; Revised and Accepted April 10, 1998

ABSTRACT

Sites of base loss in DNA arise spontaneously, are induced by damaging agents or are generated by DNA glycosylases. Repair of these potentially mutagenic or lethal lesions is carried out by apurinic/apyrimidinic (AP) endonucleases. To test current models of AP site recognition, we examined the effects of site-specific DNA structural modifications and an F266A mutation on incision and protein–DNA complex formation by the major human AP endonuclease, Ape. Changing the ring component of the abasic site from a neutral tetrahydrofuran (F) to a positively charged pyrrolidine had only a 4-fold effect on the binding capacity of Ape. A non-polar 4-methylindole base analog opposite F had a <2-fold effect on the incision activity of Ape and the human protein was unable to incise or specifically bind 'bulged' DNA substrates. Mutant Ape F266A protein complexed with F-containing DNA with only a 6-fold reduced affinity relative to wild-type protein. Similar studies are described using *Escherichia coli* AP endonucleases, exonuclease III and endonuclease IV. The results, in combination with previous findings, indicate that the ring structure of an AP site, the base opposite an AP site, the conformation of AP-DNA prior to protein binding and the F266 residue of Ape are not critical elements in targeted recognition by AP endonucleases.

INTRODUCTION

Genomic integrity depends on the efficient removal and repair of DNA lesions prior to the fixation of genetic changes (1). Base excision repair (BER), the major defense system against most oxidative and alkylation damage, is initiated by DNA glycosylases that remove inappropriate (damaged or mismatched) bases from DNA (2). Sites of base loss can also be formed via spontaneous or mutagen-induced hydrolysis of the *N*-glycosidic bond (3). In most cases, repair of these abasic sites, which are potentially cytotoxic and mutagenic (4), is initiated by class II apurinic/apyrimidinic (AP) endonucleases, essential components of BER (5,6).

AP endonucleases cleave the phosphodiester bond immediately 5′ of an AP site, producing a 3′-hydroxyl group and a 5′-abasic residue (5). The major human AP site repair enzyme is Ape (also known as Hap1, Apex and Ref1; 7–10). Following incision of a natural abasic site by Ape, the 5′-abasic residue is removed by polymerase β (polβ) (11), which also fills the single nucleotide gap (12). A 'long patch' BER pathway that operates through strand displacement and that involves Fen1 and PCNA has also been described in eukaryotes (13–16), yet the details of this pathway are still under investigation (17). To complete the BER process, the final nick is sealed by DNA ligase I (18) or an XRCC1–DNA ligase III complex (19).

AP endonucleases have been classified into two families based on amino acid sequence homology to *Escherichia coli* exonuclease III (ExoIII) or endonuclease IV (EndoIV) (5). For the ExoIII family, which includes Ape, the cleavage reaction proceeds through a hydrolytic, Mg^{2+}-stimulated mechanism (20,21). The catalytic mechanism of the EndoIV family, which includes Apn1 of *Saccharomyces cerevisiae* (5), is unknown.

Several substrate modifications have been employed to identify DNA elements that are important for recognition and/or incision (22–25). The human enzyme requires at least 4 bp 5′ and at least 3 bp 3′ of an AP site for incision and makes contacts within both the minor and major groove and with both strands of DNA around the lesion (24,25). Ape displays a minor exonuclease or 3′-diesterase activity when the target site is flanked by duplex DNA and a reduced endonuclease activity when a mismatch is present immediately 5′ of the AP site, but is not affected by a 3′ mismatch (24,26). Ape also displays a similar incision activity regardless of the base opposite in at least one sequence context, is unable to effectively incise substrates containing a branched structure [i.e. the 2-(aminobutyl)-1,3-propanediol abasic site analog] and attacks the target 5′-phosphodiester bond in a stereospecific manner (24). The homologous ExoIII protein displays a pattern of specificity similar to Ape, whereas the EndoIV homologs exhibit several significant differences (27).

*To whom correspondence should be addressed. Tel: +1 925 423 0695; Fax: +1 925 422 2282; Email: wilson61@llnl.gov

[+]Present address: Department of Cell Biology and Genetics, Erasmus University, PO Box 1738, 3000 DR Rotterdam, The Netherlands

Most notably, EndoIV family members incise branched structures efficiently and display the reverse stereospecificity (23). Different substrate specificities have also been observed for the distinct AP endonuclease families at oxidized AP sites and α-deoxyadenosine residues (28,29). While these studies have provided much useful data relating to the mechanism of AP endonuclease activity, a consensus for the recognition mechanism has not yet been reached.

There are four main hypotheses for damage-specific recognition by AP endonucleases.

Model I. Gorman *et al.* (30) suggested, based on the crystal structure of Ape, that the abasic ribose ring structure serves as the specific recognition structure by rotating out of the helix to interact with residue F266 of Ape.

Model II. Mol *et al.* (20), based on the crystal structure of ExoIII, suggested that binding might be mediated by recognition of an extrahelical base opposite the AP site.

Model III. The abasic site produces a specific local distortion in the DNA molecule that is recognized by the repair protein.

Model IV. Shida *et al.* (31) have suggested that insertion of the amino acid residue W212 of ExoIII (which, based on structural alignment between the two proteins, corresponds to F266 in Ape) into the AP site gap might serve as the essential recognition step.

To test models I–III, we examined the effect of a positively charged pyrrolidine (N) abasic site ring structure [an inhibitor of several DNA glycosylases (32)], a non-polar 4-methylindole (4meInd) residue directly opposite the prototypical abasic site analog tetrahydrofuran (F), and an additional centrally located ('bulged') base on the incision and binding activity of Ape, ExoIII and EndoIV; incision of F was used as a reference point (22–25). To test models I and IV, we measured the cleavage and binding activity of the mutant Ape protein F266A. Quantum chemical and molecular dynamics simulations were used to determine the structure, charges and hydration states of the abasic sites and to predict the solvation energies of 4meInd and the natural DNA bases. We discuss the biochemical and computational results as they relate to damage-specific recognition by AP endonucleases.

MATERIALS AND METHODS

Materials and oligonucleotides

[γ-^{32}P]ATP (3000 Ci/mmol) was purchased from Amersham (Arlington Heights, IL) and ExoIII (46 400 U/mg; see below) from Stratagene (La Jolla, CA). EndoIV was purified as described (33). The 25mer oligonucleotides containing a centrally located F or N abasic site analog were synthesized as described (32). Oligonucleotides used to construct the 'bulged' DNA substrates were obtained from Operon (Alameda, CA). The 4meInd-containing 25mer oligonucleotide was synthesized by solid state phosphoramidite chemistry on a Perkin Elmer Synthesizer using 4meInd phosphoramidite obtained from Glen Research (Sterling,VA) and was purified by HPLC to >96% purity as determined by gel analysis of a radiolabeled sample.

Fluorescence

Deoxyribose 4meInd nucleoside was obtained by treating the 4meInd phosphoramidite dissolved in acetonitrile with an equal volume of water to hydrolyze the phosphoramidite group, followed by treatment with an equal volume of glacial acetic acid to remove the DMTO group. The 4meInd nucleoside was then purified by HPLC. Fluorescence analysis was performed on a SPEX (Edison, NJ) 1681 Fluorolog 2 Spectrofluorometer using an excitation wavelength of 295 nm (slit width 1.5 mm, bandpass 2.5 nm). Emission spectra (300–400 nm, slit width 2.5 mm, bandpass 7.5 nm) were recorded and analyzed using the dm3000 software.

Purification of recombinant Ape proteins

Recombinant Ape protein was overexpressed from pETApe, which was produced by subcloning a PCR-generated *APE* cDNA fragment into the *Nde*I and *Bam*HI restriction sites of pET11a (Novagen Madison, WI). Site-specific mutations were generated using the overlapping PCR method described (34) and confirmed by direct sequencing (Lawrence Berkeley Laboratory sequencing facility). Cultures of BL21(λDE3) containing pETApe or a mutant Ape construct were diluted 1:1000 into LB broth containing 50 μg/ml ampicillin and grown for 5 h at 37°C. IPTG was added to a final concentration of 0.5 mM and gene induction carried out for 3 h at 37°C. The cells were harvested at 2000 *g*, washed with 30 ml 0.5× phosphate-buffered saline (PBS), resuspended in 30 ml buffer A (50 mM HEPES, 50 mM KCl, 5% glycerol) and sonicated for 2 × 15 s with a Misonix XL Sonicator (Farmingdale, NY). Debris was removed by centrifugation at 20 000 *g* for 20 min and the supernatant applied to a BioRad Q20 column at a flow rate of 4 ml/min with buffer A using a BioRad (Hercules, CA) BioLogic chromatography system. The Q20 flow-through was applied to a BioRad S10 column and eluted with a 50–700 mM KCl gradient at a flow rate of 2 ml/min. Fractions of 5 ml were collected and peak fractions analyzed on SDS–PAGE gels. Purified wild-type Ape protein (>95% purity) possessed a specific AP endonuclease activity of 5.3×10^5 U/mg as determined by incision at an F residue centrally located in an 18 bp duplex (1 U = 1 pmol substrate converted to product/min) (24).

Quantum chemical methods

The structures of the C3′-hydroxy-C5′-methoxy forms of tetrahydrofuran, α-C1′-hydroxy-trihydrofuran (the natural AP site), pyrrolidine and protonated pyrrolidine were fully optimized using the *ab initio* Hartree–Fock (HF) method (35) with a 6-31G* basis (HF/6-31G*) from a starting conformation with a 3′-*endo* sugar pucker and the 3′ and 5′ C–O bond vectors oriented as in B-DNA. For each structure the sugar pseudorotation phase angle was calculated as previously defined (36). More accurate single point Møller–Plesset perturbation theory energies were calculated at the optimized geometries with a 6-31G** basis set. The atom-centered charges for the endocyclic oxygen and nitrogen were calculated using the method of natural atomic orbital populations (NPA) (37), which have much less sensitivity to basis set and molecular configuration than the commonly used Mulliken populations. All quantum chemical calculations were performed using Gaussian 94 (38).

The structures of 4meInd and all four DNA bases (methyl substituted at the glycosidic bond site) were fully optimized in both the gas and aqueous phases using Hartree–Fock, second order Møller–Plesset perturbation theory and density functional theory (DFT) employing a 6-31G** basis set. The DFT was performed using the Becke three parameter exchange and Lee-Yang–Parr correlation functionals (39), as implemented in

Gaussian 94. The aqueous phase optimizations were performed with the isodensity and self-consistent isodensity polarizable continuum models (PCM) implemented in Gaussian 94 to account for the effect of aqueous solvation (40,41).

DNA substrates and incision assays

To generate duplex substrates, the damage-containing DNA strand was first 5′-end-labeled with [γ-^{32}P]ATP using T4 polynucleotide kinase (24). After kinase inactivation at 95°C for 3 min, unlabeled complementary strand was annealed to the labeled strand by heating to 95°C and then cooling to 4°C at a rate of 1°C/min in a MJ Research (Watertown, MA) PCR machine, yielding 25mer duplex oligonucleotides with the following arrangements (abasic site analog:base opposite): (F:T); (N:T); (F:A); (F:4meInd). 'Bulged' substrates were produced by annealing a 19mer single-stranded (ss) oligonucleotide (5′-GTCACCGTCG-CTACGACTG-3′) to an 18mer single-stranded oligonucleotide (5′-GAGTCGTAGGACGGTGAC-3′), generating a duplex with an extra central base.

Incision analysis of duplex DNA substrates by AP endonucleases was carried out in 10 µl reactions. Ape and ExoIII reactions were performed in 50 mM HEPES–KOH (pH 7.5), 50 mM KCl, 100 µg/ml BSA, 10 mM $MgCl_2$, 10% glycerol and 0.05% Triton X-100 (23,24) and EndoIV reactions were performed in 50 mM HEPES–KOH (pH 7.5), 50 mM KCl, 1 mM dithiothreitol and 10% glycerol (23). Incision reactions were carried out as described in the figure legends and 3 µl samples analyzed on a 16% polyacrylamide–8 M urea denaturing gel after addition of 4 µl formamide loading buffer and immediate heating at 75°C for 5 min (24). Visualization of labeled substrate on the gels was achieved using a Molecular Dynamics (Sunnyvale, CA) STORM 860 PhosphorImager and quantitative data analysis was performed using Molecular Dynamics ImageQuant v1.11 software.

Electrophoretic mobility shift assays (EMSA)

Protein was incubated with 5′-^{32}P-labeled duplex DNA substrates for 5 min at 0°C in 10 µl SBC buffer (50 mM HEPES–KOH, pH 7.5, 50 mM KCl and 10% glycerol) containing 4 mM EDTA and the binding reactions analyzed on an 8% non-denaturing gel as previously described (25), except that the electrophoresis buffer contained 4 mM EDTA and electrophoresis was performed at a reduced voltage of 8 V/cm. Band visualization and quantitative analysis was performed as described above.

RESULTS

Incision kinetics and binding of Ape to (N:T) and (F:T) substrates

To examine the role of the ring structure in recognition and to test the effectiveness of N as a global inhibitor of BER (32), duplex 25 bp DNA substrates containing a centrally located synthetic F or N abasic site analog (Fig. 1A) with a deoxythymidine residue opposite the lesion were used to measure Ape incision and binding activities (see Materials and Methods). The efficiency of incision 5′ of the abasic site analog was determined by monitoring conversion of the full-length 5′-^{32}P-end-labeled strand to the shorter, incised product on a 16% denaturing polyacrylamide gel. Time course experiments revealed that the initial reaction velocity of Ape at (F:T) was $5.2 \pm 2.3 \times 10^5$ U/mg, ~3-fold higher

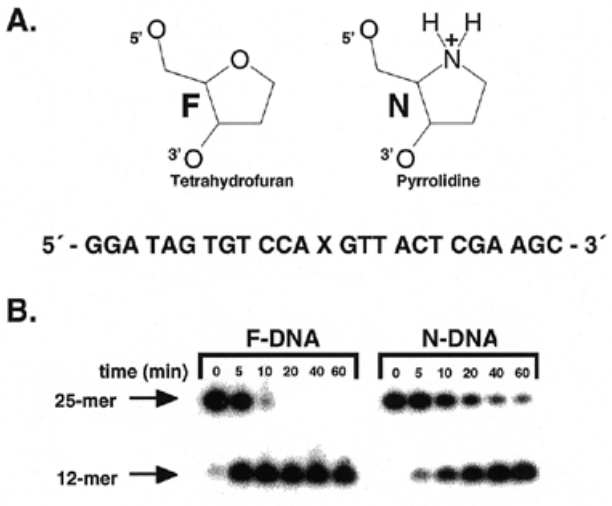

Figure 1. Ape incision of (N:T) and (F:T) substrates. (**A**) Structure of tetrahydrofuran (F) and pyrrolidine (N) abasic site analogs and sequence context of the 25mer oligonucleotide used in experiments (X, position of the abasic site analog; the complementary strand with T opposite the lesion is not shown). (**B**) Time-dependent incision at (F:T) and (N:T) by Ape. An aliquot of 0.5 pmol labeled (F:T) or (N:T) was incubated with 100 pg Ape (2.8 fmol) under standard reaction conditions with Mg^{2+} (See Materials and Methods) for 5, 10, 20, 40 or 60 min. Reactions were loaded onto 16% denaturing polyacrylamide gels and analyzed as described in Materials and Methods.

than (N:T) (Fig. 1B). The kinetic parameters of Ape for (F:T) and (N:T) duplex DNAs were determined by measuring the initial incision rates at various concentrations of target substrate. K_m and k_{cat} values were obtained from Eadie–Hofstee plots: for F-DNA, the K_m value was 27 ± 7 nM and the k_{cat} value 41 ± 19/min; for N-DNA, the K_m value was 115 ± 38 nM and the k_{cat} value 8 ± 1/min. Overall, the specificity constant (k_{cat}/K_m) was ~21-fold lower for (N:T) than (F:T).

To further characterize the interaction of Ape with (F:T) and (N:T), the ability of Ape to complex with these substrates in the presence of EDTA (under non-incising conditions) was measured by EMSA (See Materials and Methods). Several experiments using varying enzyme concentrations were performed and the results of protein–DNA complex formation plotted (Fig. 2). The concentration of Ape required for half-maximal substrate binding (~35%) was 32 nM for (F:T) and 125 nM for (N:T). We should note that precise determination of half-maximal binding is complicated by the fact that dissociation of the protein–DNA complex likely occurs during the electrophoresis process and that residual incision (~5%) occurs at high enzyme concentrations (>50 nM) (25). Ape did not form complexes with undamaged duplex DNA or ssF-DNA (Fig. 2B), although Ape does appear to interact with unmodified dsDNA, but not ssDNA, as determined by competition studies (25). The above results indicate that N-DNA is not likely to serve as a global inhibitor of BER (42). However, a strategically modified N-DNA substrate with a non-cleavable linkage 5′ of N may serve as an effective inhibitor of both DNA glycosylases and Ape *in vivo*.

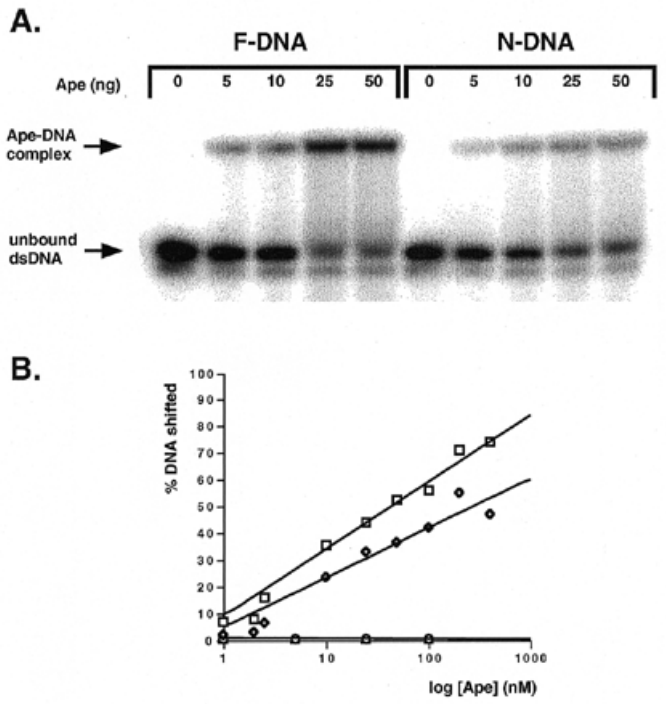

Figure 2. Ape complex formation with (F:T) and (N:T) substrates by EMSA. (**A**) An aliquot of 100 fmol labeled (F:T) or (N:T) was incubated with 1, 10, 25, 50 or 100 ng Ape (28 fmol–2.8 pmol) in EDTA and fractionated on a non-denaturing gel as described in Materials and Methods. Arrows indicate the position of Ape–DNA complexes and of unbound duplex DNA. The lower band corresponds to single-stranded and incised DNA (~10% of total label) as previously shown (25). (**B**) Graphic representation of all data obtained for Ape binding to (F:T) (8), (N:T) (◇), unmodified double stranded DNA (Δ) and single-stranded F-containing DNA (6).

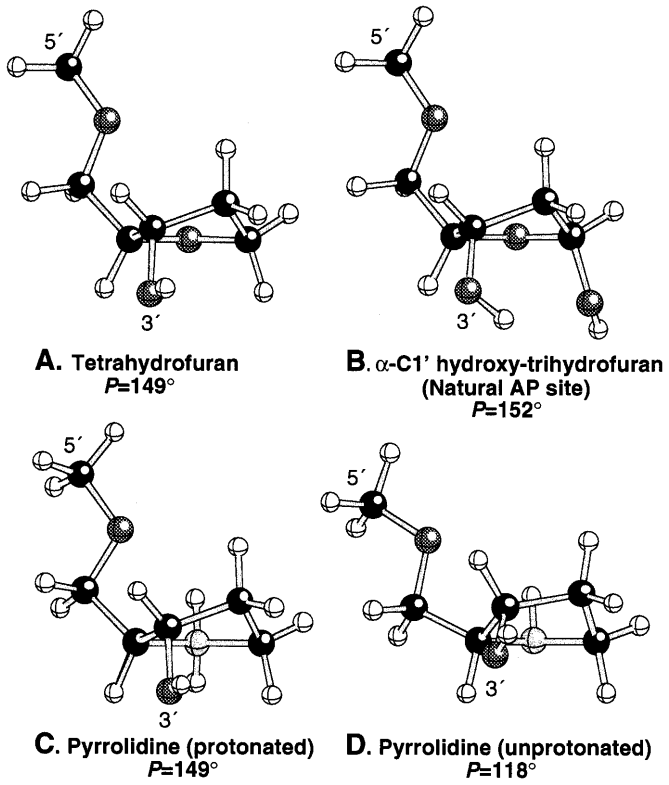

Figure 3. Hartree–Fock/6-31G* optimized structure of the lowest energy conformation of various abasic sites. *P* is the sugar pucker pseudorotation phase angle. (**A**) Tetrahydrofuran; (**B**) natural AP site; (**C**) protonated pyrrolidine; (**D**) unprotonated pyrrolidine.

Structures of abasic sites as determined by quantum chemistry

To gain further insights into structural features of AP sites that may influence recognition, the optimized HF/6-31G* lowest energy conformations for F, N and a natural AP site were determined (Fig. 3). Additional conformations including the 2′-*endo* sugar pucker and β-anomeric form of the C1′-hydroxyfuran were optimized, but were found to have higher energies (data not shown; all structures are available from the authors). Both furan structures and the protonated pyrrolidine are found to adopt a C2′-*endo* pucker (P = 149–152°), the usual value found in B-DNA helices. The unprotonated pyrrolidine ring has a somewhat different C1′-*exo* conformation (P = 118°). The calculated NPA charges on the endocyclic heteroatom in the sugar are –0.66 for the oxygen in F and –0.40 for the neutral amine of N (net charge including the hydrogen) to +0.29 if the ammonium moiety is protonated. Overall, the structures of the abasic sites, excluding the unprotonated N moiety, are very similar. It is worth noting that these findings support the biochemical evidence (22) indicating that F serves as an effective analog of a natural AP site.

Incision and binding of Ape F266A to (F:T) substrates

To test the proposed role of F266 in AP site recognition, enzymatic and binding activity of mutant Ape F266A protein was assessed. F266A protein displayed only an ~6-fold reduction in both specific activity ($0.96 \pm 0.14 \times 10^5$ U/mg) (Fig. 4A) and binding (Fig. 4B) relative to wild-type Ape protein. Targeted mutagenesis of W267 in Ape to alanine resulted in a similar 6-fold reduction in specific activity (data not shown). The possibility of bacterial AP endonuclease contamination was eliminated, since ExoIII displays different chromatographic properties than Ape on the BioLogic system (J.P.Erzberger and D.M.Wilson, unpublished data) and there was no significant exonuclease degradation of the DNA substrate (Fig. 4A). The EDTA-resistant (EndoIV-dependent) AP endonuclease activity of the F266A sample was <0.1% of the total activity.

Incision of (F:4meInd) paired duplex substrates and 'bulged' DNA substrates by Ape

The role of the opposite base in AP site recognition was tested by placing a non-polar base analog, 4meInd, opposite F. The properties of 4meInd were first evaluated by *ab initio* computer modeling (Fig. 5A). 4meInd, as determined by several different quantum chemical solvation models (Table 1), is less hydrophilic compared with the four natural bases, providing evidence that this base analog is more favored to remain stacked in the helix of duplex DNA. While these calculations do not take into account the solvent cavitation and entropy terms, the relative values for the four natural bases closely match the results obtained by free energy perturbation (43).

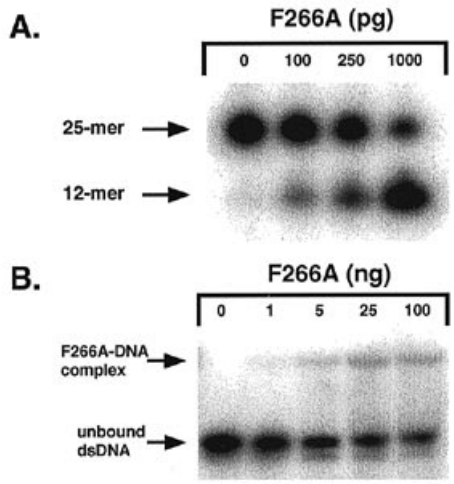

Figure 4. (F:T) incision and complex formation by F266A Ape mutant protein. (**A**) Enzymatic activity of F266A protein. Incision reactions containing 0.5 pmol labeled (F:T) and 100–1000 pg (2.8–28 fmol) F266A Ape mutant were carried out as described in Figure 1. (**B**) Binding activity of F266A protein. EMSAs containing 100 fmol (F:T) and 1, 5, 25 or 100 ng (28 fmol–2.8 pmol) F266A were performed as described in Figure 2. Arrows indicate the position of Ape–DNA complexes and of unbound duplex DNA.

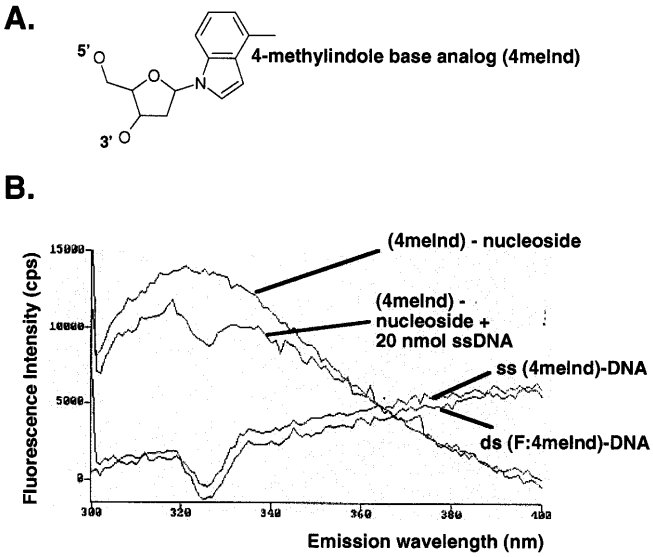

Figure 5. Structure and fluorescence of the non-polar base analog 4-methylindole (4meInd). (**A**) Chemical structure of 4meInd. The nucleotide sequence of the 4meInd oligonucleotide (GCTTCGAGTAAC[4meInd]TGGACACTATCC) used is complementary to the 25mer in Figure 1A. (**B**) Fluorescence analysis of (4meInd) nucleoside and (4meInd)-containing DNAs. Maximum emission by (4meInd) nucleoside occurred at ~320 nm. An aliquot of 20 nmol either HPLC purified (4meInd) nucleoside, (4meInd) nucleoside plus non-specific ssDNA, 4meInd-containing ssDNA or (F:4meInd) dsDNA were incubated at 37°C for 5 min in 2 ml Ape reaction buffer (see Materials and Methods) and immediately placed in a SPEX fluorometer for emission scanning. Background emission from the HEPES in the reaction buffer was subtracted from all scans. The small inverted peak observed in all the DNA-containing samples is due to the inner filter effect of DNA on the minor HEPES emission peak around 325 nm. The gradual increase in fluorescence is due to cooling of the sample during the 100 s of scanning.

To further examine the nature of the 4meInd-containing DNA, we measured the fluorescence of 4meInd nucleoside, 4meInd-containing ssDNA and (F:4meInd)-containing dsDNA. Studies have shown that fluorescent base analogs are quenched when stacked within the DNA helix, but not when present extrahelically (44). At 37°C, fluorescence peaks were observed for 4meInd nucleoside alone or for 4meInd nucleoside plus 20 nmol unmodified ssDNA, but no fluorescence was observed for ss or dsDNA containing a centrally located 4meInd residue, indicating that this base analog is stably stacked within the DNA helix (Fig. 5B). Likewise, no fluorescence was detected when (F:4meInd)-containing dsDNA was heated to 50°C (data not shown).

Table 1. Solvation profiles of 4meInd and the normal DNA bases

Base	DFT-SCIPCM	HF-SCIPCM	MP2-IPCM
4melnd	–3.0	–3.7	–4.4
A	–7.4	–8.5	–14.8
T	–8.1	–10.2	–13.4
C	–11.9	–14.7	–18.0
G	–15.7	–18.3	–23.4

Computer simulations were performed to calculate the electrostatic solvation energies (kcal/mol) using the polarizable continuum solvent model with density functional theory (DFT-SCIPCM), Hartree–Fock (HF-SCIPCM) and the second order Møller–Plesset perturbation theory (MP2-IPCM).

We subsequently measured the initial reaction velocity of Ape at duplex substrates containing three very different base moieties (4meInd, thymine and adenine) opposite F (Fig. 6A). The specific activities of Ape for these substrates were $5.2 \pm 2.3 \times 10^5$ U/mg for (F:T), $7.0 \pm 3.3 \times 10^5$ U/mg for (F:A) and $5.2 \pm 2.4 \times 10^5$ U/mg for (F:4meInd), indicating no significant preference for the moiety opposite (Fig. 6A). EMSAs demonstrated a <2-fold difference in the percentage of shifted complex when 0.25, 1, 5 or 25 ng Ape were added to (F:T), (F:A) or (F:4meInd) (data not shown). A kinetic run measuring Ape activity at five different concentrations of (F:4meInd) yielded a K_m value of 16 nM and a k_{cat} value of 21/min, both of which are similar to the values obtained for (F:T).

The structural similarities of 'bulged' DNAs (i.e. oligonucleotides with an extra centrally located base) with AP-DNA have raised the possibility that AP endonucleases may recognize 'bulged' substrates. However, no incision by Ape at 'bulged' DNA substrates was observed (Fig. 6B) and EMSAs showed no Ape–DNA complex formation with 'bulged' substrates possessing either a guanine or a cytosine as the extra base (data not shown).

Incision at various DNA chemical structures by *E.coli* AP endonucleases, ExoIII and EndoIV

The endonuclease activity of *E.coli* ExoIII, an Ape homolog, and *E.coli* EndoIV, a representative of a distinct family of AP endonucleases, was determined to reveal contrasts or trends between the different families of AP endonucleases. ExoIII incision activity for (N:T) was ~7-fold lower than for (F:T), a trend that is similar to that observed for Ape (Table 2). EndoIV, on the other hand, exhibited no significant difference in specific activity for either substrate (Table 2).

Both ExoIII and EndoIV were able to cleave (F:4meInd) with similar efficiencies to (F:T) and (F:A) (Table 2), suggesting that

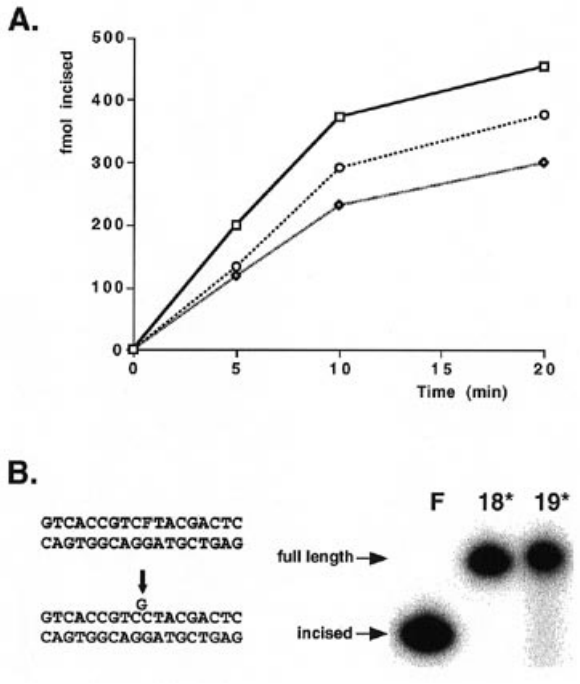

Figure 6. Ape incision at (F:4meInd) and 'bulged' DNA substrates. (**A**) Time-dependent incision of (F:A), (F:T) and (F:4meInd) substrates by Ape. Ape incision rates were determined by incubating 1 pmol either labeled (F:A) (8, solid line), (F:T) (◇, dotted line) or (F:4meInd) (6, dashed line) with 50 pg (1.4 fmol) Ape for 5, 10, 15 or 20 min. (**B**) An 18mer duplex DNA substrate containing a centrally located F abasic site analog (24) (top) and a 'bulged' DNA substrate generated by annealing an 18mer with a 'complementary' 19mer that possesses an extra, centrally located guanine (bottom, indicated by the arrow) were obtained. The F-containing strand of the 18mer duplex and either the 18mer or the 19mer of the 'bulged' substrate was ^{32}P-5′-end-labeled. These substrates were incubated with 1 ng Ape (28 fmol) for 5 min and the incision reactions analyzed (see Materials and Methods). As indicated by the arrows, complete incision is observed for F-DNA (lane 1), while no incision is detectable for either strand of the 'bulged' DNA (lanes 2 and 3).

the polarity of the base opposite F does not affect any of the endonucleases studied.

As previously published (31), ExoIII was unable to incise or bind 'bulged' DNA substrates. No incision of 'bulged' substrates was observed with EndoIV; however, in contrast to the damage-specific binding seen with ExoIII and Ape, EndoIV formed complexes with all of the DNA substrates analyzed [i.e. (U:G) mismatch, (F:T), (N:T) and 'bulged' DNA; data not shown].

Table 2. Summary of the incision activity of ExoIII and EndoIV

Substrate DNA	Specific activity (U/mg × 10^5)	
	ExoIII	EndoIV
N:T	0.07 ± 0.03	0.53 ± 0.32
F:T	0.47 ± 0.24	0.36 ± 0.25
F:A	1.07 ± 0.26	0.41 ± 0.12
F:4melnd	1.10 ± 0.24	0.24 ± 0.14

The specific activity (U/mg) on the various DNA duplex substrates was calculated by averaging the initial incision rates of at least five separate incision reactions at different DNA and protein concentrations. See Materials and Methods for unit definition.

DISCUSSION

Based on the crystal structure of Ape (30), it has been suggested that specific AP site recognition is mediated by the stacking of a phenylalanine residue at position 266 with the non-polar ribose ring of the abasic site. The fact that the mutant F266A Ape protein shows substantial incision and binding activity indicates that this residue is not essential for recognition or incision. Since a W267A mutation showed similar repair activity to F266A, it would appear that a global structural change in this region of the protein, rather than a specific amino acid change, is responsible for the lower activity of these mutant proteins. The observation that Ape complexes with (N:T) and (F:T) substrates with similar affinity in EMSAs suggests that specific recognition of the ring structure is also unlikely, since an interaction with a neutral ribose or a cationic pyrrolidine would likely proceed very differently. Moreover, previous studies demonstrating that Ape, ExoIII and EndoIV efficiently incise the abasic site analog propanediol, which lacks the ring structure, clearly indicate that the DNA sugar ring is not essential for targeted recognition by AP endonucleases (23,24).

The finding that Ape displays only a slight reduction (3- to 4-fold) in its ability to complex with (N:T) relative to (F:T) in EDTA (i.e. under non-incising conditions) implies that the more dramatic reduction (~21-fold) in catalytic efficiency for (N:T) incision stems from the pyrrolidine ring interfering with the cleavage event. Given that the distance from the N4′ of pyrrolidine to the target 5′-phosphate is only ~4.3 Å in a modeled AP site structure (D.Barsky, unpublished data), one possibility for the reduced incision efficiency of Ape and ExoIII is that the positively charged protonated endocyclic amine of N (+0.29 versus –0.66 for the F endocyclic oxygen) repels Mg^{2+}, which is necessary for the catalytic activity of these enzymes. While it is difficult to accurately measure the effective pH at the abasic site, molecular dynamics simulations (45; D.Barsky, unpublished data) indicate that the abasic site cavity is accessible to as many as three water molecules. Therefore, N is sufficiently solvated to be protonated, and thus positively charged, in the double-stranded helix. Binding of the N-containing DNA substrates by glycosylases is also consistent with the notion that the pyrrolidine residue is protonated (32). The fact that EndoIV (which does not require Mg^{2+} for incision) is unaffected by N substitution supports the idea of a metal ion repulsion effect. Furthermore, quantum chemical studies found that F and protonated N maintain similar conformations, suggesting that it is unlikely to be a structural component of the abasic site ring that leads to these incision differences. Alternatively, the positively charged sugar ring may interfere with proper positioning of the catalytic residues within the active site pockets only of Ape and ExoIII.

Many recent examples have highlighted the ability of enzymes to use extrahelical bases as a mechanism for substrate recognition (46). The identification of ExoIII and dCTP co-crystals led to the suggestion that this enzyme may use the base opposite the AP site as an anchoring point for binding (20). Although the incision activity of AP endonucleases is not significantly affected by the opposite base (23,24), it should be pointed out that T4 endonuclease V can bind its target damage equally well regardless of the base flipped into its binding site (47). We positioned a non-polar purine analog, 4meInd (which remains preferentially stacked in the DNA helix), opposite F in order to further analyze the role of the opposite base in AP site recognition. Since we observed no alteration in incision or binding by Ape, ExoIII or EndoIV of

(F:A), (F:T) or (F:4meInd), we propose that an extrahelical opposing base is not essential for the *initial* recognition of AP sites (20,46,48). Moreover, Ape has been shown to effectively incise substrates where an abasic site analog is placed directly opposite another abasic site analog (24), clearly demonstrating that a base opposite is unnecessary for recognition.

To explore the idea that specific DNA structures serve as the recognition element for AP endonucleases (49,50), we examined incision and complex formation of Ape, ExoIII and EndoIV with 'bulged' DNAs, which have been shown to possess many of the same structural features of AP-DNA by NMR (45,51–58). Previous studies found that ExoIII is unable to cleave 'bulged' substrates (32). In this study, Ape was also found not to incise or complex with 'bulged' DNAs. EndoIV, on the other hand, while also unable to incise 'bulged' DNA, formed complexes with all DNAs tested, but appears to have a binding preference for abasic site-containing DNA (23). This finding suggests that EndoIV employs a different mechanism than the ExoIII family for locating or complexing with abasic sites in DNA. Structural information on EndoIV is clearly necessary to fully interpret these findings and to determine the mechanism of recognition for this family of proteins.

NMR studies have found that AP-DNA exists in a variety of conformations that are largely dependent on the chemical structure of the abasic lesion and the stacking energetics of the flanking bases (i.e. purines are strongly retained in the helix and only a pyrimidine flanked by pyrimidines tends to be expelled from the helix due to the collapse of the AP site) (45,52,59,60). The fact that AP endonucleases recognize a variety of DNA structural conformations induced by different sequence contexts and AP site chemistries, but not 'bulged' substrates, provides compelling evidence that these repair enzymes are not in search of a specific structural form.

Another potential mechanism for recognition by AP endonucleases is specific identification of the abasic site cavity. Tryptophan (W212) insertion into the AP site gap has been proposed as a potential step in recognition by ExoIII (31). The evidence for such a mechanism is based on the fact that tryptophan is able to insert into abasic sites *in vitro* (61) and that K-W-K tripeptides are able to insert and cleave AP sites via β-elimination (62). However, the fact that a mutation in Ape at F266 (which corresponds to W212 in ExoIII, the main active site candidate for such a recognition mechanism) had only a minor effect on the incision or binding activity of Ape suggests that specific insertion into the AP site gap is unlikely an essential step in recognition.

Despite the many efforts to define the recognition mechanism of AP endonucleases, including crystallization of Ape and ExoIII and the use of substrate analogs and site-directed mutagenesis, it would appear that there is still much to learn about how these repair enzymes target AP sites in DNA. It is also apparent that the two families of AP endonucleases implement different methods to specifically recognize abasic DNA and that we have even less knowledge of how the EndoIV family members operate.

For the ExoIII family, we favor the idea that abasic DNA, as opposed to unmodified DNA, undergoes a unique conformational change *upon protein binding* that permits complex formation and, in the presence of Mg^{2+}, incision. This hypothesis rests largely on Cu-orthophenanthroline experiments, where it was observed that only protein-bound substrates exhibited increased hypersensitivity at the AP site (25).

ACKNOWLEDGEMENTS

We thank Joe Mazrimas for synthesizing and purifying the 4meInd-containing DNA substrate, Drs Monique Cosman and Michael Howard for helpful discussions and Drs Ian McConnell, James George, James Carney and Jim Felton for critical reading of the manuscript. This work was carried out under the auspices of the US Department of Energy by Lawrence Livermore National Laboratory under contract no. W-7405-ENG-48 and supported by LDRD (97-ERD-002) and NIH (CA79056) grants to DMWIII and a NIH (51330) grant to Dr Gregory L.Verdine.

REFERENCES

1 Friedberg,E.C.,Walker,G.C. and Siede,W. (1995) *DNA Repair and Mutagenesis*. ASM Press, Washington, DC.
2 Krokan,H.E., Standal,R. and Slupphaug,G. (1997) *Biochem. J.*, **325**, 1–16.
3 Lindahl,T. (1993) *Nature*, **362**, 709–715.
4 Loeb,L.A. and Preston,B.D. (1986) *Annu. Rev. Genet.*, **20**, 201–230.
5 Demple,B. and Harrison,L. (1994) *Annu. Rev. Biochem.*, **3**, 915–948.
6 Xanthoudakis,S., Smeyne,R.J., Wallace,J.D. and Curran,T. (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 8919–8923.
7 Demple,B., Herman,T. and Chen,D.S. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 11450–11454.
8 Robson,C.N., Hochhauser,D., Craig,R., Rack,K., Buckle,V.J. and Hickson,I.D. (1992) *Nucleic Acids Res.*, **20**, 4417–4421.
9 Seki,S., Hatsushika,M., Watanabe,S., Akiyama,K., Nagao,K. and Tsutsui,K. (1992) *Biochim. Biophys. Acta*, **1131**, 287–299.
10 Xanthoudakis,S., Miao,G., Wang,F., Pan,Y.C. and Curran,T. (1992) *EMBO J.*, **11**, 3323–3335.
11 Matsumoto,Y. and Kim,K. (1995) *Science*, **269**, 699–702.
12 Singhal,R.K., Prasad,R. and Wilson,S.H. (1995) *J. Biol. Chem.*, **270**, 949–957.
13 Matsumoto,Y. and Bogenhagen,D.F. (1994) *Mol. Cell. Biol.*, **14**, 6187–6197.
14 Frosina,G., Fortini,P., Rossi,O., Carrozzino,F., Raspaglio,G., Cox,L.S., Lane,D.P., Abbondandolo,A. and Dogliotti,E. (1996) *J. Biol. Chem.*, **271**, 9573–9578.
15 Kubota,Y., Nash,R.A., Klungland,A., Schar,P., Barnes,D.E. and Lindahl,T. (1996) *EMBO J.*, **15**, 6662–6670.
16 Klungland,A. and Lindahl,T. (1997) *EMBO J.*, **16**, 3341–3348.
17 Wilson,D.M.,III and Thompson,L.H. (1997) *Proc. Natl. Acad. Sci. USA*, **94**, 12754–12757.
18 Prasad,R., Singhal,R.K., Srivastava,D.K., Molina,J.T., Tomkinson,A.E. and Wilson,S.H. (1996) *J. Biol. Chem.*, **271**, 16000–16007.
19 Caldecott,K.W., McKeown,C.K., Tucker,J.D., Ljungquist,S. and Thompson,L.H. (1994) *Mol. Cell. Biol.*, **14**, 68–76.
20 Mol,C.D., Kuo,C.F., Thayer,M.M., Cunningham,R.P. and Tainer,J.A. (1995) *Nature*, **374**, 381–386.
21 Barzilay,G., Mol,C.D., Robson,C.N., Walker,L.J., Cunningham,R.P., Tainer,J.A. and Hickson,I.D. (1995) *Nature Struct. Biol.*, **2**, 561–568.
22 Takeshita,M., Chang,C.N., Johnson,F., Will,S. and Grollman,A.P. (1987) *J. Biol. Chem.*, **262**, 10171–10179.
23 Takeuchi,M., Lillis,R., Demple,B. and Takeshita,M. (1994) *J. Biol. Chem.*, **269**, 21907–21914.
24 Wilson,D.M.,III, Takeshita,M., Grollman,A.P. and Demple,B. (1995) *J. Biol. Chem.*, **270**, 16002–16007.
25 Wilson,D.M.,III, Takeshita,M. and Demple,B. (1997) *Nucleic Acids Res.*, **25**, 933–939.
26 Suh,D., Wilson,D.M.III and Povirk,L.F. (1997) *Nucleic Acids Res.*, **25**, 2495–2500.
27 Demple,B., Harrison,L., Wilson,D.M.,III, Bennett,R.A., Takagi,T. and Ascione,A.G. (1997) *Environ. Hlth Perspect.*, **105**, 931–934.
28 Haring,M., Rudiger,H., Demple,B., Boiteux,S. and Epe,B. (1994) *Nucleic Acids Res.*, **22**, 2010–2015.
29 Ide,H., Akamatsu,K., Kimura,Y., Michiue,K., Makino,K., Asaeda,A., Takamori,Y. and Kubo,K. (1993) *Biochemistry*, **32**, 8276–8283.

30 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) *EMBO J.*, **16**, 6548–6558.
31 Shida,T., Noda,M. and Sekiguchi,J. (1996) *Nucleic Acids Res.*, **24**, 4572–4576.
32 Schärer,O.D., Ortholand,J.-Y., Ganesan,A., Ezaz-Nikpay,K. and Verdine,G.L. (1995) *J. Am. Chem. Soc.*, **117**, 6623–6624.
33 Levin,J.D., Johnson,A.W. and Demple,B. (1988) *J. Biol. Chem.*, **263**, 8066–8071.
34 Ausubel,F.M. (ed.) (1997) *Current Protocols in Molecular Biology*. John Wiley & Sons, New York, NY.
35 Hehre,W.J., Radom,L., Schleyer,P.v.R. and Pople,J.A. (1986) *Ab Initio Molecular Orbital Theory*. John Wiley & Sons, New York, NY.
36 Saenger,W. (1984) *Principles of Nucleic Acid Structure*. Springer-Verlag, New York, NY.
37 Reed,A.E., Weinstock,R.B. and Weinhold,F. (1985) *J. Chem. Phys.*, **83**, 735–746.
38 Frisch,M.J., Trucks,G.W., Schlegel,H.B., Gill,P.M.W., Johnson,B.G., Robb,M.A., Cheeseman,J.R., Keith,T., Petersson,G.A., Montogomery,J.A. *et al.* (1995) *Gaussian 94, E.1 ed*. Gaussian Inc., Pittsburg.
39 Becke,A.D. (1993) *J. Chem. Phys.*, **98**, 5648–5652.
40 Miertus,S., Scrocco,E. and Tomasi,J. (1981) *Chem. Phys.*, **55**, 117–129.
41 Colvin,M.E., Seidl,E.T., Nielsen,I.M.B., Bui,L.L. and Hatch,F.T. (1997) *Chemico-Biol. Interact.*, **108**, 39–66
42 Schärer,O.D., Nash,H.M., Jiricny,J., Laval,J. and Verdine,G.L. (1998) *J. Biol. Chem.*, **273**, 8592–8597.
43 Miller,J.L. and Kollman,P.A. (1996) *J. Phys. Chem.*, **100**, 8587–8594.
44 McCullough,A.K., Dodson,M.L., Schürer,O.D. and Lloyd,R.S. (1997) *J. Biol. Chem.*, **272**, 27210–27217.
45 Coppel,Y., Berthet,N., Coulombeau,C., Coulombeau,C., Garcia,J. and Lhomme,J. (1997) *Biochemistry*, **36**, 4817–4830.
46 Roberts,R.J. (1995) *Cell*, **82**, 9–12.
47 McCullough,A.K., Schärer,O., Verdine,G.L. and Lloyd,R.S. (1996) *J. Biol. Chem.*, **271**, 32147–32152.
48 Vassylyev,D.G. and Morikawa,K. (1997) *Curr. Opin. Struct. Biol.*, **7**, 103–109.
49 Weiss,B. (1976) *J. Biol. Chem.*, **251**, 1896–1901.
50 Singer,B. and Hang,B. (1997) *Chem. Res. Toxicol.*, **10**, 713–732.
51 Woodson,S.A. and Crothers,D.M. (1988) *Biochemistry*, **27**, 3130–3141.
52 Withka,J.M., Wilde,J.A. and Bolton,P.H. (1991) *Biochemistry*, **30**, 9931–9940.
53 Patel,D.J., Kozlowski,S.A., Marky,L.A., Rice,J.A., Broka,C., Itakura,K. and Breslauer,K.J. (1982) *Biochemistry*, **21**, 445–451.
54 Kalnik,M.W., Norman,D.G., Swann,P.F. and Patel,D.J. (1989) *J. Biol. Chem.*, **264**, 3702–3712.
55 Kalnik,M.W., Norman,D.G., Zagorski,M.G., Swann,P.F. and Patel,D.J. (1989) *Biochemistry*, **28**, 294–303.
56 Kalnik,M.W., Norman,D.G., Li,B.F., Swann,P.F. and Patel,D.J. (1990) *J. Biol. Chem.*, **265**, 636–647.
57 van den Hoogen,Y.T., van Beuzekom,A.A., van den Elst,H., van der Marel,G.A., van Boom,J. H. and Altona,C. (1988) *Nucleic Acids Res.*, **16**, 2971–2986.
58 van den Hoogen,Y.T., van Beuzekom,A.A., de Vroom,E., van der Marel,G.A., van Boom,J.H. and Altona,C. (1988) *Nucleic Acids Res.*, **16**, 5013–5030.
59 Cuniasse,P., Fazakerley,G.V., Guschlbauer,W., Kaplan,B.E. and Sowers,L.C. (1990) *J. Mol. Biol.*, **213**, 303–314.
60 Goljer,I., Kumar,S. and Bolton,P.H. (1995) *J. Biol. Chem.*, **270**, 22980–22987.
61 Behmoaras,T., Toulme,J.-J. and Helene,C. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 926–930.
62 Behmoaras,T., Toulme,J.-J. and Helene,C. (1981) *Nature*, **292**, 858–859.