

The crystal structure of the human DNA repair endonuclease HAP1 suggests the recognition of extra-helical deoxyribose at DNA abasic sites

Michael A.Gorman, Solange Morera,
Dominic G.Rothwell¹, Eric de La Fortelle²,
Clifford D.Mol³, John A.Tainer³,
Ian D.Hickson¹ and Paul S.Freemont⁴

Protein Structure Laboratory, Imperial Cancer Research Fund, 44 Lincoln's Inn Fields, London WC2A 3PX, ¹Imperial Cancer Research Fund Laboratories, Institute of Molecular Medicine, John Radcliffe Hospital, Oxford OX3 9DS and ²MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK and ³Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA 92037, USA

⁴Corresponding author
e-mail: freemont@icrf.icnet.uk

The structure of the major human apurinic/aprimidinic endonuclease (HAP1) has been solved at 2.2 Å resolution. The enzyme consists of two symmetrically related domains of similar topology and has significant structural similarity to both bovine DNase I and its *Escherichia coli* homologue exonuclease III (EXOIII). A structural comparison of these enzymes reveals three loop regions specific to HAP1 and EXOIII. These loop regions apparently act in DNA abasic site (AP) recognition and cleavage since DNase I, which lacks these loops, correspondingly lacks AP site specificity. The HAP1 structure furthermore suggests a mechanism for AP site binding which involves the recognition of the deoxyribose moiety in an extra-helical conformation, rather than a 'flipped-out' base opposite the AP site.

Keywords: abasic sites/DNA repair/endonuclease/HAP1/Ref-1

Introduction

In order to maintain a high degree of genomic stability within cells, the genetic material consists of DNA as opposed to RNA. However, the small chemical modification to the ribose ring of RNA to form the more stable DNA (namely the reduction of the 2' hydroxyl group) has one deleterious consequence. Although DNA phosphodiester bonds are more stable than those in RNA, the *N*-glycosylic bond linking the deoxyribose to the base is far more labile in DNA (Lindahl, 1993). The hydrolytic breakage of the *N*-glycosylic bond to produce apurinic/aprimidinic (AP) sites is quantitatively the most significant structural abnormality to arise in cellular DNA and can be generated by at least three routes: (i) by spontaneous hydrolysis of the *N*-glycosylic bond, (ii) as an intermediate in DNA base excision repair and (iii) by the action of endogenous factors such as reactive oxygen species produced by normal cellular metabolism or exogenous damaging agents such as ionising radiation (Demple and

Harrison, 1994; for review see Barzilay and Hickson, 1995). Indeed, it is estimated that between 2000 and 10 000 purine bases alone are lost (and regenerated) in each human cell per day (Lindahl and Nyberg, 1972). Furthermore, the half life of AP sites at physiological pH is between 20 and 100 h (Lindahl, 1990). As AP sites are both cytotoxic and highly promutagenic due to a lack of coding information, and given their stability and frequency of formation, the repair of AP sites is essential to cell viability. Thus, for genome stability to be maintained, all organisms have developed repair mechanisms to eliminate AP sites.

AP sites are specifically corrected by the base excision repair (BER) pathway (Demple and Harrison, 1994; Barzilay and Hickson, 1995; Doetsch, 1995). In BER, AP sites are recognized by an AP endonuclease which cleaves the phosphodiester backbone 5' to the AP site leaving a 3'-hydroxyl nucleotide and deoxyribose 5'-phosphate as termini. To complete repair, the deoxyribose 5'-phosphate moiety is removed by a deoxyribose-phosphodiesterase, a DNA polymerase replaces the missing nucleotide and a DNA ligase joins the phosphodiester backbone (for review see Friedberg *et al.*, 1995). In humans, the major AP endonuclease is HAP1 (Robson and Hickson, 1991), also known as APE (Demple *et al.*, 1991) or Ref-1 (Xanthoudakis and Curran, 1992).

HAP1 is a 35 kDa monomeric protein that displays a wide variety of enzymatic activities relevant to DNA metabolism (Barzilay and Hickson, 1995). Besides AP endonuclease activity, HAP1 also acts as a phosphodiesterase, removing lesions such as phosphoglycolaldehyde from the 3' side of X-ray induced DNA strand breaks, as a 3' phosphatase, and has RNaseH activity (Chen *et al.*, 1991; Demple *et al.*, 1991; Walker *et al.*, 1993; Barzilay *et al.*, 1995b). However, the physiological relevance of these additional activities is unclear, as they are present at a specific activity that is between two and four orders of magnitude lower than that of the AP endonuclease activity (Demple and Harrison, 1994; Barzilay *et al.*, 1995b). Some authors have suggested that HAP1 possesses 3'-5' exonuclease activity, but this has not been confirmed in other studies (Demple and Harrison, 1994; reviewed in Barzilay and Hickson, 1995). HAP1 is homologous to a class of AP endonucleases found in many organisms, with sequence identities within their respective DNA repair domains ranging from 57% (Arp from *Arabidopsis thaliana*) to 27% (EXOIII from *Escherichia coli*) (Barzilay and Hickson, 1995). However HAP1 contains a 61 residue N-terminal domain absent from EXOIII (Barzilay and Hickson, 1995). The first insights into the catalytic mode of action of this class of AP endonucleases came from the three-dimensional structure of *E.coli* EXOIII (Mol *et al.*, 1995). EXOIII has a fold similar to that of DNase I, including the same active site location and conservation

of certain catalytic residues (Mol *et al.*, 1995), even though the overall primary sequence homology is less than 20%.

In addition to DNA repair activity, HAP1 appears to play a role in the regulation of gene expression (Xanthoudakis *et al.*, 1992). The cellular response to oxidative and other stresses is governed by a number of defence mechanisms, which act to repair DNA damage, detoxify harmful species or initiate a suicide programme to eliminate heavily damaged cells (Demple and Harrison, 1994; Barzilay and Hickson, 1995). These mechanisms are thought to be integrated into a co-ordinated sensing and signalling pathway. As part of this process, HAP1, but not the *E.coli* homologue EXOIII, has been shown to activate, *in vitro* via a reduction reaction, several important transcription factors including p53 (Rainwater *et al.*, 1995), c-Fos and c-Jun (Abate *et al.*, 1990; Xanthoudakis *et al.*, 1992; Walker *et al.*, 1993) and NF- κ B (Xanthoudakis *et al.*, 1992). These transcription factors only bind efficiently to their cognate DNA target sequences when in a 'reduced state', which can be achieved *in vitro* by incubation with either chemical reducing agents or with HAP1/Ref-1 (Abate *et al.*, 1990; Xanthoudakis *et al.*, 1992; Walker *et al.*, 1993). However, the chemical nature of this 'reduced state' and the mechanisms by which these particular transcription factors are activated by HAP1 *in vitro* are not known.

Further evidence of a role for HAP1 in cellular responses against cytotoxic stress has come from studies in which endogenous HAP1 levels were depleted following expression of HAP1 antisense-RNA (Walker *et al.*, 1994; Ono *et al.*, 1995). These HAP1 depleted cells are hypersensitive to DNA damaging agents that generate AP sites and growth under hypoxic conditions, suggesting a physiological role for the HAP1 'redox' activity in cell survival under conditions of limiting oxygen (Walker *et al.*, 1994). The HAP1 'redox' activity has been mapped to its N-terminal region (amino acids 36 and 62) by a series of deletion and mutation studies (Walker *et al.*, 1993; Xanthoudakis *et al.*, 1994). Yet Cys 65 seems to be essential for the reduction/activation process, since mutation to Ala eliminates 'redox' activity (Walker *et al.*, 1993).

To understand at the molecular level how AP sites in DNA are recognised and cleaved in human cells, and how the putative HAP1 'redox' activity is mediated, we have determined the X-ray crystal structure at 2.2 Å resolution of a truncated, but fully active form, of HAP1. The similarity in fold between HAP1 and DNase I, combined with the previously reported structures of DNase I bound to DNA (Lahm and Suck, 1991; Weston *et al.*, 1992), suggests a specific model for HAP1-DNA interactions. Based on this model and a structural comparison between HAP1, DNase I and the *E.coli* homologue EXOIII, we propose both a catalytic mechanism tested partly by site directed mutagenesis, and a new model for the recognition of AP sites by this class of AP endonucleases.

Results and discussion

Structure determination

Needle-like crystals of full-length recombinant HAP1 protein, unsuitable for structural studies, were initially obtained. However, a truncated HAP1 (HAP1³⁶⁻³¹⁸) missing the 35 N-terminal amino acids produced good quality

Table I. Crystal parameters

Crystal	Space group	Unit cell parameters			
		<i>a</i> (Å)	<i>b</i> (Å)	<i>c</i> (Å)	β°
Sm Native	C2	87.05	44.74	78.95	103.76
Sm Pt	C2	87.26	44.70	78.78	103.45
Native 'N1'	C2	89.44	45.31	78.59	103.53

Table II. Data collection and processing

	Data sets		
	Native 'N1'	Sm Native	SmPt
Resolution (Å)	3.0	2.5	2.2
Measurements	21466	33895	54729
Unique measurements	6282	10415	15089
% Complete	99.7	95.5	98.1
Mean <i>I</i> / <i>SigI</i>	6.6	16.1	12.7
<i>R</i> _{sym}	12.8	3.6	4.3
<i>R</i> _{sym(ano)}		7.7	7.7
<i>R</i> _{iso}		28.4	31.0
<i>R</i> _{cullis-iso} (acentric)		0.7	0.80
<i>R</i> _{cullis-ano}		0.41	0.46
<i>P</i> _{iso} (acentric)		1.64	1.14
<i>P</i> _{ano}		4.5	4.1
Resolution (Å)		Mean FOM	
41.41–2.2 (centric)		0.66	
41.41–2.2 (acentric)		0.41	

$$R_{\text{sym}} = \frac{\sum |I(h)_i - \langle I(h) \rangle|}{\sum \langle I(h) \rangle}$$

$$R_{\text{sym(ano)}} = \frac{\sum |I(h_+) - \langle I(h_+) \rangle|}{\sum \langle I(h_+) \rangle}$$

$$R_{\text{iso}} = \frac{\sum |F_{\text{PH}} - F_{\text{P}}|}{\sum F_{\text{P}}}$$

*R*_{cullis-iso}(acentric) = r.m.s. lack of closure/r.m.s isomorphous difference.

*R*_{cullis-ano} = r.m.s. lack of closure/r.m.s. anomalous difference.

*P*_{iso}(acentric) Phasing power = $\langle |F_{\text{H}}| \rangle / \text{r.m.s. lack of closure for acentric reflections}$.

Mean FOM, mean figure of merit.

crystals in the presence of calcium. We previously showed that truncated HAP1³⁶⁻³¹⁸ retains full 'redox' and endonuclease repair activity (Walker *et al.*, 1993). Initial molecular replacement studies using the EXOIII structure and a calcium native dataset produced one weak solution, which identified some β -strands but gave uninterpretable electron density maps for the rest of the structure (our unpublished observations). Most heavy atom derivatives were also not useful due to non-isomorphism. However, crystallisation of HAP1³⁶⁻³¹⁸ in the presence of samarium produced good quality crystals and data (Sm Native; Tables I and II), although they were non-isomorphous to our original calcium crystals (unpublished observations). A 2.2 Å resolution double heavy atom derivative data set was also obtained by derivitising the HAP1³⁶⁻³¹⁸ samarium co-crystals with platinum (SmPt; Table I and II). A 3.0 Å resolution second 'native' data set was obtained by replacing the samarium with calcium through soaking experiments (N1; Tables I and II). The HAP1³⁶⁻³¹⁸ structure was then solved by the multiple isomorphous replacement and

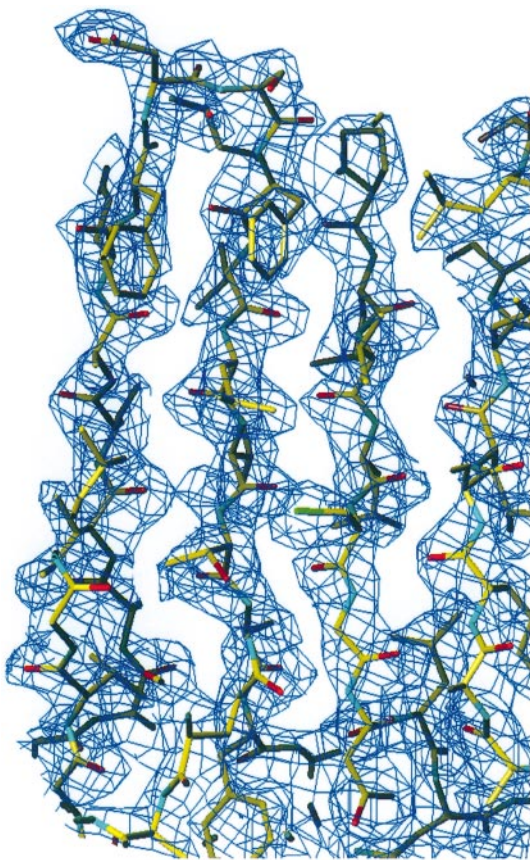


Fig. 1. Experimental electron density map and refined HAP1 atomic model. The map was calculated using the observed structure factors and phases derived from SHARP, after density modification (see Materials and methods). The map is contoured at 1σ , and was calculated using data between 20 and 2.2 Å.

anomalous scattering (MIRAS) method using the positions of the four samarium atoms and the single platinum atom (Table II). After density modification procedures, a good quality electron density map at 2.2 Å resolution was obtained (Figure 1). This electron density map allowed the interpretation of the majority of residues, although no electron density was visible for 12 N-terminal amino acids (including four vector-derived residues). The final model includes 157 water molecules, four samarium atoms and one platinum atom with all residues lying inside allowed Ramachandran regions (data not shown). Table III gives the final refinement statistics of the model as refined against the SmPt dataset, which was of higher quality and completeness compared with our other datasets (Table II).

Overall HAP1 structure and topology

With the exception of the N-terminal extension of 61 residues, which is lacking from EXOIII, HAP1 is a globular α/β protein consisting of two domains (domain 1: residues 44–136 and 295–318; domain 2: 137–260 and 282–294), with overall dimensions $40 \times 45 \times 40$ Å. Both domains display similar topologies with each comprising a six-stranded β -sheet surrounded by α -helices, which pack together to form a four-layered α/β -sandwich (Figure 2). Each β -sheet is composed of two or three anti-parallel β -strands with the β -strands in both domains flanked by topologically equivalent α -helices. β_1/β_7 are flanked by

Table III. Refinement statistics

Resolution range (Å)	20.0–2.2
R_{cryst} (%)	18.6
Protein atoms	2683
Solvent atoms	157
Metal atoms	5
Average B (Å ²)	15.40
Deviation from ideality (r.m.s.)	
Bond lengths (Å)	0.013
Bond angles (°)	1.780
Bond planarity (°)	1.622

$$R_{\text{cryst}} = \frac{\sum_h ||F_{\text{obs}}| - |F_{\text{calc}}||}{\sum_h |F_{\text{obs}}|}$$

calculated with Xplor on all reflections where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes of the reflection h .
r.m.s., root mean square.

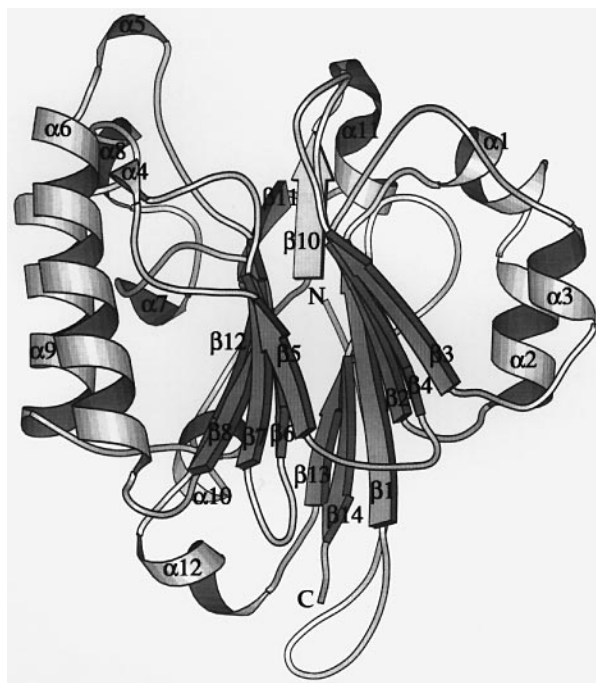


Fig. 2. Ribbon representation of the HAP1 structure. An α/β -sandwich is formed by the packing of two six-stranded β -sheets surrounded by α -helices, with strand order β_3 - β_4 - β_2 - β_1 - β_{13} and β_5 - β_6 - β_7 - β_8 - β_{12} - β_9 from domains 1 and 2, respectively. The α -helices and β -strands are numbered sequentially from the N-terminus (drawn using MOLSCRIPT; Kraulis, 1991).

kinked α -helices, α_1 - α_2/α_5 - α_6 , in domains 1 and 2 respectively, while β_3/β_8 are flanked by short α -helices α_3/α_4 , respectively. Helices α_7 , α_8 , α_9 and α_{10} have no comparable structure in domain 1, while α_{12} forms an inter-domain helix. Residues 261–281 form two extra-domain, anti-parallel β -strands, β_{10} and β_{11} , separated by a helical turn, α_{11} .

Comparison of HAP1 with EXOIII and DNase I

As predicted from the functional and limited sequence similarities between HAP1, EXOIII and DNase I, the three-dimensional structures of these proteins have similar folds (Figure 3). A superposition of 77 C α atoms of the central β -strands (excluding β_{10} and β_{11}) of HAP1 with

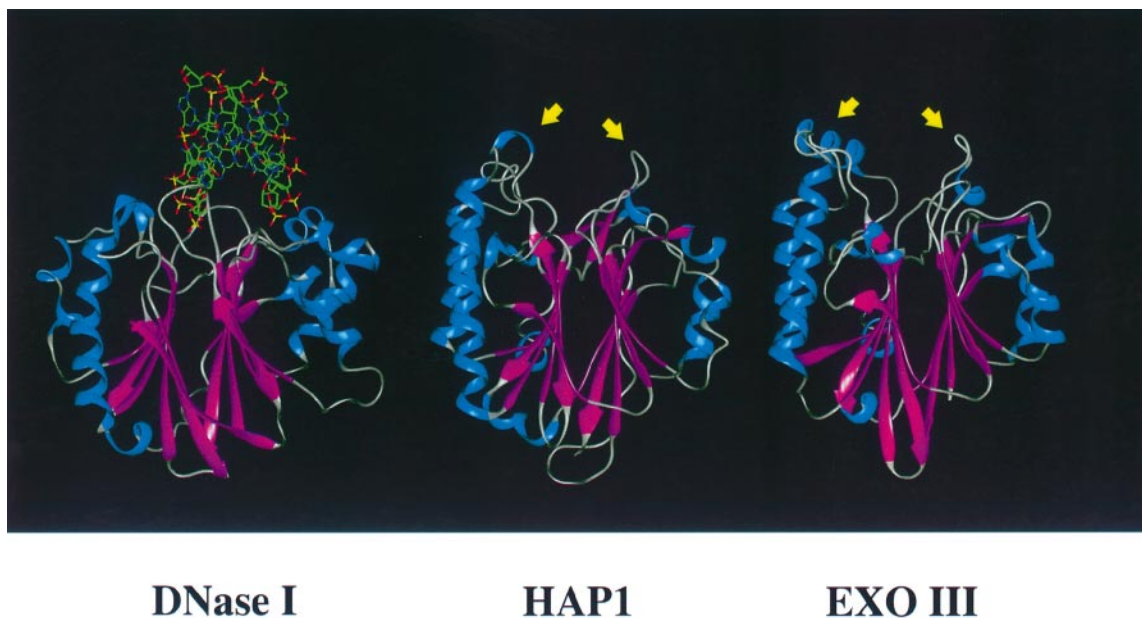


Fig. 3. Structural similarity between the overall folds of DNase I, HAP1 and EXOIII. For reference, DNase I is shown complexed to DNA (Lahm and Suck, 1991), and α -helices are coloured blue, with β -strands in magenta. The yellow arrows indicate the loop regions which are present in HAP1 and EXOIII, but are absent from DNase I (see text).

the EXOIII (PDB entry code 1AKO) and the DNase I core (PDB entry code 1DNK), gave root mean square deviations of 0.85 and 1.16 Å, respectively. Apart from the 61 residue N-terminal region, the only significant structural differences between HAP1 and EXOIII are confined to surface loop regions in which small residue insertions or deletions are evident. One interesting difference is the length of helix α_8 , which is longer by two turns in EXOIII compared with HAP1 (Figure 3), and has been implicated in DNA recognition and binding (Mol *et al.*, 1995, see below). A further comparison of HAP1 with DNase I shows the presence of three regions in HAP1 (and in EXOIII) which are absent from DNase I (Figure 3); the helical turn α_5 (176–181; residues 114–121 in EXOIII), α_8 (222–227; residues 164–173 in EXOIII) and the helical loop region α_{11} between strands β_{10} and β_{11} (267–277; residues 213–223 in EXOIII). Although these three regions are not identical between HAP1 and EXOIII, we suggest that these regions define in part the AP site recognition and cleavage specificity for AP endonucleases, since DNase I does not display specific AP site recognition.

A primary sequence alignment of HAP1, EXOIII and DNase I was obtained from the structural superposition whereupon other AP endonuclease sequences from different species could be aligned (Figure 4). The alignment shows several interesting features: (i) there is a significant degree of primary sequence conservation, (ii) the conserved residues can be grouped into categories depending upon whether they act to maintain tertiary structure, directly in catalysis or to stabilize and/or orient catalytic residues and proposed DNA binding loops and (iii) the AP endonuclease specific regions which are absent from DNase I are also missing in human L1 endonuclease (see later).

Of particular interest is the conservation in all AP endonucleases of residues that interact with the three

helical and/or loop regions which are absent from DNase I. These regions may, therefore, represent specific AP–DNA recognition elements. For example, the helical loop α_8 is stabilized and/or positioned by several conserved hydrophobic interactions in the AP family (Figure 4), which in HAP1 involve Leu220, Phe232 and Trp280. Interestingly, equivalent residues are not conserved in DNase I or L1 (Figure 4). A salt bridge between Arg237 and Glu217 is also involved in the orientation of α_8 , and is conserved in the AP endonuclease family (Figure 4). For the second helical loop region, α_5 , the hydrophobic stacking of Leu179 and Leu182 and a salt bridge between Arg181 and Glu154 stabilizes and/or positions the loop. This hydrophobic interaction appears conserved in all of the AP endonucleases, with the exception of EXOIII and L31, where an Arg and Phe form an equivalent interaction. The third helical loop region, α_{11} , is stabilized and/or positioned by a hydrogen bond between Trp267 and Asn226, as well as a hydrophobic interaction between Trp267 and Arg274. Interestingly, in EXOIII and L31, Trp267 is replaced with a Phe, as is Arg274, although Asn226 or equivalent is conserved (Figure 4), resulting in a different set of interactions, despite the conservation of loop conformation between HAP1 and EXOIII.

The sequence alignment of L1 endonuclease, DNase I and other AP endonucleases could explain some of the unique features of the L1 family. L1 endonuclease forms part of human L1 elements, which are highly abundant poly(A) (non-LTR) retrotransposons containing highly repetitive DNA sequences (Moran *et al.*, 1996; Boeke, 1997; Finnegan, 1997; Sassaman *et al.*, 1997). These elements are dispersed throughout the human genome and are implicated in a number of disease states. However the L1 endonuclease (L1 EN; Feng *et al.*, 1996) is different from HAP1, in that L1 EN has no preference for AP sites and preferentially cleaves at DNA sites which contain L1-like DNA sequences (Feng *et al.*, 1996). Although the

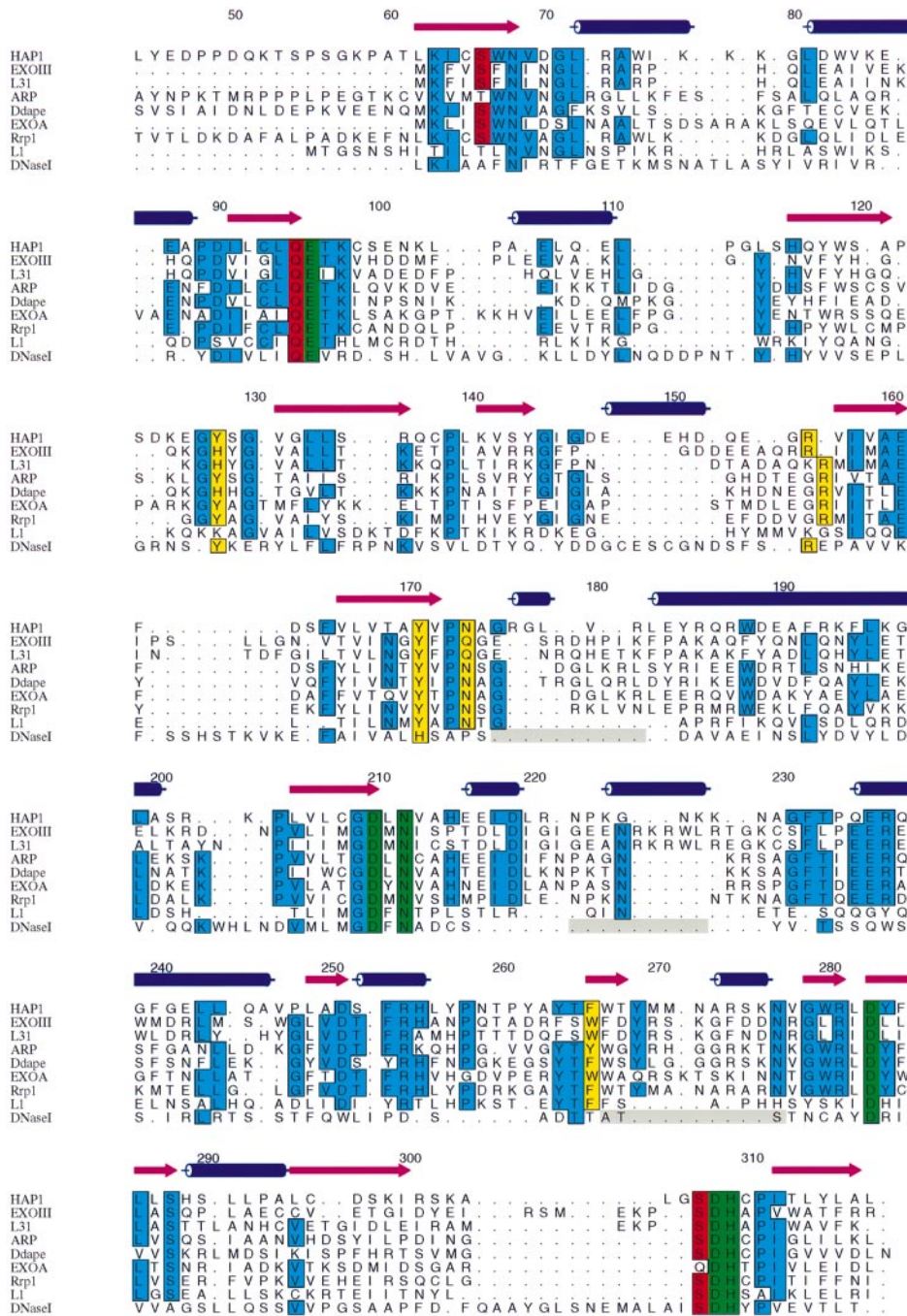


Fig. 4. Structure-based amino acid sequence alignment for AP endonucleases. An initial sequence alignment was obtained from the structural superposition of HAP1, EXOIII and DNase I. This was then used to align other members of the AP sequence family using the Clustal algorithm and displayed using ALSCRIPT (Barton, 1993). The numbering is for HAP1. All residues which are identical in at least five family members are boxed and coloured according to the following categories: residues in green correspond to catalytic residues or residues shown experimentally to be involved in enzymatic activity; residues in red are implicated in orienting and/or stabilizing catalytic residues; residues in blue are involved in tertiary structure interactions; residues in yellow are implicated in DNA–phosphate backbone interactions, based on the DNase I–DNA structure and HAP1–DNA model (see text). The three loop regions which are present in HAP1 and EXOIII, but are absent from DNase I, are shaded grey (see text). Secondary structure elements for HAP1 are shown as blue cylinders (α -helix) and magenta arrows (β -strands). The sequences are HAP1 (Human, 44–318; M99703; Robson and Hickson, 1991), EXOIII (*Escherichia coli*, 1–268; D90818) L31 (*Haemophilus ducreyi*, 1–267; U58147; Stevens et al., 1997), ARP (*Arabidopsis thaliana* 249–527; X76912; Babiychuk et al., 1994), Ddape (*Dictyostelium discoideum* 86–361; U31631; Freeland et al., 1996), EXOA (*Streptococcus pneumoniae* 1–275; J04234; Puyet et al., 1989), Rrp1 (*Drosophila melanogaster* 408–679; M62472; Sander et al., 1991), L1 (Human 1–246; U09116; Holmes et al., 1994) and DNase I (Bovine 1–260; M60606; Worrall and Connolly, 1990).

overall sequence of L1 EN is more similar to that of HAP1 (25% identity) and EXOIII (23% identity) than to DNase I (14% identity), L1 EN has no equivalent sequences to two (α 8 and α 11) of the three helical loop regions which are present in HAP1 and EXOIII, but are

absent in DNase I (Figure 4). Therefore, in these regions L1 would appear to be more structurally similar to DNase I than to HAP1 or EXOIII. This may explain why L1 has general DNA nicking activity like DNase I, and has no preference for AP sites unlike HAP1 and EXOIII.

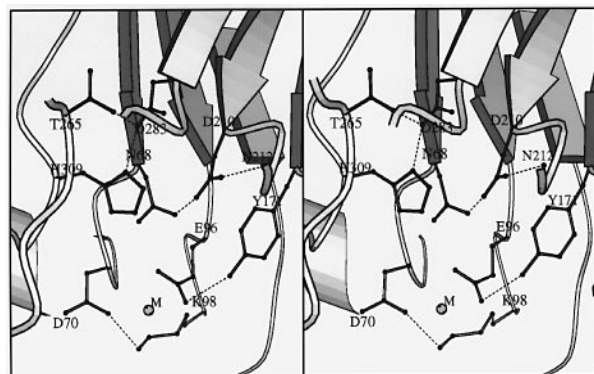


Fig. 5. Stereo view of the active site region of HAP1 showing the main residues implicated in catalysis. Hydrogen bonds are displayed as dotted lines. His309 and Asp283 are the proposed catalytic residues. A samarium metal ion is labelled M and binds to Glu96 (drawn using MOLSCRIPT; Kraulis, 1991).

Active site and proposed mechanism

Some of the active site residues of HAP1 have been determined either by site-directed mutagenesis studies, or analysis of sequence conservation between HAP1 and EXOIII (Barzilay *et al.*, 1995b; Rothwell and Hickson, 1996). The active site lies in a pocket at the top of the α/β -sandwich and is surrounded by loop regions (Figure 2). Within the active site, the imidazole ring of His309 interacts with the carboxylate of Asp283, which in turn forms a hydrogen bond with Thr265 (Figure 5). The side chains Tyr171 and Glu96 are hydrogen bonded, as are Asn68, Asp210 and the main chain amides of Asp70 and Asn212, which together form a hydrogen bonding network. In our HAP1 structure, a single samarium ion binds to the side chain of Glu96 (mean distance of 2.5 Å from the carboxyl group; Figure 5A). For both DNase I (Suck and Oefner, 1986) and EXOIII (Mol *et al.*, 1995), it was proposed that an Asp/His pair activates a water molecule which initiates an in-line nucleophilic attack on the phosphodiester bond. For HAP1, His309 would act as the general base to abstract a proton from a water molecule, while Asp283 would orient the imidazole ring and stabilize its transiently positive charged state (Figure 7D). The resulting hydroxide ion could then attack the scissile AP 5'-phosphate via an inversion of configuration. The catalytic importance of His309 and Asp283 is clear, since mutation of His309 to Asn and of Asp283 to Ala results in near elimination of enzymatic activity (Barzilay *et al.*, 1995b). The transition state intermediate could be stabilized by the divalent metal ion bound to Glu96, in a similar way to that proposed for DNase I (Suck and Oefner, 1986).

HAP1 absolutely requires divalent metal ions for catalytic activity, with a distinct preference for a magnesium ion (Barzilay *et al.*, 1995a). In our HAP1 structure, a samarium ion is bound similarly in the active site as the manganese ion seen in EXOIII (Mol *et al.*, 1995) which we consider to be representative of the catalytically essential divalent metal ion. The importance of Glu96 for metal binding is illustrated by the Glu96Ala mutant which exhibits 400-fold reduced catalytic activity and requires abnormally high amounts of exogenous divalent metal ions for activity (Barzilay *et al.*, 1995a). The carboxyl group of Asp308 is at a mean distance of 4.5 Å from the samarium ion and could therefore also participate either

indirectly or directly in metal binding, since an Asp308 Ala mutant results in both a reduced AP endonuclease activity and a preference for manganese over magnesium (Barzilay *et al.*, 1995b). Asp70 also lies close to the samarium ion (mean distance of 2.86 Å from the carboxyl group) and could participate in metal binding, but no evidence to support this is available at present. How the leaving group is stabilized is unclear, since, unlike DNase I, there is not a second His–Glu pair that can donate a proton. Possible explanations are that the metal ion remains bound to the leaving group, or that Asp210 acts as a proton donor. The observation that Asp210 lies close to a modelled AP phosphate supports this latter suggestion.

N-terminal REDOX domain

Three of the known AP endonucleases have extended N-terminal regions which are distinct from their DNA repair catalytic domains, including HAP1, which has an extra 61 N-terminal residues, Rrp1 (*Drosophila melanogaster*) with 427 extra N-terminal residues, and Arp (*A.thaliana*) with 270 extra N-terminal residues. In HAP1, this N-terminal region not only contains the nuclear localisation signal sequence (G.Barzilay and I.D.Hickson, unpublished observations), but also appears to regulate the DNA binding activity of many transcription factors *in vitro*. Specifically, the N-terminal region between residues 43 and 62 is necessary for HAP1 to act in a 'redox' mode, reductively activating oxidized proteins like p53 (Jayaraman *et al.*, 1997), c-Jun, c-Myb and c-Fos (Xanthoudakis and Curran, 1992; Xanthoudakis *et al.*, 1992, 1994; Walker *et al.*, 1993). However, the structural data indicate that this region forms an extended loop which lies across the β -strands β 13– β 14, making a number of hydrogen bond and salt bridge interactions with the globular core of the molecule. In particular, Asp50 makes a salt bridge with Arg301, and the main chain amide of Gln51 makes a hydrogen bond with the side chain carboxyl group of Asp297. Asp47 interacts with the main chain amide of Ile300 and the main chain carbonyls of Pro48 and Pro49 bind the side chain of Lys299. Notably, both Lys299 and Arg301 are generally not conserved in the AP family (Figure 4). This extended N-terminal structure is further stabilised by Pro55 and Pro59 at residue bends and by hydrophobic side chain interactions with the central domain, namely Tyr45 to Ala263 and Pro48 to Tyr257.

Cys65 is implicated in the 'redox' activity of HAP1 (Walker *et al.*, 1993). In the HAP1 structure, Cys65 is located on β 1, with the side chain pointing into a hydrophobic pocket (Trp67, Trp75, Trp83, Pro89, Leu92 and Pro311) and away from the central β -sheet (Figure 6). At one end of the pocket, Glu87 hydrogen bonds with Arg301 and Thr313, and Lys63 hydrogen bonds to Glu87 main chain carbonyl and Tyr315 side chain. Solvent accessibility calculations show that Cys65 is inaccessible to solvent and would, therefore, be unable to interact directly with residues from other proteins. Interestingly, Cys93 and Cys208 from domains 1 and 2 respectively, lie within the core β -sheet adjacent to Cys65, with their side chains only 3.5 Å apart. Nevertheless, there is no evidence for the existence of a disulphide bond between them.

The N-terminal 'redox' domain appears to pack onto the core of the molecule through a number of hydrogen bond and salt bridge interactions, forming part of the

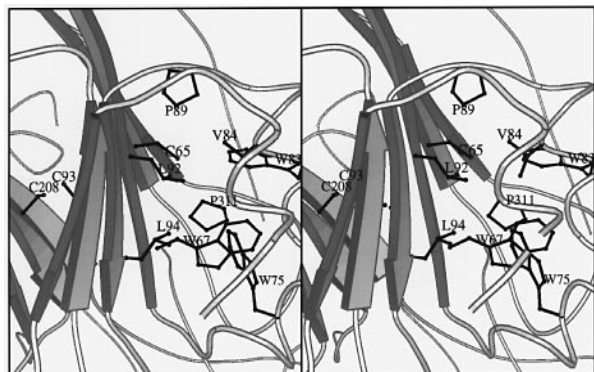


Fig. 6. Stereo view of the hydrophobic pocket which contains the essential 'redox' residue Cys65 in the HAP1 structure. Cys65 is inaccessible to solvent and is surrounded by a number of hydrophobic amino acids (see text). Cys93 and Cys208 are in close proximity to each other but are not disulfide-linked.

HAP1 molecular surface. Indeed, removal of more than 63 N-terminal residues could be deleterious to the overall HAP1 fold and structure. Consistent with this, we have shown that deletion mutants lacking 61 amino acids and more are poorly soluble (L.J.Walker and I.D.Hickson, unpublished data). The role of Cys65 in the 'redox' activity is perplexing, as it is buried in a hydrophobic pocket. However Cys65 mutations could, by disruption of a large hydrophobic core region, affect the stability and/or folding of HAP1 which then indirectly affects the apparent 'redox' activity, by altering the N-terminal domain conformation. It is also possible that the HAP1 structure is not representative of a 'redox' active conformation, since, in particular, HAP1 is required to be in a reduced state for activity, and no attempt was made to maintain such a state during crystal growth. Indeed, we observe an inter-molecular disulphide bond between crystallographically related HAP1 molecules involving Cys138, the relevance of which is unclear at present.

Proposed model for HAP1-DNA recognition and binding

The structure of DNase I has previously been determined bound to non-cleaved (Weston *et al.*, 1992) and cleaved (Lahm and Suck, 1991) DNA fragments. DNase I cleaves DNA in a generally non-sequence specific manner, but its activity is nevertheless influenced by the target DNA sequence. Consequently DNase I makes few specific protein-base interactions, but binds to DNA mainly via phosphate and non-specific hydrophobic interactions (Lahm and Suck, 1991). DNase I binds in the minor groove, and to both sides of the phosphodiester backbone, resulting in the bending of the DNA away from the enzyme (Weston *et al.*, 1992). Because AP sites occur anywhere within duplex DNA, it is likely that HAP1 DNA-binding is also primarily non sequence-specific. Since HAP1, EXOIII and DNase I all share significant structural similarity, it may be possible to derive a model of HAP1 bound to DNA based on the DNase I-DNA structures (Lahm and Suck, 1991; Weston *et al.*, 1992). From this model, there are a number of conserved amino acid residues between DNase I and AP endonucleases which could, by analogy, bind to the scissile phosphate and other phosphate groups (Figure 7A). For HAP1, these

include Tyr171, Asn174, Asn212 and His309. Interestingly, the conserved Arg156 and Tyr128 could make contact with the phosphate group 5' to the scissile phosphate in a similar way to that proposed for DNase I (Lahm and Suck, 1991). Tyr128 could also stack within the minor groove against the deoxyribose moiety two bases 5' to the AP site, an interaction synonymous with that of Tyr76 in DNase I (Figure 7A). These HAP1-DNA phosphate/base interactions are in good agreement with recent methylation and ethylation interference experiments, which showed that HAP1 makes contacts with two to three bases on either side of the AP site (Wilson *et al.*, 1997).

It has been reported that HAP1 requires at least 4 bp 5' to an abasic site and 3 bp 3' to an abasic site for cleavage at that site (Wilson *et al.*, 1995). Therefore, in order to study other potential interactions between HAP1 and DNA, an extended DNA duplex was incorporated into the HAP1-DNA model (Figure 7B). This model shows the helical loop α_8 (residues 222-229) positioned within the DNA major groove, with residues Asn226, Lys227, Lys228 and Asn229 making phosphate contacts similar to those proposed for EXOIII (Mol *et al.*, 1995). Notably, α_M in EXOIII is longer than the equivalent α_8 in HAP1, which may relate to expected subtle differences in DNA-binding and recognition between these two enzymes, since their substrate specificities differ significantly. In this HAP1-DNA model, a second helical loop region α_5 (176-181) is also positioned within the major groove (Figure 7B), with a number of polar residues again making potential DNA-phosphate backbone interactions. A third region, the extra-domain helical loop α_{11} (267-277) between strands β_{10} and β_{11} , is oriented near the DNA minor groove and could make additional phosphate backbone interactions through conserved polar residues (see Figure 4). Interestingly, the aforementioned three regions are found only in HAP1, EXOIII and, based upon sequence alignments (Figure 4), in the other AP endonucleases. They are not present in DNase I, suggesting an involvement in AP site recognition. Three other loop regions in HAP1 (residues 70-74, 98-102 and 125-129) show some similarity to DNA-binding loops in DNase I (residues 9-14, 41-44 and 72-76, respectively) and could be involved in making similar DNA interactions through phosphate backbone contacts (Figure 7B). Interestingly, Lys98 is equivalent to Arg41 in DNase I, a residue which makes base-specific contacts in the minor groove. However, in HAP1, Lys98 hydrogen bonds to the carboxyl group of Asp70, a residue possibly implicated in binding the catalytic divalent metal ion. An electrostatic surface potential map of the HAP1 structure with modelled DNA is also shown in Figure 8. This representation illustrates the surface complementarity between HAP1 and bound DNA. In particular, the helical loops α_5 and α_{11} positioned in the major and minor grooves respectively, and the DNA phosphate backbone aligned along positively charged areas (Figure 8). The AP site sits at the centre of the negatively charged active site cavity (Figure 8). It is notable that the HAP1 surface has a significant region of negative electrostatic surface potential, the functional significance of which is unclear.

How AP endonucleases recognise and specifically bind to AP sites is unknown, although a 'flipped out' base opposite the AP site may be part of the recognition process

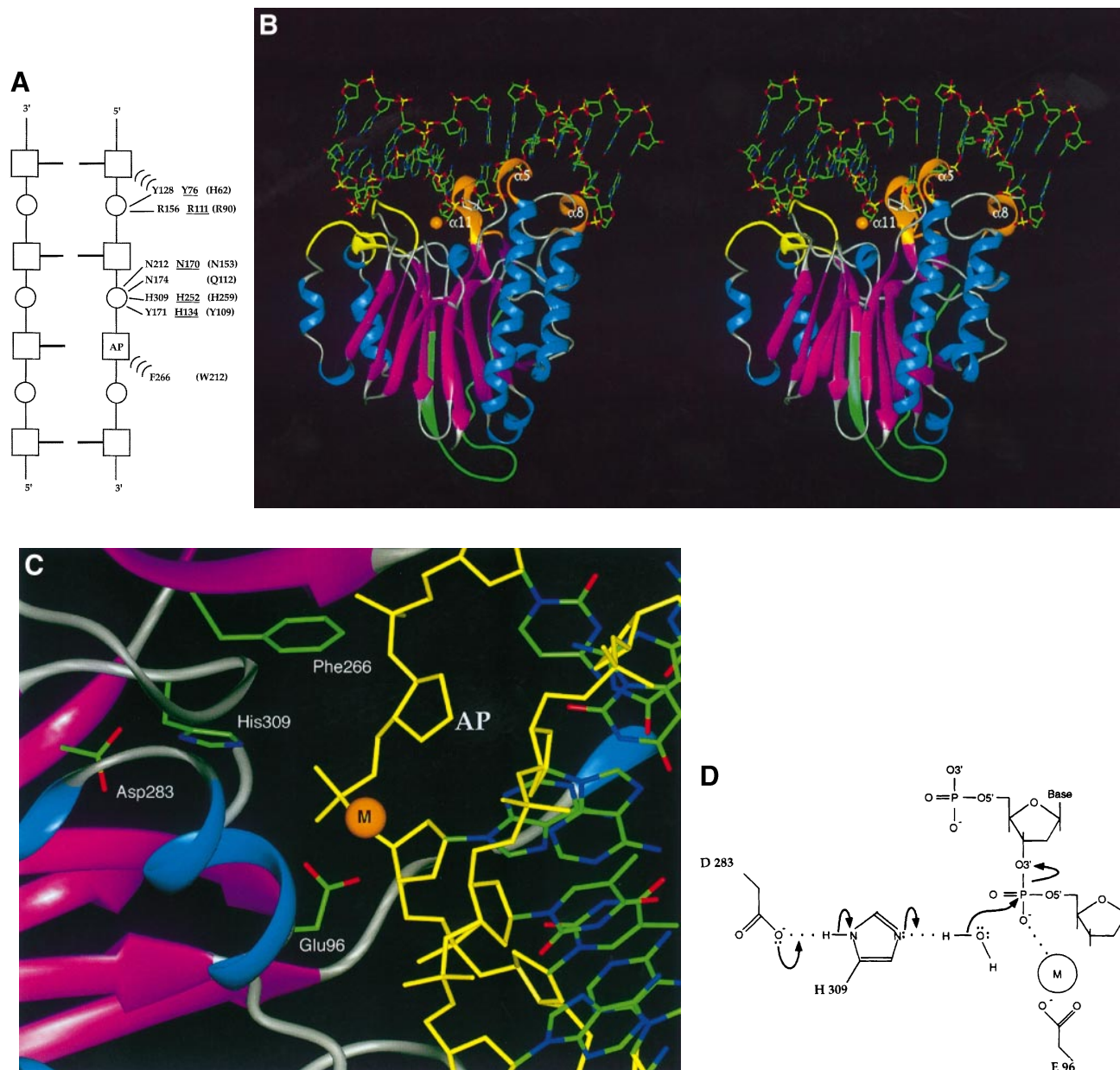


Fig. 7. Proposed model for HAP1 bound to DNA containing an AP site. **(A)** Specific amino acid–phosphate DNA interactions based on the structure of the DNase I–DNA complex (Lahm and Suck, 1991) and a structure-based sequence alignment between HAP1, DNase I and EXOIII (Figure 4). The DNA from two strands is shown represented as spheres (phosphate groups), squares (deoxyribose) and horizontal lines (bases). Amino acid–phosphate DNA interactions are shown as straight lines with deoxyribose stacking interactions as curved lines. DNA-interacting residues are shown, for HAP1 (left), DNase I (underlined) and EXOIII (in brackets). **(B)** Ribbon representation of the HAP1–DNA model in stereo. The model was obtained from superimposing HAP1 onto DNase I in complex with DNA (see text). An AP site was modelled (deoxyribose coloured white) at the observed phosphodiester cleavage site in the DNase I–DNA crystal structure. The divalent metal ion is shown as a gold sphere. The HAP1 ‘redox’ domain is coloured green. The three loop regions which are present in HAP1 and EXOIII, but not in DNase I, are coloured gold and are labelled. The loop regions which are similar between HAP1, DNase I and EXOIII are coloured yellow. **(C)** Close up view of the HAP1 active site in the HAP1–DNA model. The divalent metal ion is shown as a gold sphere labelled M. The DNA strand containing the AP site is coloured yellow. The essential catalytic residues of Glu96, Asp283 and His309 are shown. The deoxyribose of the AP site (labelled AP) is in close proximity to the conserved Phe 266 and could adopt an ‘extra-helical’ DNA conformation stabilized by the Phe side chain (see text). **(D)** Proposed catalytic mechanism for HAP1. His 309 acts as general base and abstracts a proton from a water molecule. The hydroxide ion thus formed then attacks the phosphate group 5’ to the AP site. A divalent metal ion (M), bound to Glu96, could be involved in stabilizing the transition state intermediate. Adapted from Suck and Oefner, 1986; Mol *et al.*, 1995.

(Mol *et al.*, 1995). However, considering the above HAP1–DNA model, HAP1 would probably make insufficient contacts with the DNA to allow both the recognition of the ‘flipped-out’ base and the cleavage of the phosphodiester bond at the AP site, since the AP site and the ‘flipped-out’ base would be on opposite sides of the DNA strand

(Figure 7B). Although it cannot be excluded that the DNA and/or the HAP1 protein undergo a significant conformational change allowing the specific recognition of a ‘flipped-out’ base, we consider it unlikely.

We propose an alternative model for the recognition of AP sites in which the base opposite the AP site is not

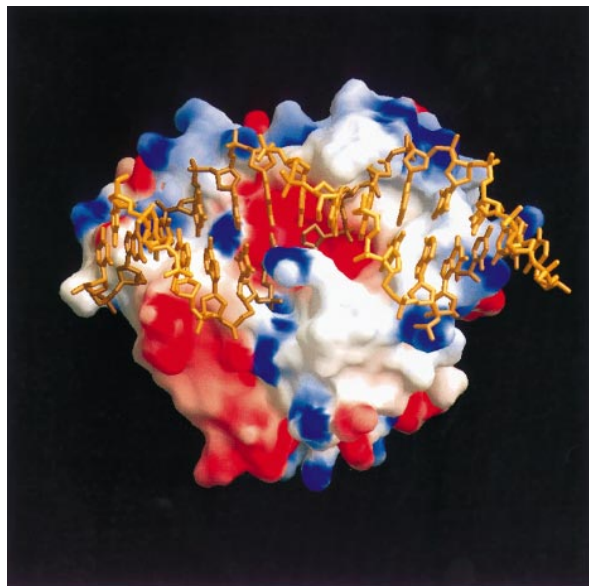


Fig. 8. Electrostatic surface potential map of HAP1 with modelled DNA. The surface was generated using the program GRASP (Nichols *et al.*, 1991) with red, blue and white corresponding to negative, positive and neutral charged regions, respectively. The modelled DNA is represented as a stick model coloured gold. The AP site lies directly in the active site cavity.

the primary recognition feature, but instead the enzyme specifically recognises the deoxyribose sugar moiety itself. In our HAP1-DNA model, the AP deoxyribose sugar could rotate around the 3'- and 5'-phosphodiester bonds and stack against Phe266 forming an 'extra-helical' deoxyribose moiety (Figure 7C). Interestingly, an aromatic amino acid at this position is conserved in all AP endonucleases but is a threonine in DNase I (Figure 4). Furthermore, the deoxyribose moiety of dCMP, as observed bound to EXOIII (Mol *et al.*, 1995), stacks against the equivalent residue Trp212. The orientation of dCMP bound to EXOIII (Mol *et al.*, 1995) may thus represent an 'extra-helical' deoxyribose. Experimental support for this hypothesis comes from NMR structures of duplex DNA fragments containing synthetic AP sites (Goljer *et al.*, 1995). These structures show two features for our hypothesis: (i) the deoxyribose ring is in equilibrium between an 'intra' and 'extra' DNA-helical orientation, and (ii) the overall conformation of the abasic site containing DNA is not significantly distorted (Goljer *et al.*, 1995). Thus, the recognition and binding of HAP1 to DNA could primarily involve three helical loop regions, which are specific to AP endonucleases, and the stabilization of an extra-helical deoxyribose ring of an AP site by a conserved aromatic residue.

Functional implications of the HAP1 structure

The recognition and repair of AP sites is of fundamental importance to cell viability and thus the AP repair pathway is highly conserved between species. Since AP sites occur frequently and spontaneously (Lindahl, 1993), the process of recognition and repair must be highly efficient. The structure of HAP1 reported here provides new insights into the first step of this pathway, namely recognition of the AP site and cleavage of the DNA backbone 5' to the AP site. By comparing HAP1 with the *E.coli* homologue

EXOIII and the functionally similar DNase I, we have identified three loop regions absent in DNase I, which we propose are involved in AP site recognition and cleavage. Furthermore, we propose an alternative model for AP site recognition, which involves recognizing the deoxyribose moiety of the AP site in an 'extra' DNA-helical conformation, rather than a 'flipped out' base opposite the AP site. Experimental support for this proposal comes from NMR studies of synthetic AP sites, which show that the deoxyribose at an AP site is in equilibrium between 'extra' and 'intra' DNA-helical conformations (Goljer *et al.*, 1995). All AP endonucleases could specifically recognize the 'extra-helical' AP deoxyribose ring, which would then lead to the formation of a stable enzyme-DNA catalytic complex. A structure of such a complex will be required to confirm the validity of this model.

Materials and methods

Overexpression and purification of HAP1

The truncated HAP1 cDNA, HAP1³⁶⁻³¹⁸, was cloned into the expression vector pT₇₋₇ (Novagen), placing the coding region under T₇ polymerase control. The construct was transformed into the *E.coli* strain, BL21(DE3)-pLysS, and induction of HAP1³⁶⁻³¹⁸ expression was achieved by the addition of 1 mM IPTG for 2 h at 37°C. Cells were lysed by adding lysozyme to a final concentration of 0.2 mg/ml placed on ice for 1 h, followed by sonication. The lysate was cleared by centrifugation and ammonium sulphate was added at 4°C to a final concentration of 45%. The HAP1³⁶⁻³¹⁸-containing supernatant was separated from the precipitate by centrifugation and HAP1³⁶⁻³¹⁸ was then precipitated in 75% saturated ammonium sulfate. HAP1³⁶⁻³¹⁸ protein was purified from the endogenous *E.coli* proteins using a succession of chromatography columns, as previously described (Barzilay *et al.*, 1995a). Assays of AP endonuclease and 'redox' activity were performed as described by Barzilay *et al.* (1995a) and Walker *et al.* (1993), respectively.

Crystallization, derivatization and data collection

HAP1³⁶⁻³¹⁸ crystals were obtained (Sm Native; Table I) using the hanging drop vapour diffusion method. Typically, drops containing 3 ml of HAP1³⁶⁻³¹⁸ (10 mg/ml in 10 mM HEPES pH 7.4) and 3 ml of well solution were equilibrated against 1 ml of well solution containing 16–20% (w/v) PEG 8000, 100 mM MES pH 6.2, 5% (v/v) 1,4-Dioxane (Sigma) and 7.5–30 mM samarium acetate (Johnson Matthey). Crystals appeared after three days as triangular plates reaching a maximum size of 0.5×0.125×0.05 mm. Double derivative crystals (SmPt; Table I) were obtained by transferring the samarium co-crystals into the same well solution containing 5 mM K₂Pt(NO₂)₄. A second 'native' crystal (Native N1; Table I) was obtained by transferring the samarium co-crystals into well solutions containing successively lower concentrations of samarium acetate and increasingly higher concentrations of calcium acetate. The crystals were placed in a final well solution containing 200 mM calcium acetate and left overnight at 19°C. All diffraction data were collected using an R-axis II image plate on a Rigaku RU200 generator with a copper rotating anode and Yale type mirrors. Crystals were flash cooled to 110 K during data collection using an Oxford Cryostream system. Crystal orientation, cell parameters and spot intensities were evaluated using the programs DENZO and SCALEPACK (Otwinowski and Minor, 1996). Data processing and scaling were carried out using the CCP4 program suite (Collaborative Computational Project Number 4, 1994).

Phase calculation and model building

The HAP1³⁶⁻³¹⁸ structure was solved by the method of multiple isomorphous replacement (MIR) and anomalous scattering (AS) using three datasets: Native N1 and two derivatives, Sm Native and SmPt (Table II). The heavy atom derivative data were interpreted by calculating difference Patterson maps using the coefficients $F_{\text{samarium}} - F_{\text{N1}}$ and $F_{\text{samarium+platinum}} - F_{\text{samarium}}$. The positions of the four samarium and single platinum heavy atoms were confirmed by calculating anomalous Patterson maps and subsequent difference Fourier maps using phases calculated from single samarium sites. Heavy atom refinement and phase calculations were carried out using the program SHARP (Eric de la

Fortelle, Cambridge). Table II outlines the final parameters of the derivatives. Density modification with the program SOLOMON (Abrahams and Leslie, 1996) was used to improve the phases. Crystallographic refinement using the SmPt dataset was performed using the conjugate-gradient method as implemented in X-PLOR (Brünger, 1992). The initial *R*-factor dropped from 40% (R_{free} 40.3%; based on 10% of the reflections not included in the refinement) to 34% (R_{free} 39.3%) after 50 cycles of Powell minimisation between 8 and 2.8 Å. Electron density maps using the Fourier coefficients ($2F_{\text{obs}} - F_{\text{calc}}$), α_{calc} and ($F_{\text{obs}} - F_{\text{calc}}$), α_{calc} from SIGMAA (Read, 1986) were calculated and examined. Corrections to the model were made manually using the program TURBO (A.Roussel and C.Cambillau, Marseille, France). Subsequent rounds of refinement and model building using simulated annealing, replacing the overall B factor by restrained individual B factors, extending the resolution to 2.2 Å and using the bulk solvent correction in the final stage (from 20 Å to 2.2 Å resolution) reduced the *R*-factor to 18.6% (R_{free} 26.8%) with the inclusion of 157 water molecules. The final refinement statistics are shown in Table III. Solvent accessibility calculations were carried out using the program ASA (A.Lesk, Cambridge) with a probe size of 1.4 Å.

Acknowledgements

We would like to thank Dr C.N.Robson for the HAP1^{36–318} cDNA and Dr L.J.Walker for initial preparations of HAP1^{36–318} protein. We are particularly indebted to Drs M.Sanderson and D.Brown for help and generous access to data collection facilities in the early part of this project. C.M. and J.T. thank T.Curran, S.Xanthoudakis and M.J.Hickey for providing protein for early crystal growth experiments. We thank Dr Gerard Bricogne for helpful discussions about the use of SHARP, and Drs Suhail Islam and Mike Sternberg for their assistance with the graphics program, PREPI. S.M. is grateful for EEC training and fellowship support (grant no. ERB CHBG-CT94–0556). J.T. is supported by NIH grant GM46312 and C.M. by a Leukemia Society of America Fellowship. The coordinates of HAP1 are being submitted to the Protein Data Bank, Brookhaven.

References

- Abate, C., Patel, L., Rauscher, F.J.D. and Curran, T. (1990) Redox regulation of fos and jun DNA-binding activity *in vitro*. *Science*, **249**, 1157–1161.
- Abrahams, J.P. and Leslie, A.G.W. (1996) SOLOMON. *Acta Crystallogr.*, **D**, 30–42.
- Babiychuk, E., Kushnir, S., Van Montagu, M. and Inze, D. (1994) The *Arabidopsis thaliana* apurinic endonuclease Arp reduces human transcription factors Fos and Jun. *Proc. Natl Acad. Sci. USA*, **91**, 3299–3303.
- Barton, G.J. (1993) ALSRIPT. A tool to format multiple sequence alignments. *Protein Eng.*, **6**, 37–40.
- Barzilay, G. and Hickson, I.D. (1995) Structure and function of apurinic/aprimidinic endonucleases. *Bioessays*, **17**, 713–719.
- Barzilay, G., Mol, C.D., Robson, C.N., Walker, L.J., Cunningham, R.P., Tainer, J.A. and Hickson, I.D. (1995a) Identification of critical active-site residues in the multifunctional human DNA repair enzyme HAP1. *Nat. Struct. Biol.*, **2**, 561–568.
- Barzilay, G., Walker, L.J., Robson, C.N. and Hickson, I.D. (1995b) Site-directed mutagenesis of the human DNA repair enzyme HAP1: identification of residues important for AP endonuclease and RNase H activity. *Nucleic Acids Res.*, **23**, 1544–1550.
- Boeke, J.D. (1997) LINEs and Alus—the polyA connection. *Nature Genet.*, **16**, 6–7.
- Brünger, A. (1992) X-PLOR (version 3.1) *A System for X-ray Crystallography and NMR*. Yale University Press, New Haven, CT.
- Chen, D.S., Herman, T. and Demple, B. (1991) Two distinct human DNA diesterases that hydrolyze 3'-blocking deoxyribose fragments from oxidized DNA. *Nucleic Acids Res.*, **19**, 5907–5914.
- Collaborative computation project Number 4 (1994) The CCP4 Suite. Programs for crystallography. *Acta Crystallogr.*, **D**, 760–763.
- Demple, B. and Harrison, L. (1994) Repair of oxidative damage to DNA: enzymology and biology. *Annu. Rev. Biochem.*, **63**, 915–948.
- Demple, B., Herman, T. and Chen, D.S. (1991) Cloning and expression of APE, the cDNA encoding the major human apurinic endonuclease: definition of a family of DNA repair enzymes. *Proc. Natl Acad. Sci. USA*, **88**, 11450–11454.
- Doetsch, P.W. (1995) What's old is new: an alternative DNA excision repair pathway. *Trends Biochem. Sci.*, **20**, 384–386.
- Feng, Q., Moran, J.V., Kazazian, H.H., Jr. and Boeke, J.D. (1996) Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell*, **87**, 905–916.
- Finnegan, D.J. (1997) Transposable elements: how non-LTR retrotransposons do it. *Curr. Biol.*, **7**, 245–248.
- Freeland, T.M., Guyer, R.B., Ling, A.Z. and Dering, R.A. (1996) Apurinic/aprimidinic (AP) endonuclease from *Dictyostelium discoideum*: cloning, nucleotide sequence and induction by sublethal levels of DNA damaging agents. *Nucleic Acids Res.*, **24**, 1950–1953.
- Friedberg, E.C., Walker, G.C. and Siede, W. (1995) DNA Repair and Mutagenesis. In *DNA Repair and Mutagenesis*. ASM Press, Washington D.C., pp. 135–190.
- 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.
- Holmes, S.E., Dombroski, B.A., Krebs, C.M., Boehm, C.D. and Kazazian, H.H. (1994) A new retrotransposable human L1 element from the LRE2 locus on chromosome 1q produces a chimaeric insertion. *Nature Genet.*, **7**, 143–148.
- Jayaraman, L., Murthy, K.G., Zhu, C., Curran, T., Xanthoudakis, S. and Prives, C. (1997) Identification of redox/repair protein Ref-1 as a potent activator of p53. *Genes Dev.*, **11**, 558–570.
- Kraulis, P. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.*, **24**, 946–950.
- Lahm, A. and Suck, D. (1991) DNase I-induced DNA conformation. 2 Å structure of a DNase I-octamer complex. *J. Mol. Biol.*, **222**, 645–667.
- Lindahl, T. (1990) Repair of intrinsic DNA lesions. *Mut. Res.*, **238**, 305–311.
- Lindahl, T. (1993) Instability and decay of the primary structure of DNA. *Nature*, **362**, 709–715.
- Lindahl, T. and Nyberg, B. (1972) Rate of depurination of native deoxyribonucleic acid. *Biochemistry*, **11**, 3610–3618.
- 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.
- Moran, J.V., Holmes, S.E., Naas, T.P., DeBerardinis, R.J., Boeke, J.D. and Kazazian, H.H., Jr. (1996) High frequency retrotransposition in cultured mammalian cells. *Cell*, **87**, 917–927.
- Nicholls, A., Sharp, K.A. and Honig, B. (1991) Protein folding and association—insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins Struct. Funct. Genet.*, **11**, 281–296.
- Ono, Y., Matsumoto, K., Furuta, T., Ohmoto, T., Akiyama, K. and Seki, S. (1995) Relationship between expression of a major apurinic/aprimidinic endonuclease (APEX nuclease) and susceptibility to genotoxic agents in human glioma cell lines. *J. Neurooncol.*, **25**, 183–192.
- Otwinowski, Z. and Minor, W. (1996) Processing of x-ray diffraction data collected in oscillation mode. In Carter, C.W.J. and Sweet, R.M. (eds) *Processing of X-ray Diffraction Data Collected in Oscillation Mode*. Academic Press, vol. 276.
- Puyet, A., Greenberg, B. and Lacks, S.A. (1989) The *exoA* gene of *Streptococcus pneumoniae* and its product, a DNA exonuclease with apurinic endonuclease activity. *J. Bacteriol.*, **171**, 2278–2286.
- Rainwater, R., Parks, D., Anderson, M.E., Tegtmeier, P. and Mann, K. (1995) Role of cysteine residues in regulation of p53 function. *Mol. Cell Biol.*, **15**, 3892–3903.
- Read, R.J. (1986) Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr.*, **A**, 140–149.
- Robson, C.N. and Hickson, I.D. (1991) Isolation of cDNA clones encoding a human apurinic/aprimidinic endonuclease that corrects DNA repair and mutagenesis defects in *E.coli* xth (exonuclease III) mutants. *Nucleic Acids Res.*, **19**, 5519–5523.
- Rothwell, D.G. and Hickson, I.D. (1996) Asparagine 212 is essential for a basic site recognition by the human DNA repair endonuclease HAP1. *Nucleic Acids Res.*, **24**, 4217–4221.
- Sander, M., Lowenhaupt, K., Lane, W.S. and Rich, A. (1991) Cloning and characterization of Rrp1, the gene encoding *Drosophila* strand transferase: carboxy-terminal homology to DNA repair. *Nucleic Acids Res.*, **19**, 4523–4529.
- Sassaman, D.M., Dombroski, B.A., Moran, J.V., Kimberland, M.L., Naas, T.P., DeBerardinis, R.J., Gabriel, A., Swergold, G.D. and Kazazian, H.H. (1997) Many human L1 elements are capable of retrotransposition. *Nature Genet.*, **16**, 37–43.

- Stevens,M.K., Klesney-Tait,J., Lumbley,S., Walters,K.A., Joffe,A.M., Radolf,J.D. and Hansen,E.J. (1997) Identification of tandem genes involved in lipooligosaccharide expression by *Haemophilus ducreyi*. *Infect. Immun.*, **65**, 651–660.
- Suck,D. and Oefner,C. (1986) Structure of DNase I at 2.0 Å resolution suggests a mechanism for binding to and cutting DNA. *Nature*, **321**, 620–625.
- Walker,L.J., Robson,C.N., Black,E., Gillespie,D. and Hickson,I.D. (1993) Identification of residues in the human DNA repair enzyme HAP1 (Ref-1) that are essential for redox regulation of Jun DNA binding. *Mol. Cell Biol.*, **13**, 5370–5376.
- Walker,L.J., Craig,R.B., Harris,A.L. and Hickson,I.D. (1994) A role for the human DNA repair enzyme HAP1 in cellular protection against DNA damaging agents and hypoxic stress. *Nucleic Acids Res.*, **22**, 4884–4889.
- Weston,S.A., 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.
- Wilson,D.M.R., Takeshita,M., Grollman,A.P. and Demple,B. (1995) Incision activity of human apurinic endonuclease (Ape) at abasic site analogs in DNA. *J. Biol. Chem.*, **270**, 16002–16007.
- Wilson,D.M.R., Takeshita,M. and Demple,B. (1997) Abasic site binding by the human apurinic endonuclease, Ape, and determination of the DNA contact sites. *Nucleic Acids Res.*, **25**, 933–939.
- Worrall,A.F. and Connolly,B.A. (1990) 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.
- Xanthoudakis,S. and Curran,T. (1992) Identification and characterization of Ref-1, a nuclear protein that facilitates AP-1 DNA-binding activity. *EMBO J.*, **11**, 653–665.
- Xanthoudakis,S., Miao,G., Wang,F., Pan,Y.C. and Curran,T. (1992) Redox activation of Fos-Jun DNA binding activity is mediated by a DNA repair enzyme. *EMBO J.*, **11**, 3323–3335.
- Xanthoudakis,S., Miao,G.G. and Curran,T. (1994) The redox and DNA-repair activities of Ref-1 are encoded by nonoverlapping domains. *Proc. Natl Acad. Sci. USA*, **91**, 23–27.

Received on July 17, 1997; revised on August 26, 1997