Conserved structural chemistry for incision activity in structurally non-homologous apurinic/apyirmidince endonuclease APE1 and Endonuclease IV DNA repair enzymes

Susan E. Tsutakawa[1], David S. Shin[2], Clifford D. Mol[2], Tadahide Izumi[34], Andrew S. Arvai[2], Anil K. Mantha[45], Bartosz Szczesny[4], Ivaylo N. Ivanov[6], David J. Hosfield[2], Buddhadev Maiti[6,] Mike E. Pique[2], Kenneth A. Frankel[1], Kenichi Hitomi[127], Richard P. Cunningham[8], Sankar Mitra[4], John A. Tainer[12]

[1]Lawrence Berkeley National Laboratory, Berkeley, CA 94720
[2]The Scripps Research Institute, La Jolla, CA 92037
[3]University of Kentucky, Lexington, KY 40536
[4]University of Texas Medical Branch, Galveston, TX 77555
[5]Present Address: Centre for Biosciences, Central University of Punjab, Bathinda, Punjab, India
[6]Georgia State University Atlanta, Georgia 30302
[7]Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka 560-8531, Japan
[8]University of Albany, SUNY, Albany, NY 12222
*Running title: Geometry of APE1 and Nfo AP endonucleolytic catalysis

To whom correspondence should be address: John A. Tainer, Life Science Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 947210 USA, Tel: (510) 486-4158; Fax: (510) 486-6880; Email: JATainer@lbl.gov

Key words: DNA repair; Base Excision Repair; Hap1; Ref-1; Nfo, magnesium, zinc, divalent metal, EndoIV

Background: DNA Apurinic/apyrimidinic (AP) sites are toxic and mutagenic if unrepaired by AP endonucleases.
Results: Structural, mutational, and computational analyses of prototypic AP endonucleases APE1 and Nfo identify surprisingly similarities.
Conclusion: APE1 and Nfo reveal functional equivalences enlightening their catalytic reaction.
Significance: A conserved catalytic geometry is specific to AP site removal despite different enzyme structures and metal ions.

SUMMARY

Non-coding apurinic/apyrimidinic (AP) sites in DNA form spontaneously and as DNA base excision repair intermediates are the most common toxic and mutagenic *in vivo* DNA lesion. For repair, AP sites must be processed by 5′ AP endonucleases in initial stages of base repair. Human APE1 and bacterial Nfo represent the two conserved 5′ AP endonuclease families in the biosphere; they both recognize AP sites and incise the phosphodiester backbone 5′ to the lesion. Yet, they lack similar structures and metal ion requirements. Here, we determined and analyzed crystal structures of a 2.4 Å resolution APE1-DNA product complex with Mg^{2+} and a 0.92 Å Nfo with three metal ions. Structural and biochemical comparions of these two evolutionarily distinct enzymes, characterize key APE1 catalytic residues that are potentially functionally similar to Nfo active site components, as further tested and supported by computational analyses. We observe a Mg-water cluster in the APE1 active site, with only Glu96 forming the only direct protein coordination to the Mg^{2+}. Despite differences in structure and metal requirements of APE1 and Nfo, comparison of their active site structures surprisingly reveals strong geometric conservation of the catalytic reaction, with APE1 catalytic side chains positioned analogously to Nfo metal positions, suggesting surprising functional equivalence between Nfo metal ions and APE1 residues. The finding that APE1 residues are positioned to substitute for Nfo metal ions is supported by the impact of mutations on activity. Collective results enlighten activities of residues, metal ions, and active site features for abasic site endonucleases.

Apurinic/apyrimidinic (AP) sites are the most common DNA lesions *in vivo*, calculated to be generated at ~10,000 lesions/cell/day in humans (1-3). AP sites form spontaneously or as central DNA repair intermediates during Base Excision Repair (BER) (4). AP sites can block replication and cause mutations, so its repair is critical for genetic integrity (5,6). Repair of AP sites is carried out via the BER pathway, which creates AP sites by uracil and alkylated base-specific monofunctional DNA glycosylases that excise the base lesion to produce AP sites (7). For the glycosylases, structures were key to revealing that specificity is encoded in nucleotide flipping and binding, as shown by MutY and uracil-DNA glycosylase (8,9). The AP sites produced by excision of the damaged base, are the substrates of AP-endonucleases. For AP-endonucleases, there are primarily two distinct families that incise the phosphodiester backbone 5′ to the lesion: the bacterial endonuclease IV (Nfo) family and the exonuclease III (Xth) family that includes Xth in bacteria and AP endonuclease (APE1) in metazoan eukaryotes (10). These pivotal nucleases detect, recognize, and cleave the DNA phosphodiester backbone 5′ of AP sites to create a free 3′-OH end for repair synthesis by a DNA polymerase. The AP-endonucleases also act as 3′ > 5′ exonucleases and catalyze nucleotide incision repair of particular oxidized bases (11-16). These two prototypic families mediate the same activities, but their structural folds and metal-dependence are different. In initial studies of Nfo, it was differentiated from Xth by its resistance to EDTA (17,18). The Xth family has a two-layered β-sheet core flanked by helices, and is Mg^{2+}-dependent (19). In contrast, Nfo has a TIM β barrel core, surrounded by helices. It has three metal ions – either three Zn^{2+} or two Zn^{2+} and one Mn^{2+} (20). In general, understanding distinct metal ion binding and activities, even in microbial systems, has lagged far behind genome sequencing and an increased knowledge of structure-function relationships is fundamental to more accurate metal ion prediction for responses to stress and DNA damage (21-25). Nfo is unusual among endonucleases in that it uses Zn^{2+} ions and that it uses three metals in its catalytic mechanism (26,27). Three metal mechanisms have also been proposed for EcoRV, RNaseH and microbial FEN1 (28-30), but in Nfo, all three metals occupy the active site at the same time. Recent biochemical, structural, and molecular dynamics (MD) studies have defined the individual roles of each Zn^{2+}: Zn1, Zn2, and Zn3 (Fig. 1) (20,27,31). The initial AP site is flipped into the active site and is bound by Zn1 and Zn3. The attacking water is deprotonated by Glu261, the one side chain directly involved in the catalytic mechanism. The resulting hydroxide ion is electrostatically stabilized by Zn1 and Zn2. All three Zn^{2+} ions stabilize the pentacoordinated transition state. At the end, Zn3 moves to coordinate the OP' and help stabilize the developing negative charge of the leaving group. Interestingly, crystallographic and biochemical studies suggest that the Zn3 position may actually be occupied by Mn^{2+} in *E. coli (Eco)*, suggesting further study.

Human APE1 (also called HAP1 and REF1) consists of a core nuclease domain that is conserved with *Eco*Xth and a 61 residue N-terminal domain that is not conserved in bacterial proteins (32,33). Crystal structures have shown that APE1 binds to both major and minor grooves of the DNA and flips out the abasic deoxyribose phosphate (34-37). In the original report of an APE1:DNA complex, we proposed a mechanism where Asp210 activates the attacking water (35). The phosphate intermediate is stabilized by the Mg^{2+} ion and contacts with His309, Asn174, and Asn212. Protonation of the 3′ ribose oxygen leaving group is through water in the first hydration shell of the Mg^{2+} (35). Yet, several subsequent papers with studies of active site mutants have proposed alternative enzyme mechanisms (38-43). These experiments were done on different mutants, in different labs using different techniques, so no clear consensus has been reached. Even the number of metal sites participating in catalysis is in question (44), which has hampered progress.

To help resolve these questions, we solved, analyzed, and compared new crystal structures of APE1 and Nfo. A 2.4 Å resolution crystal of wild type (WT) human APE1 in a product complex with Mg^{2+} revealed a Mg^{2+}-water cluster in the active site. We determined an extremely high resolution (0.92 Å) crystallographic structure for *Thermotoga*

maritima (Tma)Nfo with bound metal ions. Prompted by a common substrate, we superimposed tertiary structures of APE1 and Nfo and surprisingly observed that the scissile bond can be superimposed despite differences in protein fold, metal type, and number of metals. In fact, APE1 active site residues overlay onto metal positions in Nfo. Importantly, the geometric restraints implied by the Nfo superposition are consistent with only one of multiple proposed mechanisms. Our combined structural, biochemical, and computational analyses thus helps resolve mechanistic questions and supports a unified excision geometry and mechanism for the two prototypic AP endonucleases despite their structural differences.

EXPERIMENTAL PROCEDURES

Expression and Purification of APE1 and Nfo proteins. EcoNfo, TmaNfo and APE1 were expressed in *E. coli* and purified as previously described (27,35,45-47).

APE1 Crystallization and Data Analysis. WT APE1 (12 mg/ml) was incubated with 11mer double stranded DNA at a molar ratio of 1:1.2 for 10 min. The DNA contained a central tetrahydrofuran on one strand as previously described (35) and was purchased from Midland Inc. (USA). The protein/DNA was mixed 1:1 with 50 mM MES pH 6.0, 200 mM $LiSO_4$, and 25% mPEG 2K. Crystallographic data was collected at ALS beamline 5.0.1. Diffraction was observed to 2.1 Å but resolution was truncated to 2.4 Å resolution due to high anisotropy and overlaps. Phases were determined by molecular replacement. Refinement was performed by CNS (48) and later by PHENIX (49). NCS restraints were combined with TLS during refinement in PHENIX. There were four molecules in the asymmetric unit. The A and C chains had the most well-defined electron density with both having B factor averages for all protein atoms of 42 $Å^2$. B and D chains had significantly worse density with the average B factors of 73 and 58 $Å^2$, respectively. The N-terminus (residues 1-40) was evidently flexibly as it lacked unambiguous electron density. Clear octahedral geometry for the Mg^{2+} and coordinated waters were observable in the electron density for A, C, and D. Three other Mg^{2+} sites based on coordination geometry and distances were assigned in density near A and C chains. However, coordination was to waters and to two bases in the DNA, and these other Mg^{2+} ions are unlikely to affect the catalytic mechanism. The PDB ID for APE1 product complex is 4IEM.

Nfo crystallization and data analysis. TmaNfo (10 mg/ml) was crystallized in hanging drops mixed 1:1 with 100 mM Tris-HCl pH 9.0, 4 mM DTT, 1.5% saturated $MgSO_4$, 16% ethylene glycol, 20% mPEG2K. Crystals were frozen with 25% ethylene glycol. Crystallographic data was collected at the SSRL beamline 9-1. Diffraction data were collected to 0.92 Å resolution. Refinement was consistent with a mixture of Zn^{2+} and Mn^{2+} at metals sites 2 and 3. Tests to identify Mn in the active site were not attempted at the time, as occupancy of Mn was unexpected. Alternate metals could not be assigned the same number and thus the metals have been renumbered consecutively. Phases were determined by molecular replacement using EcoNfo PDB code 1QTW as a search model with the program AMORE (50), then refined using an improved model and the program EPMR (51) with resolution limits of 8 to 4.5 Å. The solution had an Rfactor of 0.317 and CC of 0.725, as compared to the next solution with an Rfactor of 0.52 and CC of 0.241. Refinement was performed initially with CNS (48) and later with PHENIX (49). The PDB ID for the TmaNfo is 4HNO.

Single-Turnover Kinetics. Specific endonuclease activity of APE1 Wild type (WT) and various APE1 mutants was analyzed by single-turnover kinetics under condition of excess enzyme and limiting substrate, a 52 nt tetrahydrofuran-containing oligo duplex (position 30 nt) (52-54). The reaction mixture (100 µl) contained 10 nM duplex oligo containing THF and 100 nM APE1. The ^{32}P-labeled oligo was diluted with unlabeled oligo to maximize detection range. The reaction was performed at 10°C under final buffer conditions of 50 mM Tris-HCl pH 8.0, 50 mM KCl and 2 mM $MgCl_2$. The reaction was initiated by adding enzyme to the reaction mixture and 10µl aliquots were removed at 10 sec, 20 sec, 30 sec, 1 min, 2 min, 5 min, 10 min, 30 min and 60 min into equal volume of 90% formamide denaturating loading dye containing 50 mM EDTA. The samples were heat denatured at 94°C for 3 min. Product (30 nt) was separated from substrate (52 nt) by 20% acrylamide/7 M urea gel. The

radioactivity in these bands was quantitated in a PhosphorImager (Molecular Dynamics) using ImageQuant software.

MD of APE1 and mutants. Models for the reactant APE1:abasic DNA complex and select mutants (E96A, D210N, N212A, H309A His309N and Y171F) were based on the APE1:product structure reported here and set up for classical MD using the AMBER PARM99SB force field with modified nucleic acid parameters (BSC0) (55,56) and TIP3P solvent (57). Na^+ and Cl^- ions were used for charge neutralization and to achieve a salt concentration of 0.1 M. The protonation states for histidine residues were determined by the WHATIF server (http://swiff.cmbi.ru.nl/servers/html/index.html).

The catalytically important His309 residue was set as protonated. After minimization (10,000 steps) and equilibration (4 ns dynamics with gradual scaling of positional restraints) we carried out fully unconstrained production runs for wtAPE1 for 15 ns. The mutant simulation trajectories were of the same length and employed the same simulation protocol as the wtAPE1 except that the mutated residue was allowed to move freely and equilibrate before releasing restraints on the other active site residues. All simulations were performed in the isothermal isobaric ensemble (NPT) at 1 atm and 300 K with the program NAMD 2.8 (58). A integration time step was used under a multiple time stepping scheme. The bonded and short-range interactions were calculated every third step. A short-range cutoff of 10 Å was used for the short range non-bonded interactions with a switching function at 8.5 Å. The long-range electrostatic interactions were treated with a smooth particle mesh Ewald method (58). The r-RESPA multiple timestep method (59) was adopted with a 2 fs time step for bonded, 2 fs for short-range non-bonded interactions and 4 fs for long-range electrostatic interactions. We used 2 fs time step keeping all bonds between hydrogen and heavy atoms constrained. We visualized trajectories and computed average properties with VMD (60).

Structure figures were produced using PYMOL (Schrödinger) and VMD (60).

RESULTS

Crystal structure of human WT APE1 bound to DNA. We crystallized full-length protein with Mg^{2+} and dsDNA containing a tetrahydrofuran. The structure was determined to 2.4 Å with an R_{factor} and R_{free} of 0.25 and 0.19, respectively (Table 1). This product structure is the highest resolution APE1/DNA complex known and a significant improvement compared to the published 3.0 Å Mn^{2+}:product complex (35). As before, there is only one metal ion observed in the active site in the asymmetric unit. The Mg^{2+} ion was identified by its coordination geometry and its distance to coordinating atoms (Fig. 1)(61).

In the three well-defined active sites, only one residue Glu96 directly coordinates the Mg^{2+} ion. Mutation of Glu96 reduces endonucleolytic activity by 600-fold (39,62). The 3′ ribose oxygen and the phosphate from the DNA and waters complete the tetrahedral coordination. Three waters completed the coordination. Asn68, Asp70 and Asp308 coordinate the waters in the Mg-water cluster. Mutation of these residues reduces endonucleolytic activity: Asn68, 200-fold (36); Asp308 5-fold (63,64), and Asp70 25-fold (39). Asp70 is particularly intriguing since mutation enhances 3′ phosphodiesterase activity but reduces AP endonuclease activity (65). For both activities, maximum activity requires a higher Mg^{2+} concentration, consistent with a role of Asp70 in helping to coordinate the Mg^{2+} ion through the water. Other active site waters act in indirect DNA interactions; e.g. Lys98 has a water-mediated interaction with the nucleotide adjacent to the 3′ ribose oxygen. The position of the APE1 Mg^{2+} ion, between the 5′ phosphate and the 3′ hydroxyl, is on the other side of the phosphate compared to many one-metal nucleases (66).

Ultrahigh Resolution Crystal Structure of TmaNfo. To obtain a greater understanding of the multiple metal ion Nfo catalytic mechanism in comparison to the single metal ion seen in our APE1 structures, we crystallized a hyperthermophilic ortholog from *Tma*. The optimal growth of this thermophile is ~80°C (67). TmaNfo has 33% identity with the *Eco* enzyme, suggesting overall similarity. The protein structure was solved to 0.92 Å with an R_{factor} and R_{free} of 0.12 and 0.14, respectively (Fig. 1 and Table 1). During our study, a structure of TmaNfo with Zn^{2+} and/or Cd^{2+} at a lower resolution, 2.3 Å, was reported (68).

As in EcoNfo, we found three metals in the active site, with a root mean square deviation (RMSD) of 0.2 Å from EcoNfo positions (20,27).

Previously, Mn^{2+} had been found in the Zn3 ion position in EcoNfo (27). The coordination is also more octahedral in geometry, and Nfo with Mn^{2+} is more active than with only Zn^{2+}. In our 0.92 Å structure, we could not unambiguously assign whether the Metal 2 and 3 positions were Zn^{2+} and Mn^{2+}, as the metal-ligand distances and the geometry were not definitive (61). However during refinement, we did observe negative FoFc difference density when Zn^{2+} was in Metal 2 or 3 positions and positive Fo-Fc when Mn^{2+} was in those positions, suggesting either low occupancy of Zn^{2+} or a mixture of Mn^{2+} and Zn^{2+}. In the Cd^{2+} TmaNfo structure, it was those two positions that were occupied by Cd^{2+} (68). Refinement with both Mn^{2+} and Zn^{2+} resulted in the lowest R-factors and smallest difference between R-work and R-free, as compared to other refinement scenarios.

The catalytic Glu261 and residues coordinating the three Zn^{2+} ions were absolutely conserved with EcoNfo. There was one residue different in the active site pocket. Ala30 in EcoNfo is replaced with Gln30 in TmaNfo. TmaNfo Gln30 appeared to exclude two water molecules. We postulate that this substitution is important for TmaNfo to have optimal endonucleolytic catalysis at 80 °C. To test this hypothesis, we compared the sequences from thermophilic and mesophic Nfo orthologs (Fig. 2). Although residue 30 was either alanine or glutamine in mesophilic Nfo, only glutamine was found at this position in thermophilic Nfo.

Superimposition of APE1 and Nfo active site structures. The mechanism for Nfo has been comprehensively studied biochemically, structurally, and computationally by MD (27,69). To better understand APE1, we compared the APE1 and Nfo structures. Examination of the DNA binding grooves in terms the local accessibility and for fractal dimension supported the overall similarity of the active sites (70,71) (Fig. 3). A shallow groove near the active site sterically imposes specificity for DNA lacking a bulky base. Notably, there is a significant pocket next to the active site that is present in both Nfo and APE1. We postulated that this pocket may allow the Nucleotide Incision Repair (NIR) activity and the requisite space for binding a base albeit a non-canonical base, but simple modeling with an NIR substrate, an α anomeric adenosine, did not position the base in that pocket (72). Perhaps this pocket has convergently evolved for a yet unrecognized functionality.

Since Nfo and APE1 are structurally dissimilar, we superimposed the tetrahydofuran moiety of our 2.4 Å APE1/Mg^{2+}/product complex with that from the reported *Eco* structure of Nfo/Zn^{2+}/product complex (1QUM.pdb)(20). A superposition based on only the tetrahydrofuran moiety in the DNA oriented the respective scissile 5′ phosphates and 3′ ribose oxygen atoms to overlay on top of each other, supporting the hypothesis that the two enzymes have a similar mechanism (Fig. 4A). The RMSD of the tetrahydrofuran atoms was 0.28 Å. The phosphates and ribose oxygens were 0.1 and 0,9 Å apart, respectively.

To better view any catalytic similarity of the active sites, we superimposed the scissile phosphates and ribose oxygens (Fig. 4B and C). Since EcoNfo has three metal atoms with distinct catalytic roles (31), we were interested in which metal atom geometrically matched the Mg^{2+} atom in APE1. Superimposing the scissile phosphate and the 3′ ribose oxygens from the two structures, the Zn^{2+} atoms fall in three quadrants around the phosphate with the Mg^{2+} in the fourth quadrant (Fig. 4C). Surprisingly, no one Zn^{2+} ion overlaid the Mg^{2+} ion. Notably, one Zn^{2+} atom (Zn1) is ~1.3 Å from the imidizole moiety of His309 in APE1. Mutation of His309 reduces APE1 activity over 30,000-fold (73). Another Zn^{2+} atom (Zn2) is located in between the side chains of Asn212 and Asp210. Zn2 also localizes ~1 Å from a second Mg^{2+} site proposed for APE1 (44). The side chain of Asn212 (APE1) is positioned similarly to Asp261 that directly coordinates the attacking water in Nfo. Notably, the N212A APE1 mutant lacked detectable AP endonuclease activity (74).

The third Zn^{2+} atom (Zn3) in Nfo is positioned close to APE1 Tyr171. Zn3 is postulated to act in stabilizing the pentacoordinate intermediate. Mutation of APE1 Tyr171 reduces AP catalytic activity more than 25,000-fold (42). Notably since Zn3 is in a mirror position to the Mg^{2+} relative to the axis of the reaction, replacement of Zn3 with Mn^{2+} increases Nfo activity (27). No side chains from Nfo overlay in this metal position.

The one protein residue that contributes to

the catalytic reaction in Nfo, Glu261, superimposes to lie close to Asn212 in APE1. Two Nfo residues that coordinated zinc atoms, Glu145 and His109, were 0.8 and 1.5 Å from Asp210 and His309 positions in APE1, respectively, the latter raising the possibility of the moving metal ion site. However, the measured pKa of over 8 for His309 is inconsistent with the need for a non-protonated imidizole necessary for metal ion coordination (75). His109 in Nfo is not in a His-Asp pair, as is His309 in APE1, which has Asp308 on one side and Asp283 on the other.

The superposition of the DNA substrates and overlay with Nfo catalytic metals highlighted the significance of four APE1 residues, Asp210, Asn212, His309, and Tyr171. From this structural comparison, it appears that APE1 residues may substitute catalytically for Nfo metals suggesting mutational and computation testing was merited.

Single Turnover Kinetics on Active Site Mutants

To better compare the implicated catalytic residues that have been mutated and studied in separate studies with different methods, we did single turnover kinetic studies of single site mutations on the key catalytic residues Glu96, Tyr171, Asp210, Asn212, and His309 (Fig. 5). Interestingly, we found that mutation of Glu96, the only residue that directly coordinates the Mg in the product structure, had the least effect with activity down only by 15-fold.

In contrast, Y171F and H309N showed quite large decreases in activity of ~1200 and 2500-fold respectively. Yet, Y171 and H309 were not the most important catalytic residues, as their mutation was not as severe as those of Asp210 and Asn212. Mutants D210N and N212A were decreased 10,000 and 7,000-fold in activity, respectively. Notably, D210 and N212 are located similarly to the two Zn^{2+} atoms (ZN1 and ZN2) found to coordinate the catalytic water in Nfo; they are positioned to coordinate a water that could make a linear attack on the phosphodiester.

To test our model computationally and provide an informed basis as to why specific mutations are reducing catalytic activity, we did MD simulations of WT and mutant APE1: E96A, Y171P, H309A, N212A, D210N, and H309N. In WT, we observed the catalytic water tightly hydrogen bonded to both Asp210 and Asn212 with average H-bond distances of 2.65±0.15 Å and 2.83±0.10 Å, respectively (Fig. 6). Interactions with these two residues position the oxygen atom of the catalytic water within 3.51±0.22 Å of the scissile phosphate and orient the water for an inline attack (average O*-P-O' angle of 162.6±8.2°). Notably, the hydrogen bonding network (Fig. 6) between the catalytic water, Asp210, Asn212, His309 and the abasic site phosphate is persistent throughout the MD trajectory. Indicating a less important catalytic role, mutations E96A, Y171P and H309A did not affect the H-bonded triplet of Asp210, Asn212 and catalytic water and the nucleophilic attack geometry is maintained (respective average O*-P distances of 3.52±0.25Å; 3.51±0.27Å and 4.24±0.69Å). In the N212A mutant the catalytic water is displaced from the inline position and Asp210 instead accepts H-bonds from Asn68 and Ala212 (backbone). In the D210N mutant, the catalytic water does not move. However, the mutated Asn210 side chain cannot accept a proton to activate the water. The ability of both D210N and N212A mutants to disrupt the attacking water correlates with the experimentally observed rates.

Consistent with the new APE1 crystal structure, our MD simulation of WT reveals protonated His309 stably hydrogen bonded to the abasic site phosphate (with average H-bond distance of 2.84±0.13Å and an almost linear angle of 8.9±4.7°). Experimentally, mutating H309 has a large effect on the catalytic rate. We attribute this to electrostatic polarization and charge transfer from the positively charged H309, which helps stabilize the developing negative charge on the phosphate moiety in the transition state. The H309N and H309A mutants would be much less effective in this role. Both mutants in MD are shown to create a water-filled cavity due to the smaller size of the Ala/Asn side chain compared to histidine. Although H-bonding to the equatorial oxygen of the phosphate from nearby water molecules is still possible, neutral water molecules do not stabilize the transition state either electrostatically or through charge transfer.

DISCUSSION

The value of superimposing structurally homologous proteins to understand catalytic mechanisms is well appreciated. However, it is not as intuitive to superimpose proteins that have no

structural homology and that furthermore have clearly distinct metal-dependencies in their catalytic mechanism. Nevertheless, for APE1 and Nfo, there was surprising agreement in positions of the tetrahydrofuran and leaving 3′ ribose oxygen, as well as residues and metals important for the catalytic function. The case of the evolutionary convergent catalytic triad in serine proteases is analogous except the proteins have converged on similar side chain chemistry.

For their biological functions, both APE1 and Nfo must recognize an AP site, and both flip the backbone at the AP site into a small pocket as a means of eliminating normal nucleotides with bases (10). For other DNA damage binding proteins, such as the alkylated guanine binding protein ATL, the adenine-guanine mistpair glycosylase MutY, and the uracil DNA glycosylase as well as structure specific nucleases, such as the FEN1 superfamily and Mre11, specificity is provided by flipping of nucleotides and placing the target phosphate bond into the active site (8,9,76-79). APE1 and Nfo both insert loop(s) from the minor groove side (10). Flipping the AP site ~180° into the active site pockets places the scissile phosphate into a position that is amenable for catalysis. Both endonucleases use an activated water to attack the phosphodiester bond, with an electronegative phospho-intermediate whose charge needs to be alleviated for efficient catalysis. Both endonucleases need to protonate the leaving 3′ ribose oxygen. We suggest that the requirements needed to catalyze the endonucleolytic reaction define the placement of the active site structural chemistry. It is these strict geometric requirements that made the superposition informative.

Superimpositions of APE1 and Nfo with other endonucleases that do not nucleotide flip their substrates, such as type II restriction endonucleases, were not as informative. The bonds connecting the scissile atoms diverged in angle and metal ions did not superimpose. Thus, the successful superposition between APE1 and Nfo is unique for the two prototypic AP endonucleases and defines them in their own class. The use of a magnesium-water cluster in APE1, clearly visible in this higher resolution structure of APE1:product, is not unprecedented. There have been several cases of magnesium-water clusters reported, among them mismatch VSR endonuclease, *Serratia* and *Ppo1* homing endonucleases, *Vibrio vulnificus* periplasmic nuclease Vvn, endonuclease V that initiates deaminated adenine repair, and UVC nuclear excision repair endonuclease (80-84), Possibly significant to their substrate specificity, APE1, Vvn, and Serratia homing endonuclease have been shown to be active on both RNA and DNA (85-87).

Given that the reaction and damaged DNA substrate have defined the geometry, the conservation of location of an active site residue in APE1 relative to the metals in the better structurally studied Nfo provides insight into APE1 catalytic mechanism; it allows us to plausibly assign the function of individual residues in the mechanism. The resulting mechanism is consistent with our initial model (35), but not other proposed mechanisms (38-43). Asp210 and Asn212 are positioned similarly to the key catalytic metals in Nfo, and their mutation was the most severe, suggesting that they coordinate the attacking water (Fig. 6).

There have been several papers published that have suggested alternative mechanisms that would be inconsistent with the geometric restraints. Since D210H was 15-fold more active than D210A or D210N, it was suggested that Asp210 is donating a proton to the leaving group (38,39). Asp210 is not close to the 3′ ribose oxygen so could not play this role. Histidines can activate the attacking water in nuclease mechanisms, and D210H could still activate an attacking water (40,41). It has also been suggested that Tyr171 in a phenolate form attacks the scissile phosphate directly or generates the hydroxyl that will cleave the scissile phosphate (42). This mechanism would be similar to topoisomerases. However, the angle of attack would not be linear with the position of the 3′ ribose oxygen. In a different mechanistic model, His309 has been suggested to generate the attacking nucleophile (43), but the geometry is also not consistent with this model. It is more than 3.5 Å from where the attacking water would be positioned. Moreover evidence from NMR has suggested that His309 in the APE1:DNA complex is protonated (75,88) and thus incapable of serving as a general base. Recently, MD studies suggested that there may be

a second metal binding site B where the Mg^{2+} coordinates with Asp210 and Asn212 and that the one metal moves from site B to the experimentally observed metal site A during catalysis (44,89). Our crystallographic work with one metal in site A does not show any density in position B in our 2.4 Å electron density maps. As this absence may be due to crystallographic conditions that prevent the B position from being occupied and the finding that site A was occupied, the absence in the electron density does not preclude the possibility that it exists. Yet, the superimposition with Nfo showed that site B was between Zn2 and the probable position of the attacking water, based on the product complex. Site B is too close (1.4 Å) to the attacking water and its coordination with both carboxylate oxygen atoms of Asp210 would prevent Asp210 from role of directly activating the water, as suggested from our computational work (Fig. 6). The authors also suggested that the retention of Mg dependence in an E96Q mutant indicates that site A, coordinated by Glu96, is not important (89). Our structure of the Mg^{2+} water cluster does not suggest that an E96Q mutation would prevent Mg^{2+} from occupying site A. Thus, our combined structural, mutational, and computations results do not support the existence of a second metal site B.

Here the determination and analysis of higher resolution APE1 and Nfo structures revealed a conserved active site structural chemistry despite major differences in metal ions and structural elements. In combination with mutational and computational analyses, the structures uncover functional equivalences and support a unified mechanism for AP site excision from DNA.

REFERENCES

1. Lindahl, T. (1993) Instability and decay of the primary structure of DNA. Nature 362, 709-715
2. Lindahl, T., and Barnes, D. E. (2000) Repair of endogenous DNA damage. Cold Spring Harb Symp Quant Biol 65, 127-133
3. Ciccia, A., and Elledge, S. J. (2010) The DNA damage response: making it safe to play with knives. Mol Cell 40, 179-204
4. Huffman, J. L., Sundheim, O., and Tainer, J. A. (2005) DNA base damage recognition and removal: new twists and grooves. Mutat Res. 577, 55-76
5. Loeb, L. A. (1985) Apurinic sites as mutagenic intermediates. Cell 40, 483-484
6. Sobol, R. W., Kartalou, M., Almeida, K. H., Joyce, D. F., Engelward, B. P., Horton, J. K., Prasad, R., Samson, L. D., and Wilson, S. H. (2003) Base excision repair intermediates induce p53-independent cytotoxic and genotoxic responses. The Journal of biological chemistry 278, 39951-39959
7. Hitomi, K., Iwai, S., and Tainer, J. A. (2007) The intricate structural chemistry of base excision repair machinery: implications for DNA damage recognition, removal, and repair. DNA repair 6, 410-428
8. Guan, Y., Manuel, R. C., Arvai, A. S., Parikh, S. S., Mol, C. D., Miller, J. H., Lloyd, S., and Tainer, J. A. (1998) MutY catalytic core, mutant and bound adenine structures define specificity for DNA repair enzyme superfamily. Nat Struct Biol 5, 1058-1064
9. Slupphaug, G., Mol, C. D., Kavli, B., Arvai, A. S., Krokan, H. E., and Tainer, J. A. (1996) A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA. Nature 384, 87-92
10. Mol, C. D., Hosfield, D. J., and Tainer, J. A. (2000) Abasic site recognition by two apurinic/apyrimidinic endonuclease families in DNA base excision repair: the 3' ends justify the means. Mutat Res 460, 211-229
11. Chou, K. M., and Cheng, Y. C. (2002) An exonucleolytic activity of human apurinic/apyrimidinic endonuclease on 3' mispaired DNA. Nature 415, 655-659.
12. Ishchenko, A. A., Sanz, G., Privezentzev, C. V., Maksimenko, A. V., and Saparbaev, M. (2003)

Characterisation of new substrate specificities of Escherichia coli and Saccharomyces cerevisiae AP endonucleases. Nucleic acids research 31, 6344-6353
13. Ishchenko, A. A., Yang, X., Ramotar, D., and Saparbaev, M. (2005) The 3'->5' exonuclease of Apn1 provides an alternative pathway to repair 7,8-dihydro-8-oxodeoxyguanosine in Saccharomyces cerevisiae. Mol Cell Biol 25, 6380-6390
14. Ischenko, A. A., and Saparbaev, M. K. (2002) Alternative nucleotide incision repair pathway for oxidative DNA damage. Nature 415, 183-187
15. Kerins, S. M., Collins, R., and McCarthy, T. V. (2003) Characterization of an endonuclease IV 3'-5' exonuclease activity. The Journal of biological chemistry 278, 3048-3054
16. Demple, B., Johnson, A., and Fung, D. (1986) Exonuclease III and endonuclease IV remove 3' blocks from DNA synthesis primers in H2O2-damaged Escherichia coli. Proc Natl Acad Sci U S A 83, 7731-7735
17. Cunningham, R. P., Saporito, S. M., Spitzer, S. G., and Weiss, B. (1986) Endonuclease IV (nfo) mutant of Escherichia coli. J Bacteriol 168, 1120-1127
18. Levin, J. D., Johnson, A. W., and Demple, B. (1988) Homogeneous Escherichia coli endonuclease IV. Characterization of an enzyme that recognizes oxidative damage in DNA. The Journal of biological chemistry 263, 8066-8071
19. 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
20. Hosfield, D. J., Guan, Y., Haas, B. J., Cunningham, R. P., and Tainer, J. A. (1999) Structure of the DNA repair enzyme endonuclease IV and its DNA complex: double-nucleotide flipping at abasic sites and three-metal-ion catalysis. Cell 98, 397-408
21. Cvetkovic, A., Menon, A. L., Thorgersen, M. P., Scott, J. W., Poole, F. L., 2nd, Jenney, F. E., Jr., Lancaster, W. A., Praissman, J. L., Shanmukh, S., Vaccaro, B. J., Trauger, S. A., Kalisiak, E., Apon, J. V., Siuzdak, G., Yannone, S. M., Tainer, J. A., and Adams, M. W. (2010) Microbial metalloproteomes are largely uncharacterized. Nature 466, 779-782
22. Barondeau, D. P., Kassmann, C. J., Bruns, C. K., Tainer, J. A., and Getzoff, E. D. (2004) Nickel superoxide dismutase structure and mechanism. Biochemistry 43, 8038-8047
23. McMurray, C. T., and Tainer, J. A. (2003) Cancer, cadmium and genome integrity. Nat Genet. 34, 239-241
24. Hopfner, K. P., Craig, L., Moncalian, G., Zinkel, R. A., Usui, T., Owen, B. A., Karcher, A., Henderson, B., Bodmer, J. L., McMurray, C. T., Carney, J. P., Petrini, J. H., and Tainer, J. A. (2002) The Rad50 zinc-hook is a structure joining Mre11 complexes in DNA recombination and repair. Nature 418, 562-566
25. Castagnetto, J. M., Hennessy, S. W., Roberts, V. A., Getzoff, E. D., Tainer, J. A., and Pique, M. E. (2002) MDB: the Metalloprotein Database and Browser at The Scripps Research Institute. Nucleic acids research 30, 379-382
26. Dupureur, C. M. (2008) An Integrated Look at Metallonuclease Mechanism. Current Chemical Biology 2, 159-173
27. Garcin, E. D., Hosfield, D. J., Desai, S. A., Haas, B. J., Bjoras, M., Cunningham, R. P., and Tainer, J. A. (2008) DNA apurinic-apyrimidinic site binding and excision by endonuclease IV. Nature structural & molecular biology 15, 515-522
28. Horton, N. C., Newberry, K. J., and Perona, J. J. (1998) Metal ion-mediated substrate-assisted catalysis in type II restriction endonucleases. Proc Natl Acad Sci U S A 95, 13489-13494
29. Ho, M. H., De Vivo, M., Dal Peraro, M., and Klein, M. L. (2010) Understanding the Effect of Magnesium Ion Concentration on the Catalytic Activity of Ribonuclease H through Computation: Does a Third Metal Binding Site Modulate Endonuclease Catalysis? J Am Chem Soc 132, 13702-13712
30. Tomlinson, C. G., Syson, K., Sengerova, B., Atack, J. M., Sayers, J. R., Swanson, L., Tainer, J. A., Williams, N. H., and Grasby, J. A. (2011) Neutralizing mutations of carboxylates that bind metal 2 in T5 flap endonuclease result in an enzyme that still requires two metal ions. The Journal

of biological chemistry 286, 30878-30887
31. Ivanov, I., Tainer, J. A., and McCammon, J. A. (2007) Unraveling the three-metal-ion catalytic mechanism of the DNA repair enzyme endonuclease IV. Proc Natl Acad Sci U S A 104, 1465-1470
32. 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 U S A 88, 11450-11454.
33. 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
34. Gorman, M. A., Morera, S., Rothwell, D. G., de La Fortelle, E., Mol, C. D., Tainer, J. A., Hickson, I. D., and Freemont, P. S. (1997) The crystal structure of the human DNA repair endonuclease HAP1 suggests the recognition of extra-helical deoxyribose at DNA abasic sites. Embo J 16, 6548-6558
35. Mol, C. D., Izumi, T., Mitra, S., and Tainer, J. A. (2000) DNA-bound structures and mutants reveal abasic DNA binding by APE1 and DNA repair coordination Nature 403, 451-456
36. Beernink, P. T., Segelke, B. W., Hadi, M. Z., Erzberger, J. P., Wilson, D. M., 3rd, and Rupp, B. (2001) Two divalent metal ions in the active site of a new crystal form of human apurinic/apyrimidinic endonuclease, Ape1: implications for the catalytic mechanism. Journal of molecular biology 307, 1023-1034
37. Georgiadis, M. M., Luo, M., Gaur, R. K., Delaplane, S., Li, X., and Kelley, M. R. (2008) Evolution of the redox function in mammalian apurinic/apyrimidinic endonuclease. Mutat Res 643, 54-63
38. Maher, R. L., and Bloom, L. B. (2007) Pre-steady-state kinetic characterization of the AP endonuclease activity of human AP endonuclease 1. The Journal of biological chemistry 282, 30577-30585
39. Erzberger, J. P., and Wilson, D. M., 3rd. (1999) The role of Mg2+ and specific amino acid residues in the catalytic reaction of the major human abasic endonuclease: new insights from EDTA-resistant incision of acyclic abasic site analogs and site-directed mutagenesis. Journal of molecular biology 290, 447-457
40. Galburt, E. A., Chevalier, B., Tang, W., Jurica, M. S., Flick, K. E., Monnat, R. J., Jr., and Stoddard, B. L. (1999) A novel endonuclease mechanism directly visualized for I-PpoI. Nat Struct Biol 6, 1096-1099
41. Tsutakawa, S. E., Jingami, H., and Morikawa, K. (1999) Recognition of a TG mismatch: the crystal structure of very short patch repair endonuclease in complex with a DNA duplex. Cell 99, 615-623
42. Mundle, S. T., Fattal, M. H., Melo, L. F., Coriolan, J. D., O'Regan, N. E., and Strauss, P. R. (2004) Novel role of tyrosine in catalysis by human AP endonuclease 1. DNA repair 3, 1447-1455
43. Mundle, S. T., Delaney, J. C., Essigmann, J. M., and Strauss, P. R. (2009) Enzymatic mechanism of human apurinic/apyrimidinic endonuclease against a THF AP site model substrate. Biochemistry 48, 19-26
44. Oezguen, N., Schein, C. H., Peddi, S. R., Power, T. D., Izumi, T., and Braun, W. (2007) A "moving metal mechanism" for substrate cleavage by the DNA repair endonuclease APE-1. Proteins 68, 313-323
45. Haas, B. J., Sandigursky, M., Tainer, J. A., Franklin, W. A., and Cunningham, R. P. (1999) Purification and characterization of Thermotoga maritima endonuclease IV, a thermostable apurinic/apyrimidinic endonuclease and 3'-repair diesterase. J Bacteriol 181, 2834-2839
46. Izumi, T., Malecki, J., Chaudhry, M. A., Weinfeld, M., Hill, J. H., Lee, J. C., and Mitra, S. (1999) Intragenic suppression of an active site mutation in the human apurinic/apyrimidinic endonuclease. Journal of molecular biology 287, 47-57

47. Mantha, A. K., Oezguen, N., Bhakat, K. K., Izumi, T., Braun, W., and Mitra, S. (2008) Unusual role of a cysteine residue in substrate binding and activity of human AP-endonuclease 1. Journal of molecular biology 379, 28-37
48. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr 54 (Pt 5), 905-921
49. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr 66, 213-221
50. Navaza, J. (1994) Amore - an Automated Package for Molecular Replacement. Acta Crystallogr A 50, 157-163
51. Kissinger, C. R., Gehlhaar, D. K., and Fogel, D. B. (1999) Rapid automated molecular replacement by evolutionary search. Acta Crystallogr D Biol Crystallogr 55, 484-491
52. Bicknell, R., and Waley, S. G. (1985) Single-Turnover and Steady-State Kinetics of Hydrolysis of Cephalosporins by Beta-Lactamase-I from Bacillus-Cereus. Biochemical Journal 231, 83-88
53. Malygin, E. G., Ovechkina, L. G., Zinoviev, V. V., Linstrem, U. M., and Reich, N. O. (2001) DNA-(N4-cytosine)-Methyltransferase from Bacillus amyloliquefaciens: Kinetic and substrate-binding properties. Molecular Biology 35, 35-44
54. Leipold, M. D., Workman, H., Muller, J. G., Burrows, C. J., and David, S. S. (2003) Recognition and removal of oxidized guanines in duplex DNA by the base excision repair enzymes hOGG1, yOGG1, and yOGG2. Biochemistry 42, 11373-11381
55. Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins 65, 712-725
56. Perez, A., Marchan, I., Svozil, D., Sponer, J., Cheatham, T. E., 3rd, Laughton, C. A., and Orozco, M. (2007) Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers. Biophys J 92, 3817-3829
57. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of Simple Potential Functions for Simulating Liquid Water. J Chem Phys 79, 926-935
58. Kale, L., Skeel, R., Bhandarkar, M., Brunner, R., Gursoy, A., Krawetz, N., Phillips, J., Shinozaki, A., Varadarajan, K., and Schulten, K. (1999) NAMD2: Greater scalability for parallel molecular dynamics. J Comput Phys 151, 283-312
59. Tuckerman, M., Berne, B. J., and Martyna, G. J. (1992) Reversible Multiple Time Scale Molecular-Dynamics. J Chem Phys 97, 1990-2001
60. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: Visual molecular dynamics. J Mol Graph Model 14, 33-38
61. Harding, M. M. (2001) Geometry of metal-ligand interactions in proteins. Acta Crystallogr D Biol Crystallogr 57, 401-411
62. Rothwell, D. G., Hang, B., Gorman, M. A., Freemont, P. S., Singer, B., and Hickson, I. D. (2000) Substitution of Asp-210 in HAP1 (APE/Ref-1) eliminates endonuclease activity but stabilises substrate binding. Nucleic acids research 28, 2207-2213
63. Barzilay, G., Mol, C. D., Robson, C. N., Walker, L. J., Cunningham, R. P., Tainer, J. A., and Hickson, I. D. (1995) Identification of critical active-site residues in the multifunctional human DNA repair enzyme HAP1. Nat Struct Biol 2, 561-568
64. Masuda, Y., Bennett, R. A., and Demple, B. (1998) Dynamics of the interaction of human apurinic endonuclease (Ape1) with its substrate and product. The Journal of biological chemistry 273, 30352-30359

65. Castillo-Acosta, V. M., Ruiz-Perez, L. M., Yang, W., Gonzalez-Pacanowska, D., and Vidal, A. E. (2009) Identification of a residue critical for the excision of 3'-blocking ends in apurinic/apyrimidinic endonucleases of the Xth family. Nucleic acids research 37, 1829-1842
66. Yang, W. (2008) An equivalent metal ion in one- and two-metal-ion catalysis. Nature structural & molecular biology 15, 1228-1231
67. Huber, R., Langworthy, T. A., Konig, H., Thomm, M., Woese, C. R., Sleytr, U. B., and Stetter, K. O. (1986) Thermotoga-Maritima Sp-Nov Represents a New Genus of Unique Extremely Thermophilic Eubacteria Growing up to 90-Degrees-C. Archives of Microbiology 144, 324-333
68. Tomanicek, S. J., Hughes, R. C., Ng, J. D., and Coates, L. (2010) Structure of the endonuclease IV homologue from Thermotoga maritima in the presence of active-site divalent metal ions. Acta Crystallogr Sect F Struct Biol Cryst Commun 66, 1003-1012
69. Ivanov, I., Tainer, J. A., and McCammon, J. A. (2007) Unraveling the three-metal-ion catalytic mechanism of the DNA repair enzyme endonuclease IV. Proceedings of the National Academy of Sciences of the United States of America 104, 1465-1470
70. Sanner, M. F., Olson, A. J., and Spehner, J. C. (1996) Reduced surface: An efficient way to compute molecular surfaces. Biopolymers 38, 305-320
71. Kuhn, L. A., Siani, M. A., Pique, M. E., Fisher, C. L., Getzoff, E. D., and Tainer, J. A. (1992) The interdependence of protein surface topography and bound water molecules revealed by surface accessibility and fractal density measures. Journal of molecular biology 228, 13-22
72. Aramini, J. M., Cleaver, S. H., Pon, R. T., Cunningham, R. P., and Germann, M. W. (2004) Solution structure of a DNA duplex containing an alpha-anomeric adenosine: insights into substrate recognition by endonuclease IV. J Mol Biol. 338, 77-91
73. Lucas, J. A., Masuda, Y., Bennett, R. A., Strauss, N. S., and Strauss, P. R. (1999) Single-turnover analysis of mutant human apurinic/apyrimidinic endonuclease. Biochemistry 38, 4958-4964
74. Rothwell, D. G., and Hickson, I. D. (1996) Asparagine 212 is essential for abasic site recognition by the human DNA repair endonuclease HAP1. Nucleic acids research 24, 4217-4221
75. Lowry, D. F., Hoyt, D. W., Khazi, F. A., Bagu, J., Lindsey, A. G., and Wilson, D. M., 3rd. (2003) Investigation of the role of the histidine-aspartate pair in the human exonuclease III-like abasic endonuclease, Ape1. Journal of molecular biology 329, 311-322
76. Tubbs, J. L., Latypov, V., Kanugula, S., Butt, A., Melikishvili, M., Kraehenbuehl, R., Fleck, O., Marriott, A., Watson, A. J., Verbeek, B., McGown, G., Thorncroft, M., Santibanez-Koref, M. F., Millington, C., Arvai, A. S., Kroeger, M. D., Peterson, L. A., Williams, D. M., Fried, M. G., Margison, G. P., Pegg, A. E., and Tainer, J. A. (2009) Flipping of alkylated DNA damage bridges base and nucleotide excision repair. Nature 459, 808-813
77. Tsutakawa, S. E., Classen, S., Chapados, B. R., Arvai, A. S., Finger, L. D., Guenther, G., Tomlinson, C. G., Thompson, P., Sarker, A. H., Shen, B., Cooper, P. K., Grasby, J. A., and Tainer, J. A. (2011) Human flap endonuclease structures, DNA double-base flipping, and a unified understanding of the FEN1 superfamily. Cell 145, 198-211
78. Grasby, J. A., Finger, L. D., Tsutakawa, S. E., Atack, J. M., and Tainer, J. A. (2012) Unpairing and gating: sequence-independent substrate recognition by FEN superfamily nucleases. Trends Biochem Sci 37, 74-84
79. Williams, R. S., Moncalian, G., Williams, J. S., Yamada, Y., Limbo, O., Shin, D. S., Groocock, L. M., Cahill, D., Hitomi, C., Guenther, G., Moiani, D., Carney, J. P., Russell, P., and Tainer, J. A. (2008) Mre11 dimers coordinate DNA end bridging and nuclease processing in double-strand-break repair. Cell 135, 97-109
80. Li, C. L., Hor, L. I., Chang, Z. F., Tsai, L. C., Yang, W. Z., and Yuan, H. S. (2003) DNA binding and cleavage by the periplasmic nuclease Vvn: a novel structure with a known active site. EMBO J 22, 4014-4025
81. Truglio, J. J., Rhau, B., Croteau, D. L., Wang, L., Skorvaga, M., Karakas, E., DellaVecchia, M. J., Wang, H., Van Houten, B., and Kisker, C. (2005) Structural insights into the first incision reaction during nucleotide excision repair. EMBO J 24, 885-894

82. Miller, M. D., Cai, J., and Krause, K. L. (1999) The active site of Serratia endonuclease contains a conserved magnesium-water cluster. Journal of molecular biology 288, 975-987
83. Dalhus, B., Arvai, A. S., Rosnes, I., Olsen, O. E., Backe, P. H., Alseth, I., Gao, H., Cao, W., Tainer, J. A., and Bjoras, M. (2009) Structures of endonuclease V with DNA reveal initiation of deaminated adenine repair. Nature structural & molecular biology 16, 138-143
84. Flick, K. E., Jurica, M. S., Monnat, R. J., Jr., and Stoddard, B. L. (1998) DNA binding and cleavage by the nuclear intron-encoded homing endonuclease I-PpoI. Nature 394, 96-101
85. Barnes, T., Kim, W. C., Mantha, A. K., Kim, S. E., Izumi, T., Mitra, S., and Lee, C. H. (2009) Identification of Apurinic/apyrimidinic endonuclease 1 (APE1) as the endoribonuclease that cleaves c-myc mRNA. Nucleic acids research 37, 3946-3958
86. Wu, S. I., Lo, S. K., Shao, C. P., Tsai, H. W., and Hor, L. I. (2001) Cloning and characterization of a periplasmic nuclease of Vibrio vulnificus and its role in preventing uptake of foreign DNA. Appl Environ Microbiol 67, 82-88
87. Nestle, M., and Roberts, W. K. (1969) An extracellular nuclease from Serratia marcescens. II. Specificity of the enzyme. The Journal of biological chemistry 244, 5219-5225
88. Lipton, A. S., Heck, R. W., Primak, S., McNeill, D. R., Wilson, D. M., 3rd, and Ellis, P. D. (2008) Characterization of Mg2+ binding to the DNA repair protein apurinic/apyrimidic endonuclease 1 via solid-state 25Mg NMR spectroscopy. J Am Chem Soc 130, 9332-9341
89. Oezguen, N., Mantha, A. K., Izumi, T., Schein, C. H., Mitra, S., and Braun, W. (2011) MD simulation and experimental evidence for Mg(2)+ binding at the B site in human AP endonuclease 1. Bioinformation 7, 184-198

ACKNOWLEDGEMENTS

Data was collected at the Advanced Light Source (ALS) and Stanford synchrotron radiation laboratory (SSRL) crystallography beamlines. The Berkeley Center for Structural Biology is supported in part by NIH, NIGMS, and HHMI. The ALS is supported under DOE Contract No. DE-AC02-05CH11231. SSRL is supported by DOE, OBER, NIH, NCRR, Biomedical Technology Program, and the NIGMS. Work on APE1 and Nfo is supported by NCI P01 CA92584, NIGMS GM046312, CA053791 and NSF-CAREER MCB-1149521. KH was supported in part by Japan Society for the Promotion of Science Fellowship and the Skaggs Institute for Chemical Biology. Computational resources were provided by the NSF XSEDE program (CHE110042). Numan Oezguen and Werner Braun kindly shared their coordinates for the model of the moving metal site. We thank James Holton and the Phenix team, Tom Terwilliger, Pavel Afonine, Nathaniel Echols, and the staff of the SIBLYS beamline for help and suggestions.

FIGURE LEGENDS

FIGURE 1. Crystallographic structures of human APE1 and TmaNfo showing active site details. (A) The 2.4 Å structure of an APE1:product complex reveals a Mg^{2+}-water cluster in the active site. Waters (red spheres) are shown with electron density from a PHENIX kicked map (1 σ, blue). (B) Simplified view of APE1 Mg-water cluster shows tetrahedral geometry. (C) The TmaNfo active site structure, determined at 0.92Å resolution. Electron density from a PHENIX kicked map (2.3 σ) is shown in blue. Mg^{2+} and Mn^{2+}/Zn^{2+} atoms are shown as purple and dark blue spheres. A coordinated water is shown as a red sphere. (D) The superimposition of the actives sites from TmaNfo and EcoNfo reveals a single difference in the active site. Ala30 and two waters in EcoNfo is replaced with Gln and may account for greater thermostability. These residues are next to Glu161 (Eco) that is postulated to activate the attacking water.

FIGURE 2. Alignment of Nfo sequences shows that thermophilic Nfo have substituted Gln for Ala in the active site. The region around Ala30 in *Eco* is shown, aligned with *B. thetaiotaomicron*, *T. maritima*, *A.*

aeolicus, mesophilic *B. subtilis* and *T. thermophilus*, respectively.

FIGURE 3. Comparison of the fractal dimension in the DNA binding grooves of Nfo:product (1qum.pdb), APE1:product (4iem.pdb) and UDG (1emh.pdb) highlights the shallow pockets of the AP endonucleases. Each protein's molecular surface (calculated using MSMS with a 1.4 Å probe sphere) is colored by its local atomic fractal density (70). The density is calculated using Surfractal with a 1.0-10.0 Å radius range (71). This Hausdorff-Besicovitch dimension measures the change in packing density: 2.0 indicates a flat surface, 3.0 a fully-packed volume. Intermediate values identify concave grooves and pockets. Figure created in AVS (AudioVisualSystemsInc).

FIGURE 4. Structural superimposition of APE1:product DNA (4iem.pdb) and Nfo:product DNA (1qum.pdb) based on the tetrahydrofuran moiety reveals geometic similarity. (A) Superimposition based on the tetrahydrofuran moiety shows how similar the tetrahydrofuran is deformed in the APE1 and Nfo structures and the close placement of the 3′ ribose oxygens. (B) Superimposition based on the scissile phosphate and the 3′ ribose oxygen shows how structurally similar the DNA products are, in contrast to the lack of conservation in tertiary structure of the proteins. (C) Close-up views of the active site showing the relative positioning of the scissile phosphate of the tetrahydrofuran and the 3′ ribose oxygen to the Mg in APE1 and the three Zn atoms in Nfo. The active side residues of APE1 are shown relative to the Zn^{2+} atoms in Nfo.

FIGURE 5. Single turnover kinetics of WT and APE1 mutants shows the relative importance of D210 and N212. A) Kinetics of product formation of three independent experiments for each enzyme at 10 ºC. B) Rate constants of AP site cleavage (nmol product formation per min) at 10ºC. 52nt-THF containing oligo duplex (10 nM) substrate and APE1 (100 nM) were used.

FIGURE 6 Catalytic mechanism for APE1 and Nfo. (A) MD simulation of WT APE1:Substrate DNA identified a persistent hydrogen bonding network, involving a water, residues His309, Tyr171, Asp210, Asn212 and the abasic site phosphate. A snapshot from the MD trajectory shows the positions of the active site residues; hydrogen bonds (black dotted lines) and the nucleophilic attack direction (gray arrow) are shown; all distances are in Å. Tyr171 is below but not labeled. (B) Model of Nfo:substrate, built from the substrate complex with E261Q (2nqj.pdb) and with Glu261 overlaid from the WT model (1qum.pdb).

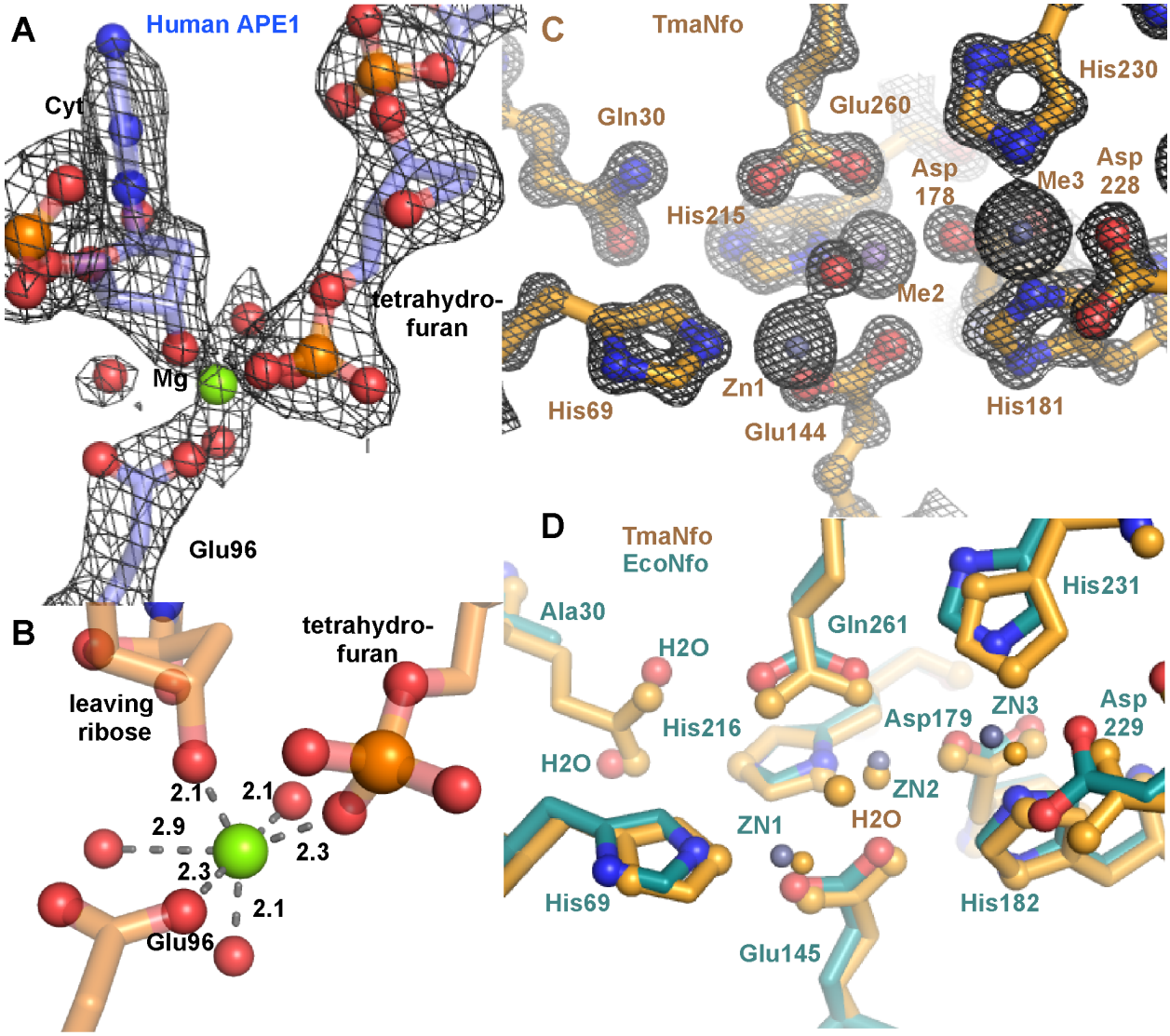

A
Human APE1
Cyt
tetrahydro-
furan
Mg
Glu96
B
tetrahydro-
furan
leaving
ribose
2.1
2.1
2.9
2.3
2.3
2.1
Glu96
C
TmaNfo
Gln30
Glu260
His230
Asp
178
Me3
Asp
228
His215
Me2
Zn1
His69
Glu144
His181
D
TmaNfo
EcoNfo
Ala30
H2O
His216
H2O
Gln261
His231
Asp179
ZN3
Asp
229
ZN2
H2O
ZN1
His69
Glu145
His182

Figure 2.

```
Eco 24- IDATAFALFTKN -35 mesophilic
Bth 24- IGANAFALFTKN -35 mesophilic
Tma 24- IGGNSFQIFPHN -35 thermophilic
Aae 24- IGAEVFQFFLRS -35 thermophilic
Bsu 26- YGANTFMIYTGA -37 mesophilic
Tth 26- LGLTAFQIFAKS -37 thermophilic
```

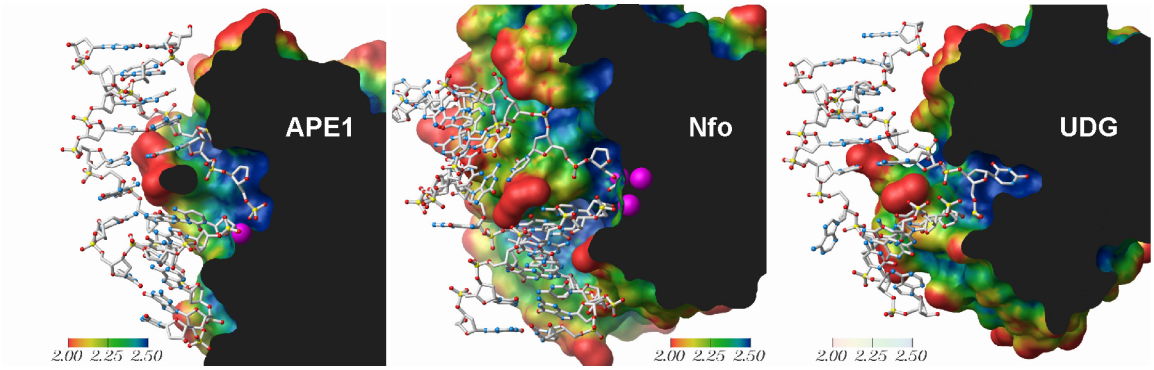

APE1
2.00 2.25 2.50
Nfo
2.00 2.25 2.50
UDG
2.00 2.25 2.50

Figure 4

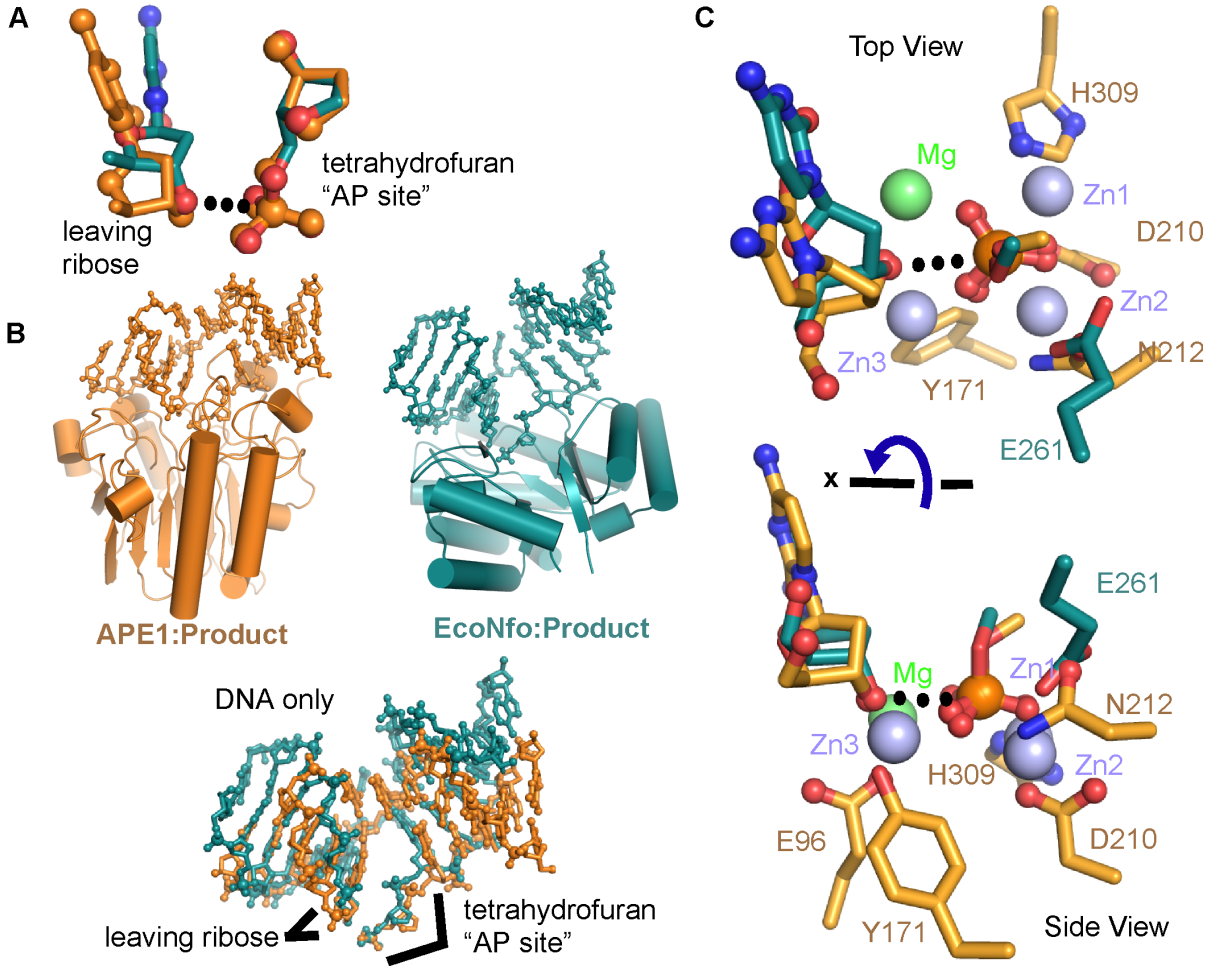
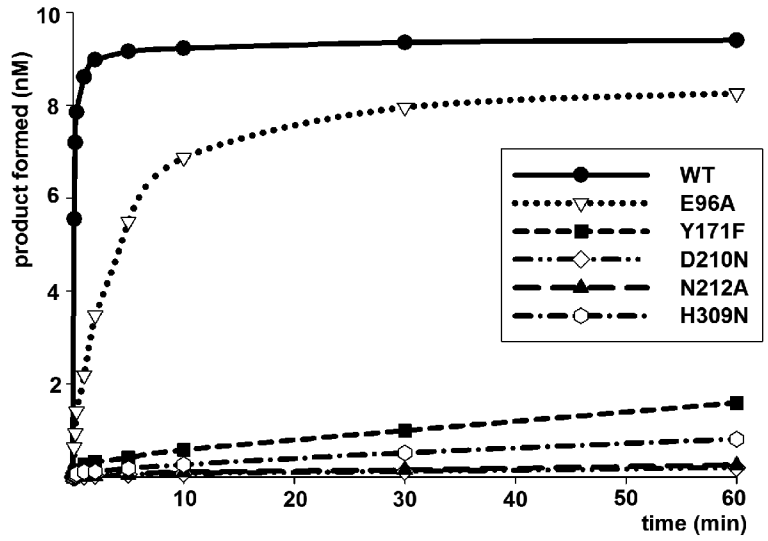

A
tetrahydrofuran
"AP site"
leaving
ribose
B
APE1:Product
EcoNfo:Product
DNA only
leaving ribose
tetrahydrofuran
"AP site"
C
Top View
H309
Mg
Zn1
D210
Zn2
Zn3
Y171
N212
E261
x
E261
Mg
Zn1
N212
Zn3
H309
Zn2
E96
D210
Y171
Side View

Figure 5

A



B

APE1 (WT and mutants)	Rate Constant for AP site cleavage [nmol/min]	Relative Activity
Wild type	33.3±3.281	1000±98.54
E96A	2.19±0.338	65.765±10.170
Y171F	0.0265±0.0065	0.796±0.197
D210N	0.0032±0.0018	0.097±0.055
N212A	0.0043±0.0013	0.129±0.039
H309N	0.0135±0.0042	0.407±0.128

Figure 6

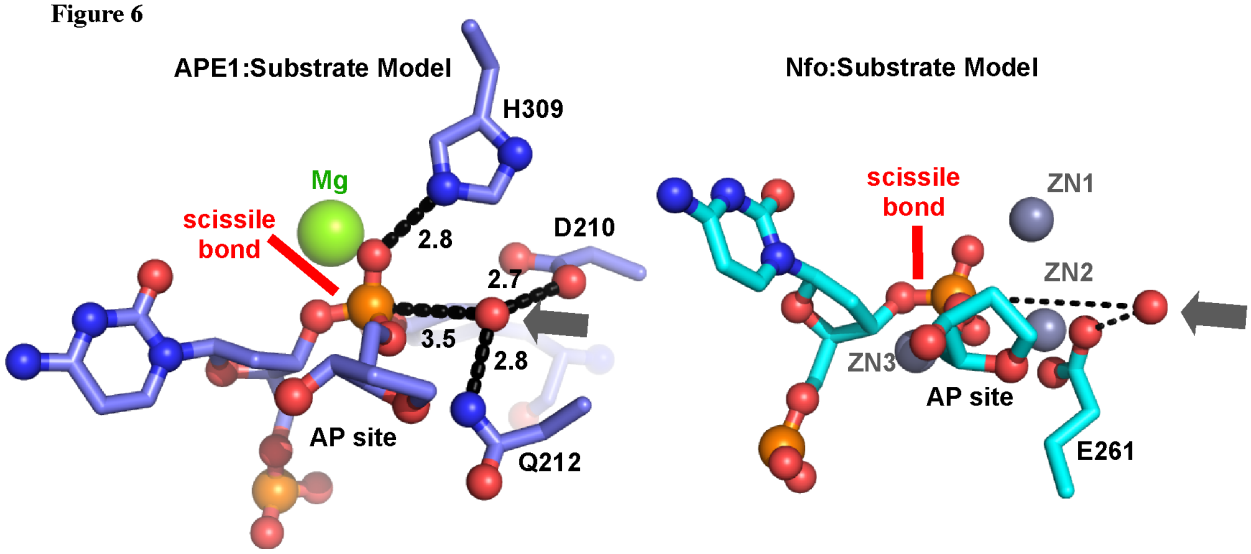

APE1:Substrate Model
H309
Mg
scissile
bond
2.8
D210
2.7
3.5
2.8
AP site
Q212
Nfo:Substrate Model
scissile
bond
ZN1
ZN2
ZN3
AP site
E261

Table 1. X-ray diffraction data collection and refinement statistics for APE1-product DNA complex and Nfo protein. Numbers in parentheses are from the high resolution shell.

Protein	**APE1 (4IEM)**	**Nfo (4HNO)**
DNA	Tetrahydrofuran product	
Species	*H. sapiens*	*T. maritima*
Synchrotron	ALS	SSRL
Data Collection		
Space Group	$P2_1$	$P6_1$
Cell dimensions		
a,b,c (Å)	104.6, 74.1, 112.1	123.370, 123.370, 35.395
a, b, g (°)	90.0, 112.0, 90.0	90.0, 90.0, 120.0
Resolution (Å)	30-2.4 Å (2.48-2.40)	30-0.92 (0.95-0.92)
Rsym[a]	0.064 (0.402)	0.053
I/σ(I)	11.85 (4.7)	26.5 (2.6)
Completeness (%)	90.1 (87.7)	95.9
Redundancy	3.1 (3.2)	2.84 (2.48)
Wilson B factor	42.2	10.2
Refinement		
Reflections (min F/σF)	56727 (1.34)	204917 (1.34)
R_{cryst} (%)[b]	0.187	0.1248
R_{free} (%)[c]	0.250	0.1390
Molecules per ASU	4 protein, 4 dsDNA	1
Number of atoms	20539	5274
Protein/DNA atoms	10530	4780
Metal atoms	7 Mg^{2+}	Zn^{2+}, 2 Zn^{2+}/Mn^{2+}, Mg^{2+}
Other non-water solvent	9	61
Water atoms	272	427
B factors		
Protein/DNA	50.0	16.4
Metals	51.0	10.9
Other non-water solvent	66.1	19.7
Water	42.4	26.5
R.m.s. deviations		
Bond length (Å)	0.004	0.017
Bond angles (°)	0.92	1.311
Ramachandran Favored (%)	96	98.3
Ramachandran Allowed (%)	3.7	1.7
Ramachandran Outliers (%)	0.3	0

[a]R_{sym} = the unweighted R value on I between symmetry mates.
[b]$R_{cryst} = \Sigma_{hkl}||F_{obs}(hkl)| - |F_{calc}(hkl)||/\Sigma_{hkl}|F_{obs}(hkl)|$.
[c]R_{free} = the cross validation R factor for 5% of reflections against which the model was not refined.

DISCLAIMER