Structural insights by molecular dynamics simulations into differential repair efficiency for ethano-A versus etheno-A adducts by the human alkylpurine-DNA *N*-glycosylase

Anton B. Guliaev, Bo Hang and B. Singer*

Donner Laboratory, Life Sciences Division, Lawrence Berkeley National Laboratory, University of California, Berkeley, CA 94720, USA

Received May 3, 2002; Revised and Accepted July 8, 2002

ABSTRACT

1,*N*⁶-ethenoadenine adducts (ϵ A) are formed by known environmental carcinogens and found to be removed by human alkylpurine-DNA *N*-glycosylase (APNG). 1,*N*⁶-ethenoadenine (EA) adducts differ from ϵ A by change of a double bond to a single bond in the 5-member exocyclic ring and are formed by chloroethyl nitrosoureas, which are used in cancer therapy. In this work, using purified recombinant human APNG, we show that EA is a substrate for the enzyme. However, the excision efficiency of EA was 65-fold lower than that of ϵ A. Molecular dynamics simulation produced similar structural motifs for ϵ A and EA when incorporated into a DNA duplex, suggesting that there are no specific conformational features in the DNA duplex which can account for the differences in repair efficiency. However, when EA was modeled into the APNG active site, based on the APNG/ ϵ A-DNA crystallographic coordinates, in structures produced by 2 ns molecular dynamics simulation, we observed weakening in the stacking interaction between EA and aromatic side chains of the key amino acids in the active site. In contrast, the planar ϵ A is better stacked at the enzyme active site. We propose that the observed destabilization of the EA adduct at the active site, such as reduced stacking interactions, could account for the biochemically observed weaker recognition of EA by APNG as compared to ϵ A.

INTRODUCTION

The saturated exocyclic adduct of adenine, 1,*N*⁶-ethanoadenine (EA) (Fig. 1A), has been identified as one of the products of the reaction of 1,3-bis(2-chloroethyl)nitrosourea (BCNU) with DNA (1,2). BCNU belongs to the family of therapeutic nitrosourea compounds used in cancer treatment. The ethano adducts in DNA structurally resemble the exocyclic etheno adducts formed from the reaction of the

chemical carcinogen vinyl chloride with DNA (3,4) or by lipid peroxidation (5). The etheno adducts, particularly 1,*N*⁶-ethenoadenine (ϵ A), have been extensively studied biochemically and structurally (6). It has been shown that this adduct can be efficiently removed from DNA by rodent or human alkylpurine-DNA *N*-glycosylase (APNG) (also termed alkyladenine DNA glycosylase, AAG) (7–10). The mechanism of ϵ A excision by APNG has been proposed based on the 2.1 Å crystal structure of an APNG mutant protein (E125Q) complexed to ϵ A-containing DNA (11). Crystallization of the protein–substrate complex was made possible by substitution of Glu125 with a glutamine residue, which prevents activation of the active site bound water acting as a nucleophile. The authors showed that flipped-out ϵ A has the ability to stack in a stable position between the aromatic side chains in the enzyme active site (11). The position of the adduct was also stabilized by a key hydrogen bond between the main chain of His136 and N9 of ϵ A, which offered a unique acceptor lone pair essential for hydrolysis of the C1'–N glycosylic bond. The His136 side chain forms hydrogen bond interactions to Tyr157 and the phosphate group of ϵ A (Fig. 1B).

The ethano adducts differ from etheno adducts by the change of a double bond to a single bond in the 5-member exocyclic ring (Fig. 1A). In this work we have addressed the issue of whether such a small structural change could affect the recognition and repair efficiency of EA compared to ϵ A by human APNG. Recent work in this laboratory showed that a small structural change in the adduct structure has an effect on DNA glycosylase activity (12). Addition of a hydroxymethyl group to the C8 position of 3,*N*⁴-ethenocytosine (ϵ C) to form 8-(hydroxymethyl)-3,*N*⁴-ethenocytosine (8-HM- ϵ C), a product of the reaction with the mutagen/carcinogen glycidaldehyde, reduced the repair efficiency by *Escherichia coli* mismatch uracil-DNA glycosylase (Mug) by 2.5-fold as compared with that of the structurally related ϵ C. However, molecular dynamics simulation showed similar alignment and hydrogen bonding patterns for both adduct pairs in the 25mer oligomer duplexes used in the biochemical studies (12). The lower Mug activity toward 8-HM- ϵ C suggests some degree of steric hindrance to the binding or catalytic activity as a result of the hydroxymethyl group on the etheno ring.

*To whom correspondence should be addressed. Tel: +1 510 642 0637; Fax: +1 510 486 6488; Email: abguliaev@lbl.gov

Figure 1. (A) Chemical structures of the εA and EA adducts. (B) The APNG active site structure showing the stacking between εA and aromatic side chains of Tyr127, His136 and Tyr157. The black dashed lines show hydrogen bonds between εA N9 and His136 N, Tyr157 O4 and His136 Nτ and εA OP1 and His136 Nτ. The picture was generated using the atomic coordinates of the crystallized εA-DNA/APNG complex [PDB ID code 1f4r (11)].

In this work the reparability by human APNG of EA, incorporated into a 25mer DNA duplex, was investigated and compared to the repair efficiency of εA by the same enzyme using a DNA glycosylase assay. EA was found to be a substrate for the human enzyme, but a much weaker one than εA. The observed difference in rate of excision of εA versus EA adducts was correlated with the structural data obtained by molecular modeling. The availability of crystal data for the APNG enzyme complexed to εA-containing DNA allowed us to use it as a starting point in our molecular modeling. The observation of structural perturbations caused by replacement of εA by EA in the enzyme active site might have an effect on the substrate preference of εA over EA. The complementarity between the substrate and enzyme active site should be one of the factors responsible for the catalytic specificity and efficiency of repair. However, a number of other events, such as initial lesion binding/recognition, ease of rotating the damaged base from the DNA ladder and stabilization of the extrahelical conformation, also contribute to the efficiency of repair for a particular adduct. These factors can be influenced by the conformational features of the adduct-containing duplexes. To evaluate the effects of the EA adduct on the local and global structural features of the DNA duplex we performed simulation of an EA-T-containing 25mer DNA duplex. These data were compared to the εA-T- and A-T-containing duplexes.

MATERIALS AND METHODS

Oligonucleotides

Synthesis of the EdA phosphoramidite and its site-specific incorporation into oligonucleotides was described by Maruenda *et al.* (13). The edA phosphoramidite was purchased from Glen Research (Sterling, VA). Both derivatives were placed in the sixth position from the 5'-end of a 25mer sequence (X): 5'-CCG CTX GCG GGT ACC GAG CTC GAA T-3'. The unmodified 25mer and complementary strands with T opposite the modified base were purchased from Operon Technologies (Alameda, CA). All the oligomers were purified by HPLC and denaturing PAGE.

DNA glycosylase assay

The enzymatic assay used to test APNG-mediated cleavage of EA or εA from oligonucleotides was carried out essentially as previously described (14,15). Briefly, 25mer oligonucleotides were 5'-end-labeled with [γ -³²P]ATP (specific activity 6000 Ci/mmol, 1 Ci = 37 GBq; Amersham Pharmacia Biotech) and annealed to a complementary strand in a 1:1.5 molar ratio. The standard reactions (10 μ l) contained 2 nM 5'-³²P-end-labeled oligomer duplex in 10 mM HEPES-KOH, pH 7.4, 100 mM KCl, 1 mM EDTA, 1 mM EGTA, 1 mM DTT, 0.1 mg/ml acetylated BSA and varying amounts of human APNG protein (a gift from Dr Tim O'Connor, Beckman Research Institute, Duarte, CA) (16). In these reactions, a 5' AP endonuclease, the major human AP endonuclease (HAP1) (a gift from Dr Ian Hickson, Oxford University, Oxford, UK), was added to cleave the apurinic (AP) site resulting from the excision of EA or εA by APNG protein. The reactions were stopped by adding equal amounts of F/E solution (90% formamide plus 50 mM EDTA) and then heated at 95–100°C for 3 min. Reactions were then resolved by 12% polyacrylamide–8 M urea denaturing PAGE. For band quantitation, a Bio-Rad FX molecular phosphorimager and Quantity One software (v.4.0.1) were used.

Molecular modeling

EA- and εA-containing 25mer DNA duplexes. A set of force field parameters for εA was previously developed using an *ab initio* quantum mechanical calculation and procedure described in earlier publications from this laboratory (12,17). The EA adduct was built by saturating the C7=C8 double bond in the imidazole ring of εA. Atom-centered charges were calculated with the RESP module of AMBER 6.0 using the partial charges obtained by Hartree-Fock calculation using the 6-311G* basis set in the Spartan 5.0 suite (Wavefunction, Inc., Irvine, CA). Prior to the charge calculations, the conformation of EA was geometry optimized using the 6-31G* basis set. The εA and EA adducts were incorporated in the sixth position into the 25mer sequence used in biochemical studies (for sequence see Materials and Methods). The topology and coordinate files for the three DNA duplexes (εA-T-DNA, EA-T-DNA and A-T-DNA, used as a control) were generated with the xLeap module of AMBER 6.0 (18). Forty-eight Na⁺ ions were placed around the phosphate groups to neutralize negative charges, and an aqueous environment was represented by a rectangular water box, which provided no less than 10 Å of TIP3P water molecules around the solute. Two

Figure 2. (Left) Protein-dependent cleavage of a 25mer oligonucleotide containing either EA or ϵ A by human APNG protein. Increasing amounts of APNG protein (0.3–4.8 ng for EA and 0.01–0.15 ng for ϵ A) were incubated with 2 nM 32 P-end-labeled oligomer substrates for 30 min at 37°C. The AP site produced by DNA glycosylase action was further cleaved by adding HAP1 (5 ng), a 5' AP endonuclease, to the reaction mixture. Note that the use of HAP1 alone had no detectable effect on either EA- or ϵ A-containing templates. (Right) Time-dependent cleavage of a 25mer oligonucleotide containing either EA or ϵ A. Oligomer duplexes were reacted with 3 (for EA) or 0.15 ng (for ϵ A) APNG protein for varying times at 37°C. The scanning data were normalized as nM oligomer substrate cleaved per ng APNG protein. (Inset) Detailed time-dependent response of EA excision by APNG.

nanosecond molecular dynamics simulation runs at 310 K, using particle-mesh Ewald (PME) to treat Coulombic interactions and a 2 fs time step, were generated after the system achieved the correct density and volume (17).

EA-DNA/APNG and ϵ A-DNA/APNG complexes. In this work the high resolution X-ray crystal structure of the ϵ A-DNA/APNG complex (PDB code 1f4r) served as the starting structure. Hydrogens were added using the xLeap module of AMBER 6.0. To generate the EA-DNA/APNG complex, ϵ A was replaced by the geometry optimized EA adduct, using Insight II (Biosym/MSI, San Diego, CA). Two sets of topology and coordinate files for the APNG protein complexed to ϵ A-DNA and EA-DNA were generated using the xLeap module of AMBER 6.0. A rectangular box of TIP3P water molecules was added, providing at least 10 Å of explicit solvent around each DNA/protein complex, yielding 9292 water molecules. The complete system consisted of approximately 31 728 atoms and has the initial dimensions 73.439, 70.948 and 76.668 Å in the *x*, *y* and *z* directions, respectively. The initial density of the water around the protein was 0.806 g/cm³. Molecular dynamics simulations were carried out using the SANDER module of AMBER 6.0 with a 2 fs time step. SHAKE was applied to all hydrogen atoms and a 10 Å cut-off was used for Lennard–Jones interactions. Constant pressure was maintained with isotropic scaling. All long-range electrostatic interactions were handled using the PME method. In the beginning of the simulations, the water box was subjected to a series of equilibration molecular dynamics runs while holding the DNA/APNG complex fixed, and was similar to the procedure used for the DNA duplexes. The equilibration runs began with 1000 steps of minimization followed by 10 ps of simulation, during which the temperature was slowly raised from 0 to 310 K and kept at this temperature for another 50 ps. During the first 30 ps of simulation the water density and pressure converge to the correct values (1.01 g/cm³

and 1 atm, respectively). This was followed by a second set of 1000 steps of minimization and 3 ps of simulation, which were carried out with the restraints on the solute molecule reduced to 25 kcal/mol. Finally, five rounds of 800 steps of conjugate gradient minimization were performed with the positional restraints reduced by 5.0 kcal/mol in each round. The unrestrained molecular dynamics production runs of 2 ns were initiated after the last round of minimization. The final structures representing the conformational family for the DNA/enzyme complexes produced by molecular dynamics simulation were generated by averaging the molecular dynamics trajectories based on root mean square deviation (RMSD) profiles (from 0.4 to 2 ns).

Structural analysis and calculations. The molecular dynamics trajectories were processed using the analytical modules of AMBER 6.0 and visually analyzed with the VMD program (19). Nucleic acid structural parameters were derived using CURVES 5.1 (20). Production runs for the 25mer DNA duplexes and DNA/APNG complexes were carried out on 64 processors (16 processors per node) using the IBM SP RS/6000 supercomputer available at the National Energy Research Scientific Computing Center, Lawrence Berkeley National Laboratory. The equilibration runs and trajectory analysis were performed on a Silicon Graphics Origin 200 server interfaced with a dual processor Octane workstation.

RESULTS

Biochemical assay

We first tested the excision activity of APNG protein towards EA since this enzyme excises the closely related adduct ϵ A as well as another ethano adduct, *N*²,3-ethanoguanine (8,10,21). As shown in Figure 2 (left), APNG protein showed a protein-dependent cleavage of a 32 P-end-labeled EA-containing 25mer oligomer duplex (EA-T). The cleavage products

Figure 3. Autoradiogram of gel electrophoresis of 5'-³²P-labeled oligonucleotides after reaction with human APNG for varying times (0–60 min). The amount of APNG used for these reactions was 3 ng for EA and 0.15 ng for ϵ A excision. For reaction and gel procedure details see Materials and Methods. The arrows show the position of the 5mer cleavage product. M, a 5mer marker with the same sequence as the expected cleavage product.

from EA- and ϵ A-containing oligonucleotides after 5' AP endonuclease treatment were both ³²P-labeled 5mers are shown in Figure 3 (arrows). These are the expected products resulting from the 5' hydrolysis by HAP1 of an AP site at the sixth position. However, the extent of EA excision by APNG was much lower than that of ϵ A excision, as indicated in Figure 2 (left).

Excision of EA from the 25mer DNA duplex by APNG protein as a function of time is shown in Figure 3, in which a comparison was made between the rate of excision of EA and ϵ A under the same assay conditions except that the amount of APNG used was different (3 and 0.15 ng for EA and ϵ A, respectively). In Figure 2 (right) the scanning results were normalized per ng protein in order to compare the rates of these two activities. It is evident that the excision of EA occurs much more slowly than that of ϵ A, with an ~65-fold difference.

Conformation of the EA- versus ϵ A-containing duplexes

Two nanosecond molecular dynamics calculations were performed for the two 25mer DNA duplexes used in the biochemical studies (ϵ A-T-DNA and EA-T-DNA) and a corresponding control A-T-DNA (25mer DNA duplex with an unmodified A at the sixth position). The conformational stability was evaluated by calculating RMSD values of each picosecond relative to the coordinates of the initial energy minimized structures for all three DNA duplexes. RMSD values for all atoms, the five central base pairs and the adduct-containing base pair are shown in Figure 4A and B for ϵ A-T- and EA-T-DNA, respectively. Based on RMSD values, both structures reached conformational equilibrium after the first 400 ps and showed a plateau for the rest of the simulation. To

Figure 4. Time dependence of RMS deviations of the 25mer DNA duplexes containing ϵ A-T (A) and EA-T lesions (B). The data is shown for all atoms (black), the lesion (red) and 5 bp with the lesion in the middle (blue) ($C_4T_5\epsilon A/EA_6G_7C_8/G_{47}A_{46}T_{45}C_{44}G_{43}$). Both structures reached conformational equilibrium after the first 400 ps.

monitor integrity of the duplex during the simulation we calculated Watson–Crick hydrogen bond distances and percentage occupancy for all base pairs in the duplexes. All hydrogen bonds, including 5'-TA and 3'-GC base pairs flanking the adduct site, were 98–100% occupied during the entire simulation (data not shown). Terminal bases were not included in hydrogen bond calculations due to known fraying effects, which were also observed in our simulations. No hydrogen bonding was observed in the ϵ A-T and EA-T mismatches. Top and side views for the $T_5X_6G_7/A_{46}T_{45}C_{44}$ motifs, where X = A, ϵ A or EA, are shown in Figure 5. In both lesion-containing duplexes, compared to the unmodified duplex, the adduct was displaced towards the major groove, while the opposite T remained stacked between A and C bases. Figure 6 shows average values for the inter- and intra-base pair parameters (Fig. 6A and B, respectively) of the 5 bp for the ϵ A- and EA-containing duplexes and corresponding control. Average values were calculated over the simulation trajectory. Presence of the adduct had a similar effect on the conformation of the mismatch and neighboring bases in ϵ A-T-DNA and EA-T-DNA. A positive shear (SHR) value was observed for both ϵ A-T and EA-T base pairs, indicating the magnitude of displacement of the adduct towards the major groove (Fig. 5). Another two intra-base pair parameters affected by the presence of either ϵ A or EA in the DNA duplex were buckle and propeller twist. Considerable propeller twist (15–23°), compared with the unmodified DNA (<8°) was observed for the T_5-A_{46} , $\epsilon A/EA-T_{45}$ and C_8-G_{43} base pairs in

Figure 5. Top (left) and major groove (right) views for the 3 bp motifs for the A-T-, ϵ A-T- and EA-T-containing 25mer DNA duplexes produced by 2 ns molecular dynamics simulations. EA and ϵ A adducts are shown in red and A and T bases are shown in blue. Both ϵ A-T and EA-T base pairs produced similar structural motifs with the adduct displaced toward the major groove and a non-planar alignment of the bases, as compared to the A-T pair. No hydrogen bonds were observed between the bases in the ϵ A-T and EA-T pairs. Yellow dashed lines show Watson-Crick hydrogen bonds for the A-T base pair.

the lesion-containing duplexes (Fig. 6A). Buckling around the lesion site was also larger in magnitude than for the unmodified DNA. Perturbations in inter-base pair parameters, which probably best describe stacking interactions, were similar in both the ϵ A and EA duplexes. The most noticeable differences from the unmodified duplex were observed for the tilt (TLT), roll (ROL) and twist (TWS) parameters (Fig. 6B). The high magnitude of TWS for the T₅-A₄₆/ ϵ A₆-T₄₅ and T₅-A₄₆/EA₆-T₄₅ base pair steps (58° and 50°, respectively) indicated a larger helical twist at these steps compared with the rest of the adduct-containing and unmodified duplexes. The succeeding steps, ϵ A₆-T₄₅/G₇-C₄₄ and EA₆-T₄₅/G₇-C₄₄, showed much smaller TWS values (2° and 4°, respectively), characteristic of untwisting of the DNA at the lesion site. In previous modeling work from our laboratory (17) we reported smaller TWS values at the ϵ A-T base pair in a 15mer DNA duplex. Moreover, the magnitude of TWS was sequence-dependent (17). The curvature of the DNA was calculated

using the CURVES 5.1 algorithm and was not affected by presence of the adduct. To avoid a contribution from the highly flexible DNA ends, the terminal base pairs were not included in the curvature measurements. The values for the EA-T-DNA and ϵ A-T-DNA duplexes were 14° and 11°, respectively. The sugar conformation of EA falls in the C2'-*endo* conformation, while ϵ A was closer to the C1'-*exo* range. Both adducts stack in an *anti* orientation into the DNA helix.

The conformation of the ϵ A-T base pair produced by our modeling approach was similar to the conformation of that lesion reported based on NMR data (22,23). However, in our modeling we observed a slightly bigger shift of ϵ A towards the major groove than was reported by NMR. The displacement of ϵ A along the *x*-axis towards the major groove was 2.5 Å, while EA was displaced by 2.0 Å, compared to unmodified A in the control duplex. The differences in the conformation of the ϵ A-T pair between NMR and modeling can be attributed to sequence-dependent effects (17).

Figure 6. (A) Average values for the intra-base pair parameters describing the geometry of base pairing for the 5 bp in the A-T-, ϵ A-T- and EA-T-containing duplexes. (B) Average values for the inter-base pair parameters describing the stacking interactions for the 4 bp steps in the A-T-, ϵ A-T- and EA-T-containing duplexes. The tick marks on the x-axis indicate the base pair step. For example, label C₄-T₅ corresponds to the C₄-G₄₇/T₅-A₄₆ base pair step.

Effect of the EA adduct on the APNG active site

The availability of crystal data for human APNG complexed to ϵ A-containing DNA allowed us to use this structure as a starting point in our molecular modeling study in which we

Figure 7. Time dependence of RMS deviations (RMSD) of the ϵ A-DNA/APNG (A) and EA-DNA/APNG complexes (B). Black, complex (DNA + enzyme); blue, enzyme alone; green, DNA alone; red, active site. The conformational families produced by molecular dynamics simulation for the ϵ A-DNA/APNG and EA-DNA/APNG complexes deviate minimally from the crystal coordinates. High RMSD fluctuations for the DNA duplex (green traces) can be explained by the contribution of the more flexible ends. (C) The RMSD values for the ϵ A (blue) and EA (gray) adducts.

addressed the question of substrate preference of this enzyme for ϵ A over EA. Simple superimposition of EA over ϵ A did not reveal any conformational effects which EA might have on the active site of APNG enzyme. First, to validate our modeling protocol, we performed 2 ns simulation of the APNG/ ϵ A-DNA complex (PDB ID code 1f4r). The analysis of the overall structure and position of the adduct in the active site showed that the averaged minimized structure produced by molecular dynamics simulation deviates minimally from the crystal coordinates. All averaged RMSD values were <2.0 Å, with a value of 0.9 ± 0.06 Å for the enzyme active site, 1.7 ± 0.17 Å for the enzyme, 1.87 ± 0.31 Å for the DNA and 1.97 ± 0.25 Å for the all-atom RMSD for the entire structure (Fig. 7A). The largest RMSD fluctuations observed for the DNA duplex bound to enzyme can be explained by the

Figure 8. Superimposition of the ϵ A-DNA/APNG active site from the crystal structure (green) (PDB code 1b4r) over the ϵ A-DNA/APNG active site produced by 2 ns molecular dynamics simulation (red). The RMSD between active site conformations is <0.91 Å. The yellow dashed lines indicate the key hydrogen bond between the ϵ A adduct and main chain amide of His136 and two hydrogen bonds which stabilize the position of the His136 side chain. All three hydrogen bonds remained intact during molecular dynamics simulation.

contribution of more flexible DNA ends. All stacking and key hydrogen bond interactions in the active site remained intact during this simulation. The superimposition of the active sites of the crystal structure of ϵ A-DNA/APNG and the ϵ A-DNA/APNG complex produced by molecular dynamics simulation is shown in Figure 8. Note that molecular dynamics simulation produced more pronounced plane-to-plane stacking between His136 and the imidazole ring of ϵ A than in the crystal structure.

The RMSD values for the EA-DNA/enzyme complex showed a similar profile to that observed for the ϵ A-DNA/enzyme complex and indicated overall conformational stability for the system when ϵ A was replaced by the EA adduct (Fig. 7B). Slightly higher averaged RMSD values were observed for the active site of the enzyme complexed with EA-DNA than the active site of the ϵ A-DNA complex (1.1 ± 0.1 versus 0.9 ± 0.06 Å, respectively). Pertinent observations can be drawn from monitoring the RMSD values for the adduct itself during the course of the simulations. Only corresponding atoms between the structures were compared. The EA adduct showed significantly higher flexibility and larger deviation from the starting position in our simulation than the ϵ A adduct (Fig. 7C). The average RMSD value for EA was 0.36 ± 0.13 Å, while for the ϵ A adduct it was 0.2 ± 0.13 Å. The main conformational feature observed for the APNG active site complexed with EA-DNA was displacement and almost 70° rotation of the His136 side chain. This created an edge-to-edge packing interaction with EA, rather than the

much more stable face-to-face stacking observed between the planar ϵ A and His136 in the crystal structure (Fig. 9). Face-to-face stacking was also supported during our molecular dynamics simulation of the ϵ A-DNA/APNG complex. The change in the stacking interaction between EA and His136 resulted in a weakening of two hydrogen bonds: between the side chain of His136 and the 5'-phosphate of EA (ϵ A OP1-His136 N τ) and the side chain of His136 and Tyr157 (Tyr157 O4-His136 N π). The evolution of these hydrogen bonds over simulation time is shown in Figure 10. However, the key hydrogen bond between N9 of EA and NH of His136 remained intact in the EA-DNA/APNG complex (Fig. 10).

DISCUSSION

One of the most important steps in DNA base excision repair (BER) is recognition and excision of the damaged base from the DNA ladder by DNA glycosylases. This step is the key determinant of BER activity against a specific lesion. Recent crystallographic studies revealed details of the base excision mechanism of DNA glycosylases, including human APNG, showing that damaged DNA bases are excised by hydrolysis of the C1'-N glycosylic bond. The result of this reaction is a free DNA base and an abasic sugar residue, which is hydrolyzed by an AP endonuclease, followed by DNA synthesis and ligation, which restores the correct DNA sequence (24,25). The position of the adduct in the enzyme

Figure 9. APNG active site structure for the ϵ A-DNA/APNG (blue) and EA-DNA/APNG (yellow) complexes produced by molecular dynamics simulations. The green dashed lines indicate hydrogen bonds between ϵ A N9 and His136 NH in the APNG/ ϵ A-DNA complex and EA N9 and His136 NH in the APNG/EA complex. The steric clash between the EA exocyclic ring and His136 side chain (indicated by red arrow) resulted in reduced stacking interactions (edge-to-edge packing between EA and His136) and destabilized the position of the ethano adduct in the enzyme active site.

binding pocket is achieved by rotation of the damaged base out of the DNA in order that it can be inserted into the enzyme active site. Correct alignment of the modified base in the enzyme active site is one of the key steps for successful removal of that base from the DNA. Structural data on the enzyme complexed to adduct-containing DNA provide essential information on the interaction between the substrate and enzyme active site. The use of molecular modeling has allowed refinement of the conformation of DNA/enzyme complexes with adduct structures, which were not used in the X-ray crystallography studies. Additionally, structural data on adduct-containing DNA duplexes should provide valuable information on some initial steps of BER. Pronounced structural perturbation around the lesion might be a signal for the DNA repair enzyme to act on the substrate to prevent binding to the adduct-containing DNA motif. The stacking interaction between the adduct and the flanking bases, together with hydrogen bonding with the opposite base, should influence the ability of the modified base to be flipped out from the DNA duplex into the enzyme binding site. The observed conformational features of the adduct-containing duplexes and enzyme active site bound to the modified base should be carefully examined and compared with the biochemical data, thus providing a possible explanation for differential repair by the particular enzyme.

In this work we have demonstrated that human APNG recognizes and excises an EA adduct in a defined oligonucleotide (Figs 2 and 3). Previously this enzyme was also found to act on the ϵ A adduct (7,8,10), a structural analog of EA, although these two adducts are produced by completely unrelated compounds. Human APNG, as well as homologs in cells from eukaryotic and prokaryotic species, represents a family of enzymes with a wide substrate range (for a review see 26). This work showed that the substrate range of APNG is still expanding.

Human APNG excises ϵ A from DNA with high efficiency (8,9). We previously reported (9) that ϵ A is even preferred by

Figure 10. Evolution over time of the three hydrogen bond distances in the APNG binding pocket for the ϵ A- (green) and EA-containing (blue) DNA/APNG complexes. The ϵ A N9–His136 NH and EA N9–His136 NH hydrogen bonds remained intact during the entire simulation.

APNG over 3-methyladenine, after which the enzyme was originally named. The kinetic comparison made in this work between ϵ A and EA showed that ϵ A is excised much faster

than EA (Fig. 2). Such biochemical data prompted us to explore the structural basis for the observed difference.

In this work we employed molecular dynamics simulation to provide structural insights on the EA- and ϵ A-containing 25mer DNA duplexes and the effect of EA on the APNG active site conformation. Molecular modeling did not reveal any significant conformational features which can distinguish between the ϵ A and EA adducts when incorporated opposite T in 25mer DNA duplexes. Both duplexes have similar structural motifs around the lesion sites. Both adducts adopted the *anti* orientation, were displaced towards the major groove and formed a non-planar, sheared base pair with the opposite T. It has been proposed that sheared base pairs can be a structural feature important for recognition by some DNA glycosylases (27). No hydrogen bonds were observed between the bases in the ϵ A-T and EA-T pairs. The sugar pucker of the EA and ϵ A adducts falls in the C2'-*endo*/C1'-*exo* range. The smaller twist values observed for both lesions should contribute to unwinding of the DNA upon binding to the repair enzyme. The unwound DNA around the lesion site allows easy access for the repair enzyme to continue further adduct recognition and discrimination (28). The overall conformation of the ϵ A-T base pair produced by modeling was in general agreement with NMR data on an ϵ A-T-containing 9mer duplex (23).

Saturation of the imidazole ring in the EA adduct partially reduced the stacking ability of this molecule, as compared to ϵ A, which favors π - π stacking interactions with amino acids in the enzyme active site. The extra, non-planar hydrogens at the C7 and C8 positions of EA, as compared to ϵ A, contribute an additional van der Waals surface area that makes it more difficult to accommodate the adduct between the conformationally constrained Tyr127 and more flexible His136. The replacement of ϵ A by EA in the APNG active site resulted in an edge-to-edge packing interaction between His136 and EA. The conformation produced by molecular modeling shows that in order to accommodate EA in the enzyme active site the active site required a structural rearrangement involving His136. A comparison of the APNG crystal structure with ϵ A-DNA/APNG and abasic pyr-DNA/APNG complexes showed that the Tyr127, Tyr157 and His136 side chains are in the same orientation, suggesting that the conformation of the APNG active site is predetermined and not influenced by adduct binding (29). The extra energy required to overcome the steric clash between the aromatic side chain of His136 and the 7,8-dihydro-imidazole ring of EA should prevent an easy fit of EA adducts into the APNG active site. However, the mechanism of EA excision by APNG may be similar to that reported for ϵ A (11,29). Both ϵ A and EA have a lone pair acceptor nitrogen (N9), which is unique to the alkylated base. The position of Glu125 is not changed in the presence of EA and this residue should be able to deprotonate the active site bound water for nucleophilic attack on the C1' sugar carbon of EA. The hydroxide nucleophile will be stabilized by Arg182, the position of which also remains unchanged in the EA-DNA/APNG complex as compared to the ϵ A-DNA/APNG complex.

Based on the conformations of the ϵ A- and EA-containing duplexes, it can be suggested that the glycosylase does not distinguish between these adducts based on local DNA distortion. Similar structural motifs for these adducts serve as an initial signal for the enzyme to test the base by forcing

Tyr162 into the helix and displacing the modified base into the enzyme active pocket. The enzyme active pocket requires tight interaction between the adduct and the neighboring amino acids and thus is sensitive to the adduct structure and conformation. It was shown that an APNG mutant (H136Q), engineered to eliminate aromatic stacking interactions with ϵ A, has very low repair efficiency (11,29). Moreover, it has been proposed that base stacking interactions between the damaged bases and the aromatic side chains of amino acids in the active site may provide a basis for recognition and excision by *E. coli* 3-methyladenine DNA glycosylase II (30,31), which also excises EA (B.Hang, A.B.Guliaev and B.Singer, manuscript in preparation). The observed destabilization of the EA adduct in the active site, such as the weaker stacking interaction of the adduct with the aromatic side chains of His136, is likely to contribute to the lower efficiency of repair and explain why this adduct is a less preferable substrate than ϵ A for human APNG.

ACKNOWLEDGEMENTS

We would like to thank the staff of the National Energy Research Scientific Computing Center (NERSC), Lawrence Berkeley National Laboratory, especially David Skinner, for help in setting up AMBER 6.0 on the IBM SP RS/6000 supercomputer at NERSC for our calculations. This work was supported by NIH grants CA 47723 (to B.S.) and CA 72079 (to B.H.) and was administered by the Lawrence Berkeley National Laboratory under Department of Energy contract DE-AC03-76SF00098.

REFERENCES

1. Ludlum,D.B. (1986) Chapter 2. In Singer,B. and Bartsch,H. (eds), *The Role of Cyclic Nucleic Acid Adducts in Carcinogenesis and Mutagenesis*. IARC Scientific Publications no. 70. IARC, Lyon, France, pp. 137-146.
2. Ludlum,D.B. (1990) DNA alkylation by the haloethylnitrosoureas: nature of modifications produced and their enzymatic repair or removal. *Mutat. Res.*, **223**, 117-126.
3. Bartsch,H. (1999) Keynote address. In Singer,B. and Bartsch,H. (eds), *Exocyclic DNA Adducts in Carcinogenesis and Mutagenesis*. IARC Scientific Publications no. 150. IARC, Lyon, France, pp. 1-16.
4. Nair,J., Barbin,A., Velic,I. and Bartsch,H. (1999) Etheno DNA-base adducts from endogenous reactive species. *Mutat. Res.*, **424**, 59-69.
5. El Ghissassi,F., Barbin,A., Nair,J. and Bartsch,H. (1995) Formation of 1,N⁶-ethenoadenine and 3,N⁴-ethenocytosine by lipid peroxidation products and nucleic acid bases. *Chem. Res. Toxicol.*, **8**, 278-283.
6. Singer,B. and Bartsch,H. (eds) (1999) *Exocyclic DNA Adducts in Carcinogenesis and Mutagenesis*. IARC Scientific Publications no. 150. IARC, Lyon, France.
7. Hang,B., Singer,B., Margison,G.P. and Elder,R.H. (1997) Targeted deletion of alkylpurine-DNA-N-glycosylase in mice eliminates repair of 1,N⁶-ethenoadenine and hypoxanthine but not of 3,N⁴-ethenocytosine or 8-oxoguanine. *Proc. Natl Acad. Sci. USA*, **94**, 12869-12874.
8. Saparbaev,M., Kleibl,K. and Laval,J. (1995) *Escherichia coli*, *Saccharomyces cerevisiae*, rat and human 3-methyladenine DNA glycosylases repair 1,N⁶-ethenoadenine when present in DNA. *Nucleic Acids Res.*, **23**, 3750-3755.
9. Dosanjh,M.K., Roy,R., Mitra,S. and Singer,B. (1994) 1,N⁶-ethenoadenine is preferred over 3-methyladenine as substrate by a cloned human N-methylpurine DNA glycosylase (3-methyladenine DNA glycosylase). *Biochemistry*, **33**, 1624-1628.
10. Singer,B., Antoccia,A., Basu,A.K., Dosanjh,M.K., Fraenkel-Conrat,H., Gallagher,P.E., Kusmierek,J.T., Qiu,Z.H. and Rydberg,B. (1992) Both purified human 1,N⁶-ethenoadenine-binding protein and purified human

- 3-methyladenine-DNA glycosylase act on 1,N⁶-ethenoadenine and 3-methyladenine. *Proc. Natl Acad. Sci. USA*, **89**, 9386–9390.
11. Lau,A.Y., Wyatt,M.D., Glassner,B.J., Samson,L.D. and Ellenberger,T. (2000) Molecular basis for discriminating between normal and damaged bases by the human alkyladenine glycosylase, AAG. *Proc. Natl Acad. Sci. USA*, **97**, 13573–13578.
 12. Hang,B., Downing,G., Guliaev,A. and Singer,B. (2002) Novel activity of *Escherichia coli* mismatch uracil-DNA glycosylase (Mug) excising 8-(hydroxymethyl)-3,N⁴-ethenocytosine, a potential product resulting from glycidaldehyde reaction. *Biochemistry*, **41**, 2158–2165.
 13. Maruenda,H., Chenna,A., Liem,L.K. and Singer,B. (1998) Synthesis of 1,N⁶-ethano-2'-deoxyadenosine, a metabolic product of 1,3-bis(2-chloroethyl)nitrosourea, and its incorporation into oligomeric DNA. *J. Org. Chem.*, **63**, 4385–4389.
 14. Hang,B., Sagi,J. and Singer,B. (1998) Correlation between sequence-dependent glycosylase repair and the thermal stability of oligonucleotide duplexes containing 1,N⁶-ethenoadenine. *J. Biol. Chem.*, **273**, 33406–33413.
 15. Rydberg,B., Dosanjh,M.K. and Singer,B. (1991) Human cells contain protein specifically binding to a single 1,N⁶-ethenoadenine in a DNA fragment. *Proc. Natl Acad. Sci. USA*, **88**, 6839–6842.
 16. O'Connor,T.R. (1993) Purification and characterization of human 3-methyladenine-DNA glycosylase. *Nucleic Acids Res.*, **21**, 5561–5569.
 17. Guliaev,A.B., Sagi,J. and Singer,B. (2000) Sequence-dependent conformational perturbation in DNA duplexes containing an ϵ A-T mismatch using molecular dynamics simulation. *Carcinogenesis*, **21**, 1727–1736.
 18. Case,D.A., Pearlman,D.A., Caldwell,J.W., Cheatham,T.E., Ross,W.S., Simmerling,C.L., Darden,T.A., Merz,K.M., Stanton,R.V., Cheng,A.L., Vincent,J.J., Crowley,M., Tsui,V., Radmer,R.J., Duan,J., Pitner,I., Massova,G.L., Seibel,G.L., Singh,U.C., Weiner,P.K. and Kolman,P.A. (1999) *AMBER 6*. University of California, San Francisco, CA.
 19. Humphrey,W., Dalke,A. and Schulten,K. (1996) VMD—visual molecular dynamics. *J. Mol. Graphics*, **14**, 33–38.
 20. Lavery,R. and Sklenar,H. (1996) *CURVES5.1. Helical Analysis of Irregular Nucleic Acids*. Laboratory for Theoretical Biochemistry, CNRS, Paris.
 21. Habraken,Y., Carter,C.A., Sekiguchi,M. and Ludlum,D.B. (1991) Release of N²,3-ethanoguanine from haloethylnitrosourea-treated DNA by *Escherichia coli* 3-methyladenine DNA glycosylase II. *Carcinogenesis*, **12**, 1971–1974.
 22. Santos,C.D., Kouchakdjian,M., Yarema,K., Basu,A., Essigmann,J. and Patel,D.J. (1991) NMR studies of the exocyclic 1,N⁶-ethenodeoxyadenosine adduct (ϵ dA) opposite deoxyguanosine in a DNA duplex— ϵ dA(Syn)-dG(Anti) pairing at the lesion site. *Biochemistry*, **30**, 1828–1835.
 23. Kouchakdjian,M., Eisenberg,M., Yarema,K., Basu,A., Essigmann,J. and Patel,D.J. (1991) NMR studies of the exocyclic 1,N⁶-ethenodeoxyadenosine adduct (ϵ dA) opposite thymidine in a DNA duplex—nonplanar alignment of ϵ dA(Anti) and dT(Anti) at the lesion site. *Biochemistry*, **30**, 1820–1828.
 24. Parikh,S.S., Mol,C.D. and Tainer,J.A. (1997) Base excision repair enzyme family portrait: integrating the structure and chemistry of an entire DNA repair pathway. *Structure*, **5**, 1543–1550.
 25. Mol,C.D., Parikh,S.S., Putnam,C.D., Lo,T.P. and Tainer,J.A. (1999) DNA repair mechanisms for the recognition and removal of damaged DNA bases. *Annu. Rev. Biophys. Biomol. Struct.*, **28**, 101–128.
 26. Singer,B. and Hang,B. (1997) What structural features determine repair enzyme specificity and mechanism in chemically modified DNA? *Chem. Res. Toxicol.*, **10**, 713–732.
 27. Cullinan,D., Johnson,F. and de los Santos,C. (2000) Solution structure of an 11-mer duplex containing the 3,N⁴-ethenocytosine adduct opposite 2'-deoxycytidine: implications for the recognition of exocyclic lesions by DNA glycosylases. *J. Mol. Biol.*, **296**, 851–861.
 28. Zou,Y., Luo,C. and Geacintov,N.E. (2001) Hierarchy of DNA damage recognition in *Escherichia coli* nucleotide excision repair. *Biochemistry*, **40**, 2923–2931.
 29. Lau,A.Y., Scharer,O.D., Samson,L., Verdine,G.L. and Ellenberger,E. (1998) Crystal structure of a human alkylbase-DNA repair enzyme complexed to DNA: mechanisms for nucleotide flipping and base excision. *Cell*, **95**, 249–258.
 30. Yamagata,Y., Kato,M., Odawara,K., Tokuno,Y., Nakashima,Y., Matsushima,N., Yasumura,K., Tomita,K., Ihara,K., Fujii,Y., Nakabeppu,Y., Sekiguchi,M. and Fujii,S. (1996) Three-dimensional structure of a DNA repair enzyme, 3-methyladenine DNA glycosylase II, from *Escherichia coli*. *Cell*, **86**, 311–319.
 31. Labahn,J., Scharer,O.D., Long,A., Ezaznikpay,K., Verdine,G.L. and Ellenberger,T.E. (1996) Structural basis for the excision repair of alkylation-damaged DNA. *Cell*, **86**, 321–329.