Published online 25 November 2011

Nucleic Acids Research, 2012, Vol. 40, No. 6 ***2759–2770***
doi:10.1093/nar/gkr1094

Structure and stability of DNA containing an aristolactam II-dA lesion: implications for the NER recognition of bulky adducts

Mark Lukin, Tanya Zaliznyak, Francis Johnson and Carlos de los Santos*

Department of Pharmacological Sciences, School of Medicine, Stony Brook University, Stony Brook, NY 11794-8651, USA

Received September 12, 2011; Revised November 2, 2011; Accepted November 3, 2011

ABSTRACT

Aristolochic acids I and II are prevalent plant toxicants found in the *Aristolochiaceae* plant family. Metabolic activation of the aristolochic acids leads to the formation of a cyclic *N*-hydroxylactam product that can react with the peripheral amino group of purine bases generating bulky DNA adducts. These lesions are mutagenic and established human carcinogens. Interestingly, although AL-dG adducts progressively disappear from the DNA of laboratory animals, AL-dA lesions has lasting persistence in the genome. We describe here NMR structural studies of an undecameric duplex damaged at its center by the presence of an ALII-dA adduct. Our data establish a locally perturbed double helical structure that accommodates the bulky adduct by displacing the counter residue into the major groove and stacking the ALII moiety between flanking bases. The presence of the ALII-dA perturbs the conformation of the 5′-side flanking base pair, but all other pairs of the duplex adopt standard conformations. Thermodynamic studies reveal that the lesion slightly decreases the energy of duplex formation in a sequence-dependent manner. We discuss our results in terms of its implications for the repair of ALII-dA adducts in mammalian cells.

INTRODUCTION

Aristolochic acids (AAs) are nitrophenanthrene carboxylic acid derivatives (Figure 1) commonly found in plants of the *Aristolochiaceae* family. These compounds are potent human mutagens (1) and established human carcinogens (2). Chronic dietary poisoning by AA leads to terminal renal failure and is associated with transitional cell carcinoma of the upper urinary tract (3–5). *In vivo*, AA can undergo enzymatic reduction of its nitro group to the corresponding aristolactam (Figure 1), a reaction that is mainly catalyzed by the cytosolic NAD(P)H:quinone oxidoreductase (NQO1) or the microsomal enzymes CYP1A2 and CYP1A2 (6). The cyclic hydroxamic acid **2** (*N*-hydroxyaristolactam) formed in this multistep reaction can lead, without further activation by phase II enzymes, to the formation of a carbenium ion **3** that is capable of attacking cellular DNA and generating purine adducts **4** (Figure 1) (7,8). In contrast to other nitroaromatic arenes that preferentially react with the C-8 position of guanine, activated AA forms adducts with the exocyclic amino groups of purine bases, namely AL-N^6-dA and AL-N^2-dG (Figure 1) (9). Remarkably, whereas the latter adduct gradually disappears from laboratory animals after discontinuation of AA administration, AL-dA exhibits life-long persistence in rat genomic DNA (10) and stays for years in human DNA within the renal cortex (4).

The ALII-N^6-dA adduct is highly mutagenic predominantly leading to A → T transversions (11) that are rarely observed in human tumors (12). Similarly, this type of transversion was shown to be a 'signature' of AA-induced mutations in humans exposed to low levels of AA in their diet (13). Based on its unique mutational signature, a differential rate of AL-dA repair has been measured for the transcribed and non-transcribed strands of genomic DNA, suggesting that the transcription-coupled repair (TCR) pathway of the nucleotide excision repair (NER) system, but not the global genomic repair (GGR) pathway, removes AL-dA adducts from the genome (13).

To understand the molecular mechanisms responsible for the mutagenic properties of AL-dA lesions and their resistance to GGR, it is necessary to establish the structural and physicochemical properties of adduct-containing duplexes. Our recent progress in the chemical synthesis of the ALII adducts and their incorporation into the oligonucleotides (11) now makes possible the preparation of

*To whom correspondence should be addressed. Tel: +1 631 444 3649; Fax: +1 313 444 3218; Email: cds@pharm.stonybrook.edu

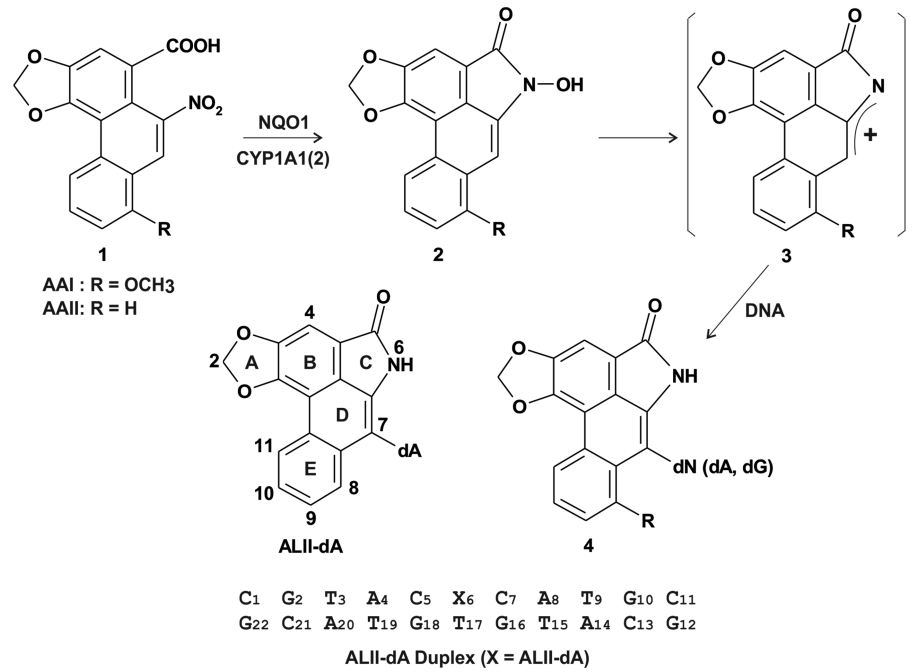

Figure 1. Schematic representation of the metabolic activation of aristolochic acids and formation of DNA adducts. The structure of the arisotlactam-II-(N^6)-dA lesion indicates the position of ALII atoms with Arabian digits and labels the aristolactam rings with upper case letters. Sequence of the ALII-dA duplex characterized in the present study.

large amounts of damaged duplexes, thus opening an avenue for such studies. In this article, we describe the NMR structure determination of an undecamer duplex containing the dA-N^6-adduct of aristolactam II (ALII-dA), the most abundant and persistent DNA lesion generated by AAII (1). Our data reveal that the ALII-dA adduct causes local structural perturbations at the site of damage by intercalating the ALII moiety between the lesion-flanking guanine bases on the complementary strand and displacing the lesion partner nucleotide into the major groove of the duplex. We also report the thermodynamic parameters of three ALII-dA duplexes, which, in combination with our structural data and the repair studies in mammalian cells (V. Sidorenko, *et al.*, companion manuscript), challenge the current paradigm that a lesion-induced disruption of the normal DNA structure causing a decrease of local duplex stability elicits NER recognition.

EXPERIMENTAL SECTION

Sample preparation

Of the three ALII-dA lesion-containing duplexes of different composition that were prepared for this work, the d[(CGTACXCATGC)•(GCATGTGTACG)] duplex (termed hereafter the ALII-dA duplex), where X is the damaged residue, was the one chosen for full structure/stability studies. Two other duplexes having the lesion in the context of mutational hotspots (13), namely d[(CTCCTXGGTTG)•(CAACCTAGGAG)] (TXG duplex) and d[(ATGACXGAAAC)•(GTTTCTGTCAT)] (CXG duplex), were also synthesized and partially characterized. Preparation of oligonucleotides containing the ALII-dA lesion has been reported elsewhere (11,14). Non-modified oligonucleotides were prepared following standard solid-phase synthesis methods and purified by ion-pair reverse-phase HPLC (15). The quality and composition of the oligonucleotides were assessed by ESI–MS. The purified samples were dissolved in 1 M NaCl solution containing 1 mM EDTA and desalted using the Sephadex G-25 column (2.5×80 cm). A 1 : 1 strand ratio on the duplex was obtained by monitoring the NMR signal intensity of resolved thymine methyl protons during the gradual addition of the unmodified to the lesion-containing oligomer. One milliliter of a 25-mM phosphate buffer, pH 6.5, containing 50 mM NaCl was added to the duplex solution and lyophilized to dryness. The powder was dissolved in 1 ml of 99.5% D_2O, lyophilized, and dissolved again in 0.6 ml of '100%' D_2O (Aldrich). For NMR experiments recorded in 10% D_2O buffer, the sample was lyophilized and dissolved in a 1:9 mixture of D_2O : MilliQ H_2O.

NMR experiments

The ALII-dA duplex concentration used for the NMR studies was ~1 mM. NMR spectra were collected on a Bruker Avance spectrometers operating at 700, 800 and 900 MHz. Proton chemical shifts were referenced relative to sodium 3-(trimethylsilyl)propionate-2,2,3,3-d_4 at 0 ppm. Phase-sensitive NOESY (90, 150, 200 and 300 ms mixing time), COSY, TOCSY (30 and 70 ms isotropic mixing time) and C-H-HSQC spectra in D_2O buffer were recorded at 27°C. The residual water signal was suppressed by presaturation during the relaxation delay of

1.5 s. Phase-sensitive proton NOESY spectra (200 and 300 ms mixing time) in 10% D_2O buffer were collected at 5°C, using the 'excitation sculpting' pulse sequence (16). The temperature dependence of imino protons was studied at 700 MHz from 5°C to 60°C using the 'jump-return' pulse sequence. NMR data were processed and analyzed using NMRPipe (17) or Felix (Accelrys Inc., San Diego, CA, USA). Two-dimensional data sets consisted of 2048 and 380 complex points in the t2 and t1 dimensions, respectively. Shifted sine-bell window functions were used to smooth the time domain data prior to the Fourier transformation. Polynomial baseline correction was applied to the frequency domain spectra as needed.

Molecular dynamics

Restrained molecular dynamics (rMD) calculations were performed with Xplor-NIH (18) using an all-atom force field derived from CHARMM (19) and a dielectric constant set to four (20). Partial electric charges of the aristolactam moiety were calculated using the GAUSSIAN module of the HyperChem program package (HyperCube Inc.) in the minimal (3–21 G) basis set. A starting B-form duplex structure was built using the Biopolymer module of Insight 2000 (Accelrys Inc.). The central A•T base pair of this duplex was modified by attaching the ALII moiety to the adenine N^6 atom, displacing the T nucleotide toward the major groove and partially intercalating the aromatic ring system between G16 and G18 residues of the complementary strand. The initial model was energy minimized by 1000 steps of conjugated gradient minimization to remove the unfavorable atomic interactions.

Experimental interproton distances were calculated using a full relaxation-matrix approach (21). Briefly, hydrogen atom positions on the initial structure were energy minimized, using a sole potential energy function that was proportional to the difference between back-calculated and experimental NOE intensities. Interproton distances were then computed at the end of the minimization. Distance-refined structures were calculated using an rMD protocol similar to that employed for α-OH-PdG•dC duplexes (22). Briefly, the simulations were initiated at five different temperatures (100, 105, 110, 115 and 120 K) and the system was heated to 500 K on 100 ps. During this time, potential energy penalty function enforcing interproton distances were introduced with a penalty constant that increased from 10 to 300 kcal/(mol A^2) and remained at this value until the end of the simulation. Simulations were run at 500 K for 100–120 ps. After this time, the system was cooled down to 300 K in 100 ps and run at this temperature for an additional 130 ps. A total of 25 independent structures of the dA-ALII duplex were computed by starting the simulations at five different temperatures (100, 105, 110, 115 and 120 K) and running five different time periods (100, 105, 110, 115 and 120 ps) of high-temperature step. Atom coordinates from the last 30 ps of each simulation were averaged and energy minimized, yielding a set of refined models that exhibited pairwise root-mean-square deviations (RMSDs) of <0.7 Å. Given their convergence, the models were averaged and energy minimized generating final refined structure presented here. Structures were analyzed with curves (23) and visualized with Chimera (24) or InsightII (Accelrys Inc.).

UV melting studies

Thermal denaturations of the lesion-containing and control duplexes were carried out using a CARY100 Bio UV–vis spectrophotometer, equipped with a multicell block temperature regulation unit and a fluid circulation thermal regulation enhancement (Varian, Inc.). Temperature readings were stable and accurate within 1°C. Initial temperatures, 10°C or 80°C depending upon the experiment, were allowed to equilibrate for at least 10 min. The rate of temperature change was set to 0.2°C/min. Sample concentrations varied between 0.2 and 2.2 OD_{260} units of duplex dissolved in 1 ml of 25 mM sodium phosphate buffer solution, pH 6.8, containing 150 mM NaCl and 0.5 mM EDTA. Duplex melting temperatures (T_m) were computed from the first derivative of the Abs versus temperature plots. Three independent melting profiles were obtained at each of 12 different sample concentrations, resulting in a set of 36 independent T_m values for each duplex. Plots of $(1/T_m)$ versus $\ln(C_t)$, where C_t is the duplex molar concentration, were fit to a straight line, and the ΔH^o and ΔS^o then derived from the slope and intercept, respectively (25). The Gibbs free energy (ΔG^o) of duplex formation was calculated from these values.

RESULTS

Nonexchangeable proton spectra

The assignment of non-exchangeable protons of the duplex resulted from the analysis of NOESY, COSY and TOCSY spectra following established procedures (26). Figure 2 shows an expanded region of a 300-ms mixing time NOESY spectrum illustrating the interactions between the base (6.5–8.5 ppm) and sugar-H1′ (4.5–6.6 ppm) protons of the duplex. The most striking feature of this spectrum is that the ALII-dA lesion has virtually no effect on NOE interactions between the sugar and aromatic protons of the modified strand. Each purine-H8 or pyrimidine-H6 of this strand shows NOE cross-peaks to the H1′ proton of the same and 5′-flanking nucleotides (Figure 2, top), with relative intensities and chemical shift values within the limits characteristic of a normal DNA duplexes. Proton signals from the damaged adenine residue experience a slight chemical shift increase (~0.5 ppm) that, however, is not accompanied by any appreciable intensity change of its NOE interactions with the protons of adjacent nucleotides. All these facts suggest that the duplex conserves a right-handed helical conformation throughout the modified strand, with no appreciable perturbations caused by the presence of the lesion. By contrast, although the normal aromatic to sugar-H1′ NOE walk is essentially preserved in the complementary strand, the

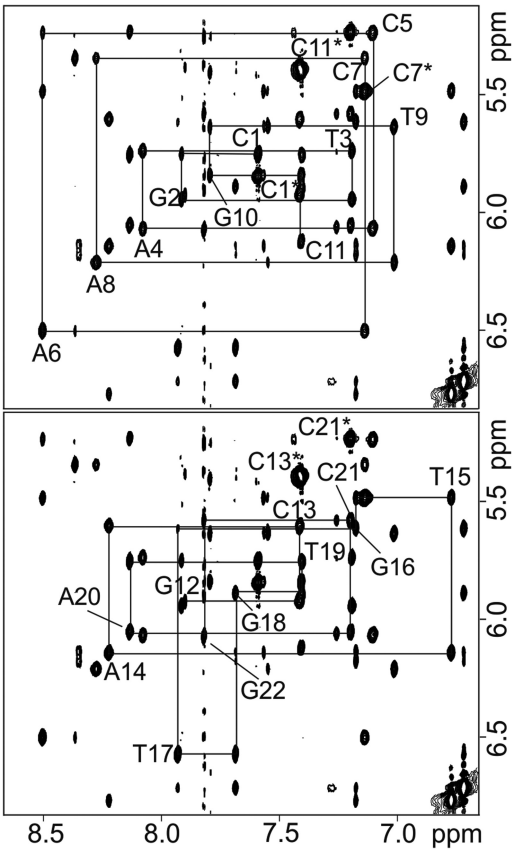

Figure 2. Expanded region of an 800 MHz NOESY (300-ms mixing time) spectrum of the ALII-dA duplex recorded at 25°C with the sample dissolved in 25 mM phosphate buffer, pH 6.6, 50 mM NaCl, in 100% D_2O. The figure shows sequential interactions between the base (6.6–8.6 ppm) and the sugar H1′ (5.1–6.8 ppm) protons, from C1 through C11 on the lesion containing strand (top panel) and from G12 through G22 on the complementary strand (bottom panel). Nucleotide labels identify the intraresidue NOE peaks and asterisks indicate cytosine H6–H5 interactions.

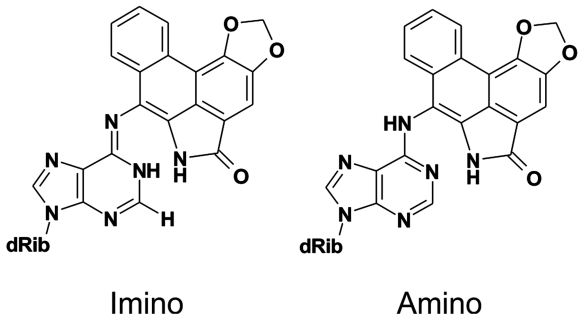

Figure 3. Tautomeric forms of the ALII-dA adduct (the *syn*, *syn*-conformation is shown).

chemical shifts of both the anomeric and the aromatic protons of the T17 nucleotide experience a significant low-field shift (~1 ppm), indicating that the ALII adduct disrupts the location of the lesion-partner nucleotide (Figure 2, bottom). In addition, although T17 exhibits internucleotide NOE interactions typically observed in double-stranded DNA, the intensity of some cross-peaks is altered, indicating that some proton distances are significantly changed opposite to the ALII-dA lesion.

Assignment of the non-exchangeable protons of the artistolactam moiety followed the analysis of COSY, NOESY and natural abundance ^{13}C,^{1}H-HSQC spectra. The aristolactam methylene protons are readily distinguished on the ^{13}C,^{1}H-HSQC spectrum due to the fact that, in the anisotropic DNA environment, these protons have close but clearly different chemical shifts. Protons of ring **E** exhibit a characteristic COSY cross-peak pattern that allows tracing their connectivity, despite the fact that H8 and H10 chemical shifts are close (Supplementary Figure S1, peaks A–C). Of these protons, H11 is easily distinguished by its NOE contact with ALII H2s (Supplementary Figure S2, peaks B) leading, in turn, to the assignment of the remaining ring **E** protons. The lack of scalar spin–spin interactions with other protons and the presence of NOE contacts with the methylene group (Supplementary Figure S2, peaks G) identified the H4 proton of ALII moiety.

The ALII protons exhibited numerous NOE interactions with protons of adjacent nucleotides, which established its location with high confidence. The aristolactam methylene group (H2) and the aromatic H4 proton display strong NOE interactions with base and sugar protons of nucleotides on the complementary strand, specifically G16, T17 and G18 (Supplementary Figure S3), indicating that the C2–C5 edge of the aristolactam moiety is deeply buried within the DNA base stack. By contrast, ring **E** protons H8–H11 exhibit a low number of NOE cross-peaks of weak intensity (Supplementary Figure S2), suggesting that they are more exposed to solvent in the major groove of the duplex. These data allow us to conclude that the damaged 2′-deoxyadenosine adopts an *anti*-conformation and is properly stacked within the DNA double helix. The aristolactam moiety is directed toward the complementary strand where it stacks between G16 and G18, displacing the T17 nucleotide from its normal position. Given that steric interactions prohibit an *anti* conformation of the ALII-dA moiety around the adenine C6–N6 bond, the only possible conformational isomerism results from rotation around the adenine(N6)–ALII(C7) bond. However, the experimental NOE cross-peak pattern indicates a unique *syn, syn*-conformation for the ALII-dA residue in duplex DNA (Figure 3). The chemical shifts of the nonexchangeable protons on the ALII-dA duplex are listed in Supplementary Table S1.

Conformation of the T17 nucleotide

The location of the ALII moiety, stacked between residues of the undamaged strand, and the chemical shifts of T17-H6 and T17-H1′ suggest that this nucleotide is displaced from its normal position in duplex DNA. The relatively weak NOE interactions between T17 and adjacent nucleotides, especially G16, further indicate that the conformation of the former nucleotide is drastically perturbed in the ALII-dA duplex. As a result, its position relative to neighboring residues and the conformation of its sugar–phosphate backbone become less

well defined. We realized that NOE interactions involving the H5′ and H5″ protons of T17 could provide additional experimental restraints for establishing its conformation with greater confidence. In order to obtain stereospecific assignments of the T17-H5′ and T17-H5″ resonances and to use their NOE distance information fully, we employed our recently proposed isotope-labeling strategy (14) to prepare an ALII-dA duplex having T17-H5″ selectively deuterated. Comparison of the NOESY spectra of the deuterated and nondeuterated duplexes (Supplementary Figure S3) yielded the unambiguous assignment of both diastereotopic 5′ protons and allowed the computation of five additional NOE restraints, including two key distances involving the ALII-H4 proton. As a result, rMD defined the conformation of T17 with greater accuracy.

Exchangeable proton spectra

The imino proton region (13.8–10.5 ppm) of the 1D spectrum recorded at 5°C in 10% D_2O buffer solution, pH 6.5, shows 10 well-resolved signals that account for the imino protons of the ALII-dA duplex (Figure 4, top). A NOESY spectrum recorded under the same experimental conditions displays a characteristic pattern of interactions that permits the assignment of the exchangeable protons of the duplex (Figure 4, bottom). Thymine imino protons exhibit strong NOE peaks from their interaction with H2 protons of complementary adenines (Figure 4, peaks A–D), whereas the imino protons of guanine nucleotides interact with both the hydrogen-bonded and non-hydrogen-bonded protons of complementary cytosines (Figure 4, peaks E–H). In addition, the imino protons of adjacent base pairs exhibit sequential NOE interactions (Supplementary Figure S4) indicating proper base pair stacking throughout the duplex. These observations establish the formation of A•T and C•G Watson–Crick alignments for all nondamaged base pairs with proper stacking throughout the ALII-dA duplex. The proton signal at 10.9 ppm displays a strong NOE cross-peak with the H2 proton of the damaged adenine base (Figure 4, peak I) and with the imino protons of G16 and G18, the lesion adjacent guanines (Supplementary Figure S4, peaks F and J, respectively), thereby demonstrating that it belongs to the base pair containing the lesion. We discuss its assignment below.

Exchangeable protons of the lesion-containing base pair

The central ALII-dA•dT base pair has three exchangeable protons that, in principle, might produce signals in this region of the spectrum, namely, the thymine imino proton, the H6 proton of the ALII moiety and, based on an early report (27), the dA(H1) proton of the damaged nucleotide in the imino tautomeric form (Figure 3, left). The absence of an intraresidue NOE peak with the thymine methyl protons (Figure 4, middle panel) immediately rules out the possibility that the 10.9 ppm resonance belongs to T17(H3). This observation is not surprising, since the extrahelical orientation of the T17 will certainly facilitate the fast water exchange of its imino proton. Regarding the alternative possibilities, both

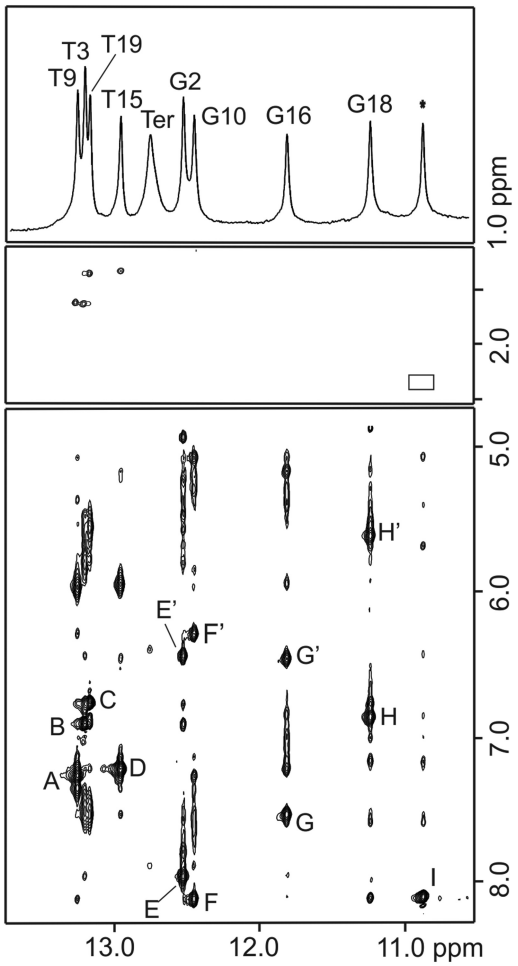

Figure 4. The top panel displays the imino proton region of a 700 MHz 1D spectrum of the ALII-dA duplex recorder at 5°C with the sample dissolved in 25 mM phosphate buffer, pH 6.6, 50 mM NaCl, in 10% D_2O. The bottom and center panels show expanded regions of a 900 MHz NOESY (300-ms mixing time) spectrum of the ALII-dA duplex recorded under the same experimental conditions. Labeled peaks are assigned as follows: A, T9(N3H)-A14(H2); B, T3(N3H)-A20(H2); C, T19(N3H)-A4(H2); D, T15(N3H)-A8(H2); E, G2(N1H)-C21(N4H)b; E′, G2(N1H)-C21(N4H)nb; F, G10(N1H)-C13(N4H)b; F′, G10(N1H)-C13(N4H)nb; G, G16(N1H)-C7(N4H)b; G′, G16(N1H)-C7(N4h)nb; H, G18(N1H)-C5(N4H)b; H′, G18(N1H)-C5(N4H)nb. ('b' and 'nb' indicate bonded and non-hydrogen-bonded amino protons, respectively). The asterisk labels the imino proton signal of the damaged base pair and its NOE interactions are depicted in Figure 5. The empty box indicates the region of the spectrum where a cross-peak should appear had the resonance denoted with the asterisk belongs to T17(H3). 'Ter' labels the overlapping imino protons of the terminal guanine bases.

the ALII(H6) and dA(H1) protons are expected to resonate within this region and produce an intense NOE cross-peak with dA(H2), similar to the one observed experimentally (Figures 4 and 5, peak I). Further analysis of the water NOESY spectrum revealed a resonance at 7.5 ppm, which was missing from the NOESY spectra in D_2O, indicating that it originates from an exchangeable proton of the duplex. This proton displays strong NOE cross-peaks with ALII(H8) and ALII(H9) (Figure 5, peaks 8 and 9, respectively), indicating they are very close in

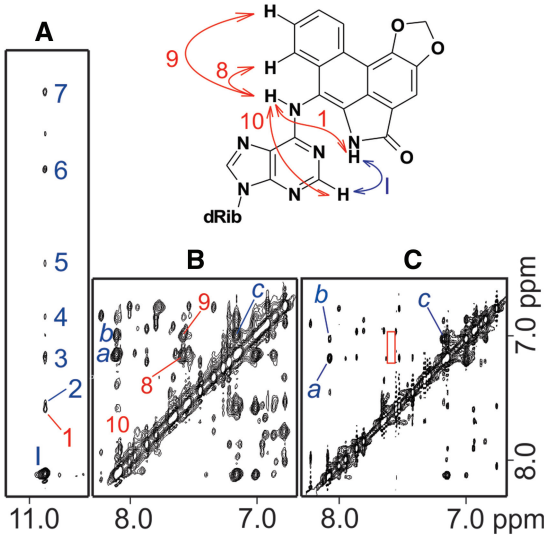

Figure 5. Regions of NOESY spectra used for the assignment of exchangeable protons at the lesion site. (**A**) Displays a segment of the NOESY spectrum recorded in 10% D_2O buffer (also shown on Figure 4) depicting interactions of the 10.9 ppm signal. (**B**) Shows the symmetrical aromatic region of the same NOESY spectrum. (**C**) Shows the same aromatic region on a NOESY spectrum recorded in 100% D_2O buffer. Blue letters or numbers denote NOE cross peaks involving the ALII(N6H) imino proton and red numbers indicate those involving the ALII-dA(H6) amino proton. Labeled peaks are assigned as follows: I, ALII(N6H)-A6(H2); 1, ALII(N6H)-A6(N6H); 2, ALII(N6H)-C7(N4H)b; 3, ALII(N6H)-ALII(H8); 4, ALII(N6H)-C5(N4H)b; 5, ALII(N6H)-ALII(H4); 6, ALII(N6H)-G18(H1′); 7, ALII(N6H)-C7(H1′); 8, A6(N6H)-ALII(H8); 9, A6(N6H)-ALII(H9); 10, A6(H2)-A6(N6H); a, ALII(H11)-ALII(H8); b, ALII(H11)-ALII(H9); c, ALII(H8)-ALII(H9). The red rectangle in (C) indicates the A6(N6H)-ALII(H8) and A6(N6H)-ALII(H9) interactions that are absent in 100% D_2O.

space, and also weak interactions with ALII(H6) and dA(H2) (Figure 5, peaks 1 and 10, respectively), thus establishing a separation of at least 4 Å between them. These observations forced us to discard the previously reported imino tautomeric form of ALII-dA (Figure 3, left) (27) and we conclude that, in double-stranded DNA, it adopts the more common amino form (Figure 3, right). Accordingly, the resonance at 7.5 ppm belongs to dA(H6) and that at 10.9 ppm to the H6 imino proton of the ALII moiety. Analysis of the cross-peaks pattern observed between these protons and the surrounding exchangeable and nonexchangeable protons (Figure 5, left panel) is in full agreement and confirms these assignments. Exchangeable proton chemical shifts of the ALII-dA duplex are listed in Supplementary Table S1.

Structure of the ALII-dA duplex

The sharp and uniform shape of the proton signals observed during the NMR characterization of the ALII-dA duplex demonstrates that the latter exists as a single, well-defined structure without experiencing any appreciable conformational exchange on the NMR timescale. Under this condition, the set of experimental NOE-derived distance restraints can drive MD simulations to a low-energy model that accurately represents the structure of the damaged duplex in solution. All of the refined structures are in excellent agreement with the experimental NMR constraints, having no interproton distance violations >0.1 Å. The refined models are very similar (Supplementary Figure S5) showing a heavy atom pairwise RMSD in the range of 0.18–0.67 Å for the whole duplex, of 0.09–0.80 Å for the three central base pairs and of 0.13–0.60 Å for the undamaged base pairs after excluding the three central and terminal base pairs (Table 1). The ALII-dA duplex structure is a B-form right-handed helix that is perturbed at the central base pair by the presence of the lesion (Figure 6, left panel). The duplex structure preserves W–C alignments on all of the unmodified base pairs, although there is a visible 23° buckle at C5•G18, a base pair that is adjacent to the ALII-dA lesion (Figure 6, top right panel). Additional structural perturbations of this lesion-flanking base pair includes the anomalous sugar conformation of G18, which appears in the C1′-endo region, and the value of the C5 glycosidic (χ) torsion angle, namely −73°. The ALII-dA nucleotide has a χ angle in the high *anti* range (−100.86°), close to the value observed in undamaged duplexes, but the conformation of the associated deoxyribose moiety is in the unusual C4′-endo range (Table 1). The ALII moiety, which is almost coplanar with the damaged adenine base, intercalates between G16 and G18 nucleotides taking up the space normally occupied by T17, the counter base, and widening the groove dimension of the duplex at the lesion site (Figure 6). Additionally, ALII establishes extensive stacking interactions with the flanking guanine bases in the complementary strand, with rings **A** and **B** exhibiting the most extensive contacts and ring **E** appearing in the major groove. Positioning of the ALII moiety inside of the undamaged strand of the duplex completely displaces T17 from the helical stack into the major groove of the duplex (Figure 6). Despite the limited number of specific interactions with other bases, T17 adopts a well-defined conformation that is very similar in all of the refined structures. Statistics of the refinement and relevant structural parameters of the ALII-dA duplex are listed in Table 1.

Duplex melting and thermodynamics

The temperature dependence of the exchangeable proton spectra confirms the hypothesis that the presence of the lesion causes limited destabilization to the ALII-dA duplex (Figure 7). Whereas increasing the temperature immediately affects the signals of the terminal imino protons, which broaden and disappear first due to their solvent exchange, the signals of the five central base pairs of the duplex retain their sharp shape even at 50°C. Broadening and disappearance of the central imino proton signals are evident at 60°C, when the duplex as a whole starts to melt. Such melting behavior is atypical of damaged DNA that generally melts from the termini toward the center of the duplex and, simultaneously, from the lesion site toward the ends. Interestingly, although the temperature increase does not affect the shape of the imino protons adjacent to the modified base pair, it leads to appreciable chemical shift changes that reach 0.4 ppm in case of

Table 1. ALII-dA duplex structure

Statistics of the refinement[a]		Relevant structural parameters at the lesion site[c]	
RMSD heavy atom position[b]	<0.7 (<0.8)	C5•G18 base pair buckle (°)	23
RMSD NOE distances (Å) $\times 10^{-2}$	8.7 (1.6)	C5; G18 sugar pucker	C2′-endo; C1′-endo
NOE energy (kcal/mol Å^2)	11.5 (43)	C5; G18 χ angle (°)	−72; −158
RMSD bond distances (Å) $\times 10^{-3}$	5.6 (6.5)	ALII-dA; T17 Sugar Pucker	C4′-endo; C2′-endo
RMSD bond angles (°)	2.2 (2.4)	ALII-dA; T17 χ angle (°)	−101; −96
RMSD dihedral angles (°)	1.9 (1.6)	C7; G16 Sugar Pucker	C1′-exo; C2′-endo
van der Waals energy (kcal/mol)	−341 (−332)	C7; G16 χ angle (°)	−119; −93

[a]Computed on the minimized averaged structure. The values for the initial minimized structure are shown in brackets.
[b]Maximum values computed among the 25 rMD structures (between brackets is the value for the central three-base pair segment.
[c]Computed from the minimized averaged structure.

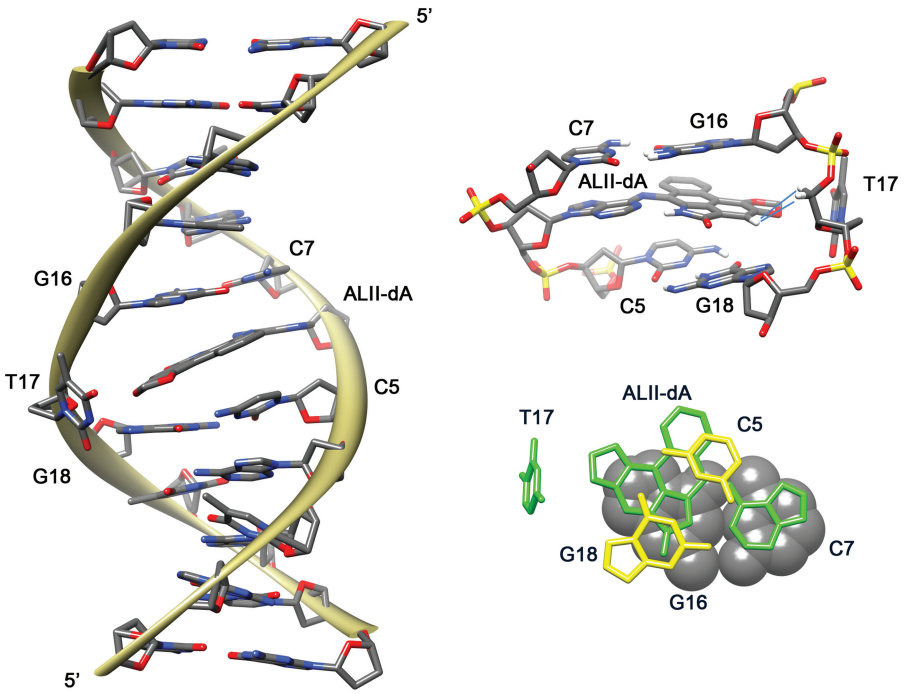

Figure 6. Solution structure of the ALII-dA containing duplex. The left panel is a view of the whole duplex colored by atom types and displayed with the major groove prominent. The picture depicts T17 complete displaced in the major groove of the duplex and the ALII moiety inside the helix. The top right panel is a close up view of the central trimeric segment of the duplex displayed with the minor groove prominent. The image depicts W–C alignments of the flanking C7•G16 and C5•G18 base pairs and indicates with blue lines the distance restrains obtained following deuteration of T17(H5″) that helped define de conformation of this nucleotide. The bottom right panel is a view, along the helical axis, of the central trimeric segment of the duplex depicting stacking interactions at the lesion site. The picture displays only the base atoms with the C7•G16 base pair shown as solid spheres, ALII-dA•T17 in green and C5•G18 in yellow.

G16(H1). Such changes can be attributed to hydrophobic effects that are known to change at higher temperatures (28). It is highly plausible that, as the temperature increases, hydrophobic interactions may compact the central part of the duplex changing, in turn, the relative positions of the ALII moiety and adjacent base pairs.

UV-melting profiles for the lesion-containing and undamaged duplexes (data not shown) confirm that the ALII-dA lesion induces the thermal destabilizing of the duplex. The magnitude of this effect varies from 3° to 6.4°, being largest in the (C–X–C) duplex case and least in the (C–X–G) duplex case. Plots of $(1/T_m)$ versus $\ln(C_t)$ for the damaged duplexes display straight lines without any visible signs of curvature (Figure 8), indicating the occurrence of a one-step duplex dissociation process that is suitable to van't Hoff analysis. The lesion increases (less negative values) the enthalpy of formation of all three duplexes though the magnitude of this effect that ranges from 4 to 20 kcal/mol is highly dependent on the sequence context of the lesion. Simultaneously, there is an increase in duplex entropy that is also sequence dependent and largely compensates for the destabilizing action of ALII-dA. As a result, the decreases of free energy caused by the ALII-dA adduct at 37°C are in the 0.8–3.2 kcal/mol range, with the largest value observed for the (C–X–C) duplex, when the lesion is flanked by hydrophilic cytosine nucleotides. Table 2 lists thermodynamic parameters for the damaged and undamaged duplexes.

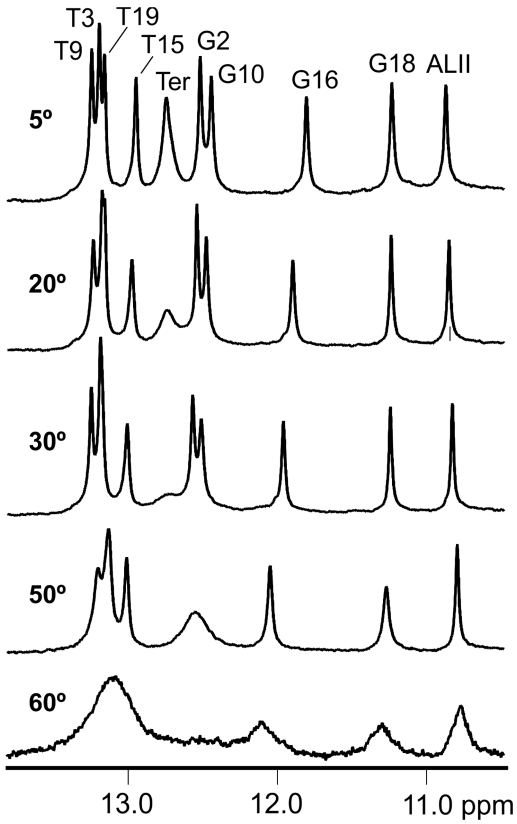

Figure 7. Temperature dependence of the imino proton signals of the ALII-dA duplex recorded at 600 MHz with the sample dissolved in 25 mM phosphate buffer, pH 6.6, 50 mM NaCl, in 10% D_2O.

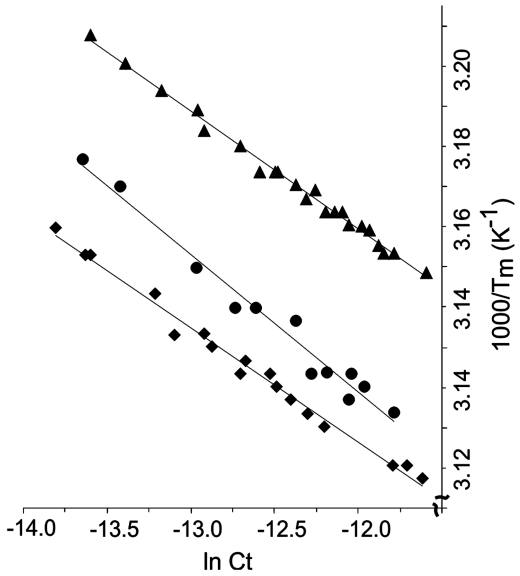

Figure 8. Plots of $1/T_m$ versus ln (C_t) for the ALII-dA (filled diamond), (TXG)- (filled circle) and (CXG)- (filled triangle) duplexes.

DISCUSSION

Tautomeric form of the ALII-dA lesion

An early study of the ALII-dA nucleoside in DMSO solution (27) demonstrated that it exists predominantly in its imino tautomeric form (Figure 3, left). Our NMR results, however, persuaded us to reconsider the tautomeric state of the ALII-dA adduct in double-stranded duplexes. The presence of the imino tautomer is inconsistent with the observation that only one imino proton signal (at 10.9 ppm) exists in close proximity to the damaged adenine H2 proton, whereas a second exchangeable signal at 7.5 ppm is close in space to the aristolactam ring **E** protons (Figure 5). In addition to the NMR evidence, molecular-modeling calculations reveal a strong steric repulsion between the adenine-H1 and aristolactam-H6 protons, making impossible a coplanar conformation of the aristolactam and adenine moieties in the ALII-N^6-dA adduct. They also demonstrate that, although the most stable conformation of the amino tautomeric form is far from planarity, the adduct can adopt a planar conformation with an energetic penalty of only 2.4 kcal/mol (Supplementary Figure S6). Taken together, our results establish that the tautomeric equilibrium of the ALII-dA shifts toward the amino form when the adduct is incorporated into DNA. It also deserves mention that, since the aristolactam imino proton is just 1.7 Å apart from the adenine N1, hydrogen bonding between these atoms is probable, a fact that would contribute further to the stabilization of this tautomer. Although the MM^+ force field used for our calculations does not take into account the possibility of hydrogen bond formation between these atoms, the observation of a sharp ALII-H6 signal at elevated temperatures (Figure 7) provides indirect evidence for the presence of such bond in the ALII-dA duplex.

NMR spectra and duplex structure

The NMR-derived MD structure of the ALII-dA duplex represents a new structural motif for DNA duplexes carrying a bulky dA-N^6 lesion. Although several solution structures of duplexes containing polyaromatic hydrocarbon (PAH) adducts at adenine-N^6 have been reported, all of them, to the best of our knowledge, are lesions formed by metabolically activated hydrocarbons. This means that the ultimate species leading to the formation of the DNA lesions are cyclic diol-epoxides and, therefore, the polyaromatic moiety is connected to the adenine amino group via a saturated (sp^3) alicyclic carbon. The geometry of such a link imposes considerable steric restraints on the mutual position of the polyaromatic and the adenine moieties; in particular, it prohibits a conformation in which they are coplanar. A major consequence of these restraints is that in such lesions the polyaromatic moiety tends to partially intercalate into the double helix, either above or below the damaged base pair. This type of conformation have been observed in the case of benzo[a]pyrenediolepoxide-N^6-dA (29–31), benzo[a]anthracenediolepoxide-N^6-dA (32–34) and benzo[c]phenanthrenediolepoxide-N^6-dA (35,36) adducts. Although the structures preserve W–C hydrogen bond alignments, either partially or completely, adduct intercalation leads to significant distortion of base

Table 2. Stability of ALII-dA-containing duplexes

Duplex sequence	T_m (°C)	ΔH° (kcal/mol)	ΔS° (cal/mol K)	$\Delta G^\circ_{25°C}$ (kcal/mol)	$\Delta G^\circ_{37°C}$ (kcal/mol)	$\Delta\Delta G^\circ_{37°C}$ (kcal/mol)
C–X–C	45.2	−76	−211	−12.7	−10.2	−2.5
C–A–C	51.6	−98	−273	−15.9	−12.7	
T–X–G	43.1	−58	−157	−11.2	−9.3	−1.2
T–A–G	46.7	−78	−219	−13.1	−10.5	
C–X–G	39.9	−68	−190	−11.0	−8.7	−0.8
C–A–G	42.9	−72	−203	−11.9	−9.5	

stacking at and around the lesion site, usually causing significant destabilization of the DNA duplex and, in some cases, heterogeneous conformations (30,31,37).

In contrast to bulky PAH lesions, formation of the ALII-dA adduct follows a quite different mechanism that involves an aromatic carbenium ion (Figure 1) as the active intermediate. Consequently, an aromatic carbon (sp^2) of the ALII moiety connects to the adenine amino group directly, resulting in a link that has a partially conjugated amino group and whose geometry is close to trigonal. Theoretically, then, the ALII and purine moieties can adopt a coplanar conformation like the one we experimentally observed in the ALII-dA duplex (Figure 6).

A direct consequence of this coplanarity is that the damaged duplex can accommodate the bulky lesion with minimal perturbation of its helical structure. The damaged nucleotide occupies mostly a normal position in the ALII-dA duplex preserving good stacking interactions with neighboring bases. Only, the unusual C4′-endo conformation adopted by its 2′-deoxysugar is outside the standard value range for B-form structures. Adjacent base pairs conserve all their W–C hydrogen bonds although one of them, C5•G18, is buckled. In addition, C5 has an uncommon glycosidic torsion angle and G18 deoxyribose exhibits a C1′-endo conformation (Table 1), indicating that some structural adjustments occur fundamentally at the 5′-side of the lesion site.

Significant structural perturbations occur at T17, the counter nucleotide of the ALII-dA adduct (Figure 6), whereby a displacement from its normal position into the major groove of the duplex, makes room for the stacking of the aristolactam moiety between G16 and G18. Although T17 loses stacking interactions with neighboring residues, it adopts a well-defined conformation and, judging by the shape of its cross-peaks in the 2D spectra, its mobility is not any different from that of other residues of the duplex. It is likely that some hydrophobic interactions also exist between the thymine base and the aristolactam methylene group that would contribute to stabilization of T17 in the major groove of the ALII-dA duplex.

The refined structure reveals extensive hydrophobic interactions between the polyaromatic ALII moiety and adjacent guanine bases that, in principle, can change significantly with different duplex sequences, especially those with less hydrophobic nucleotides stacking with the adduct. However, the NMR spectra of the TXG and CXG duplexes show that they conserve the key chemical shifts and cross-peak intensity patterns previously observed at the lesion site of the ALII-dA duplex (Supplementary Figure S7). Although computation of refined solution structures for such duplexes is beyond the scope of the present study, the similarity of these key NMR parameters suggests that the conformation at the lesion site is essentially the same in the three sequence contexts examined. It is necessary to note, that, whereas rings **B–D** of ALII are deeply buried within the base stack, ring **E** is partially exposed to the duplex major groove (Figure 6, bottom right panel), suggesting that bulkier substitutions at positions 8–11 of the AL moiety will be tolerated without affecting significantly the lesion site conformation. Therefore, we propose that the structure and physical properties of dA adducts induced by other AAs, including the more prevalent aristolochic acid I, will be similar to those of the ALII-dA adduct. Experimental verification of this hypothesis is underway.

Duplex stability

The refined structure of the ALII-dA duplex identified several factors that affect the thermal and thermodynamic stability of the damaged duplex. Stabilizing features include a minimally perturbed double-stranded helix, W–C alignments on all unmodified base pairs and the extensive hydrophobic interactions of the ALII-dA nucleotide. Traits likely leading to duplex destabilization comprise the disruption of hydrogen bonds across the damaged base pair, changes of 2′-deoxysugar conformations at the lesion site, loss of T17 stacking, partial water exposure of the ALII ring **E** in the major groove of the duplex and the almost coplanar conformation of the ALII and adenine moieties. Thermal denaturation studies of the ALII-dA duplex confirmed the compounding effect of all of these features.

Since the bulky ALII lesion disrupts the normal W–C A–T base pair at the lesion site, we expected a considerable reduction of duplex stability but, surprisingly, the destabilization effect was rather small. The temperature dependence of the imino proton signals demonstrates that the central region of the duplex conserves a double-helical structure until the duplex has completely melted. Therefore, the ALII-dA makes no appreciable contribution to the initiation of the melting process from the lesion site toward the duplex termini, as it is generally observed with other DNA lesions. The most plausible explanation for this observation is that the energetic loss resulting from the disruption of the ALII-dA•T base pair is partially compensated for by the extensive

hydrophobic interactions between the ALII moiety and the guanine bases, e.g. G16 and G18, that flank the lesion site in the complementary strand. It is also plausible that the extrahelical conformation of the T17 residue in close proximity to the aristolactam [O–CH_2–O] methylene group partially protects the aromatic moiety from solvent exposure somehow increasing stability at the lesion site.

Van't Hoff thermodynamic parameters (Table 2) quantify the impact of ALII-dA adducts in duplex stability and, since the structural perturbation are limited to the lesion-containing and flanking base pairs, reflect lesion-induced stability effects at the damaged site. As expected there is an increase (less negative) on the enthalpy of duplex formation that is sequence dependent. The entropic component of the Gibb's free energy compensates almost completely for the enthalpic variation, leading to only a moderate decrease of duplex stability and establishing the significant role played by hydrophobic interactions at the lesion site. Several polyaromatic diol epoxide dA lesions that are capable of partially intercalate into the DNA helix also display compensatory effects between the entropic and enthalpic components of duplex formation (38). In summary, our NMR-derived structure and UV-melting analysis indicate that, whereas the ALII-dA lesion causes appreciable structural and energetic perturbations of the duplex, it does not trigger local melting of the duplex, a fact that possibly has significant biological implications.

BIOLOGICAL IMPLICATIONS

A uniqueness of the NER pathway is its ability to recognize a virtually infinite number of bulky lesions in double-stranded DNA. The mechanism responsible for this characteristic remains a subject of active investigation, especially the initial lesion recognition step mediated by XPC/hR23B that operates during global genomic NER (39). Although the crystal structure of RAD4/RAD23, the yeast ortholog of XPC/hR23B, in a complex with a UV-damaged duplex has been determined (40) and many solution structures of duplexes containing bulky lesions are also available (41), it is still unclear what features of damaged duplexes elicit their recognition by the NER system. As a rule, recognizable NER substrates contain a covalently bound bulky adduct that induces structural distortions of the DNA double helix (42,43). In addition, lesions that decrease duplex stability creating a 'bubble' at the lesion site of the double helix are better NER substrates, suggesting that a thermodynamic probing mechanism operates during the initial lesion recognition step (44,45). The observation that the crystal structure of the RAD4/RAD23 in complex with UV-damaged DNA shows a β-hairpin wedging into the DNA helix and assisting the local melting of the duplex further supports the stability-sensing mechanism of NER recognition (40). This mechanism also explains the fact that some bulky lesions that increase the duplex stability, such as the 3-(deoxyguanosin-N^2-yl)-2-acetylaminofluorene (dG(N^2)-AAF), are persistent in cellular DNA (46,47) or the observation that removal of the lesion counter base cytosine increases the thermal stability and hinders NER incision of a duplex containing a *cis*-benzo[a]pyrene adduct (48). Other hypotheses that postulated alternative recognition factors, such as lesion-induced changes of local DNA flexibility (49), increased helical dynamics (50) or poor base stacking (51) were rapidly eclipsed by the thermodynamic destabilization mechanism of NER recognition. Thus, the current paradigm states that XPC/hR23B would sense locally destabilized duplex structures and, since duplex destabilizing lesions facilitate DNA strand separation, such lesion-induced destabilization would be the hallmark for NER recognition (45). Although the thermodynamic destabilization hypothesis of NER recognition satisfactorily explains the detection of a large range of bulky lesions, it fails in the case of the ALII-dA adduct. This lesion disrupts the duplex structure at the damaged site and decreases the free energy of duplex formation but, nonetheless, is a poor substrate for *in vitro* NER incision (V. Sidorenko, *et al.*, companion manuscript), has long persistence in genomic DNA (1) and seems to be recognized exclusively by the transcription-coupled NER pathway (13). In order to reconcile our observations with the thermodynamic hypothesis of NER recognition, one could suggest that there is a 'destabilization threshold' that must be overcome for successful NER recognition. If this is the case, the local destabilization energy needed for NER recognition should exceed 2.5 kcal/mol at 37°C. A comparable destabilization value was measured for the (-)-*trans* isomer of the DB[*a,l*]pyrene-N^6-dA adduct (38) that, as the ALII-dA lesion, also adopts an intercalative structure (52) that seems refractory to global NER removal in human cells (53). Alternatively, our results may support one or more of the other NER recognition hypotheses being the base-stacking hypothesis (51) the most relevant to the case of the ALII-dA adduct. The solution structure of the ALII-dA duplex demonstrates that both the damaged adenine and ALII moieties form extensive hydrophobic contacts with flanking residues without perturbing stacking of the undamaged base pairs. The significant role of such hydrophobic effects is evident from the results of our thermodynamic study (Table 2). It would be possible to suggest also that, since the ALII-dA nucleotide is tightly packed inside the base pair stack, it would reduce the internal mobility of the DNA double helix at the lesion site, resulting in lack of recognition of the ALII-dA adduct by the XPC/hR23B dimmer (49). The sharp and uniform cross-peak shape observed on the NMR spectra may serve as indirect evidence of the absence of conformational transitions detectable in the NMR timescale. However, further experiments, such as ^{13}C relaxation-based dynamics studies of damaged duplexes, are needed to establish the role of internal dynamics during NER recognition of ALII-dA.

In closing, the discovery of the structural and energy destabilizing effects of the ALII-dA adduct, a lesion that is poorly recognized by the global genomic NER pathway, provide new arguments for the need to put forward additional DNA factors during NER recognition, and emphasize the role that stacking interactions and possibly internal dynamics have in this process.

ACCESSION NUMBERS

PDB ID: 2LGM, RCSB ID: RCSB102368.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Table 1, Supplementary Figures 1–7.

AKNOWLEDGEMENTS

We are grateful to Drs Mike Goger, Shibani Bhattacharya and Kaushik Dutta from New York Structural Biology Center for assistance with NMR experiments, and to Mrs Irina Zaitseva for assistance with the preparation of oligonucleotides. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Environmental Health Sciences or the National Institutes of Health.

FUNDING

Funding for open access charge: NIH Grants ES004068 (to F.J. and to C.D.L.S.) and ES017368 (to C.D.L.S.).

Conflict of interest statement. None declared.

REFERENCES

1. Arlt,V.M., Stiborova,M. and Schmeiser,H.H. (2002) Aristolochic acid as a probable human cancer hazard in herbal remedies: a review. *Mutagenesis*, **17**, 265–277.
2. National Toxicology Program Report on Carcinogens. (2009) *Aristolochic Acids*, 12th edn. Research Triangle Park, North Carolina, US Department of Health and Human Services, Public Health Service.
3. Cosyns,J.-P., Jadoul,M., Squifflet,J.P., de Plaen,J.F., Ferluga,D. and van Ypersele de Strihou,C. (1994) Chinese herbs nephropathy: a clue to Balkan endemic nephropathy? *Kidney Int*., **45**, 1680–1688.
4. Grollman,A.P., Shibutani,S., Moriya,M., Miller,F., Wu,L., Moll,U., Suzuki,N., Fernandes,A., Rosenquist,T., Medverec,Z. *et al*. (2007) Aristolochic acid and the etiology of endemic (Balkan) nephropathy. *Proc. Natl Acad. Sci. USA*, **104**, 12129–12134.
5. Nortier,J.L., Martinez,M.C., Schmeiser,H.H., Arlt,V.M., Bieler,C.A., Petein,M., Depierreux,M.F., De Pauw,L., Abramowicz,D., Vereerstraeten,P. *et al*. (2000) Urothelial carcinoma associated with the use of a Chinese herb (Aristolochia fangchi). *N. Engl. J. Med*., **342**, 1686–1692.
6. Stiborová,M., Frei,E., Arlt,V.M. and Schmeiser,H.H. (2008) Metabolic activation of carcinogenic aristolochic acid, a risk factor for Balkan endemic nephropathy. *Mutat. Res*., **658**, 55–67.
7. Schmeiser,H.H., Frei,E., Wiessler,M. and Stiborova,M. (1997) Comparison of DNA adduct formation by aristolochic acids in various in vitro activation systems by 32P-post-labelling: evidence for reductive activation by peroxidases. *Carcinogenesis*, **18**, 1055–1062.
8. Stiborová,M., Mareîs,J., Frei,E., Arlt,V.M., Martínek,V. and Schmeiser,H.H. (2011) The human carcinogen aristolochic acid I is activated to form DNA adducts by human NAD(P)H: quinone oxidoreductase without the contribution of acetyltransferases or sulfotransferases. *Environ. Mol. Mutagen*, **52**, 448–459.
9. Fernando,R.C., Schmeiser,H.H., Scherf,H.R. and Wiessler,M. (1993) Formation and persistence of specific purine DNA adducts by 32P-postlabelling in target and non-target organs of rats treated with aristolochic acid I. *IARC Sci. Publ*., **124**, 167–171.
10. Bieler,C.A., Stiborova,M., Wiessler,M., Cosyns,J.-P., van Ypersele de Strihou,C. and Schmeiser,H.H. (1997) ^{32}P-post-labelling analysis of DNA adducts formed by aristolochic acid in tissues from patients with Chinese herbs nephropathy. *Carcinogenesis*, **18**, 1063–1067.
11. Attaluri,S., Bonala,R.R., Yang,I.-Y., Lukin,M.A., Wen,Y., Grollman,A.P., Moriya,M., Iden,C.R. and Johnson,F. (2010) DNA adducts of aristolochic acid II: total synthesis and site-specific mutagenesis studies in mammalian cells. *Nucleic Acids Res*., **38**, 339–352.
12. Nedelko,T., Arlt,V.M., Phillips,D.H. and Hollstein,M. (2009) *TP53* mutation signature supports involvement of aristolochic acid in the aetiology of endemic nephropathy-associated tumours. *Int. J. Cancer*, **124**, 987–990.
13. Moriya,M., Slade,N., Brdar,B., Medverec,Z., Tomic,K., Jelaković,B., Wu,L., Truong,S., Fernandes,A. and Grollman,A.P. (2011) TP53 Mutational signature for aristolochic acid: An environmental carcinogen. *Int. J. Canc*, March 16 (doi: 10.1002/ijc.26077; epub ahead of print).
14. Lukin,M. and de los Santos,C. (2010) Stereoselective nucleoside deuteration for NMR studies of DNA. *Nucleosides Nucleotides Nucleic Acids*, **29**, 562–573.
15. Beaucage,S.L. (1993) Oligodeoxyribonucleotide synthesis: Phosphoramidite approach. In: Agrawal,S. (ed.), *In Protocols for Oligonucleotides and Analogs, Methods in Molecular Biology Series*, Vol. 20. Human Press, Totowa, NJ, pp. 33–61.
16. Hwang,T.L. and Shaka,A.J. (1995) Water suppression that works. Excitation sculpting using arbitrary wave-forms and pulsed-field gradients. *J. Mag. Reson*, **112**, 275–279.
17. Delaglio,F., Grzesiek,S., Vuister,G.W., Zhu,G., Pfeifer,J. and Bax,A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**, 277–293.
18. Schwieters,C.D., Kuszewski,J.J., Tjandra,N. and Clore,G.M. (2003) The Xplor-NIH NMR molecular structure determination package. *J. Mag. Reson*., **160**, 66–74.
19. Brooks,B., Bruccoleri,R., Olafson,B., States,D., Swaminathan,S. and Karplus,M. (1983) CHARMM: a program for macromolecular energy minimization and dynamics calculations. *J. Comp. Chem*, **4**, 187–217.
20. Friedman,R.A. and Honig,B. (1992) The electrostatic contribution to DNA base–stacking interactions. *Biopolymers*, **32**, 145–159.
21. Yip,P. and Case,D.A. (1989) A new method for refinement of macromolecular structures based on nuclear Overhauser effect spectra. *J. Mag. Reson*., **83**, 643–648.
22. Zaliznyak,T., Bonala,R., Attaluri,S., Johnson,F. and de los Santos,C. (2009) Structure of duplex DNA containing α-OH-PdG: the mutagenic adduct produced by acrolein. *Nucleic Acids Res*., **37**, 2153–2163.
23. Lavery,R. and Sklenar,H. (1988) The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dyn*, **6**, 655–667.
24. Pettersen,E.F., Goddard,T.D., Huang,C.C., Couch,G.S., Greenblatt,D.M., Meng,E.C. and Ferrin,T.E. (2004) 'UCSF Chimera – A visualization system for exploratory research and analysis.'. *J. Comput. Chem*., **25**, 1605–1612.
25. Plum,E.G., Breslauer,K.J. and Roberts,R.W. (1999) Thermodynamics and kinetics of nucleic acids association/dissociation and folding processes. In: Kool,E.T. (ed.), *Comprehensive Natural Products Chemistry*, Vol. 7. Elsevier Science Ltd, Oxford, pp. 15–53.
26. de los Santos,C. (1999) Probing DNA structure by NMR spectroscopy. In: Kool,E.T. (ed.), *Comprehensive Natural Products Chemistry*, Vol. 7. Elsevier Science Ltd, Oxford, pp. 55–80.
27. Pfau,W., Schmeiser,H.H. and Wiessker,M. (1991) N6-Adenyl arylation of DNA by aristolochic acid II and a synthetic model for the putative proximate carcinogen. *Chem. Res. Toxicol*., **4**, 851–586.
28. Muller,N. (1990) Search for a realistic view of hydrophobic effects. *Acc. Chem. Res*., **23**, 23–28.
29. Zegar,I.S., Kim,S.J., Johansen,T.N., Horton,P.J., Harris,C.M., Harris,T.M. and Stone,M.P. (1996) Adduction of the human N-ras codon 61 sequence with (-)-(7S,8R,9R,10S)-7,8-dihydroxy-9,10-epoxy-7,8,9,10-tetrahydrobenzo[a] pyrene: structural refinement of the intercalated SRSR(61,2) (-)-

(7S,8R,9S,10R)-N6-[10-(7,8,9,10- tetrahydrobenzo[a]pyrenyl)]-2′-deoxyadenosyl adduct from 1H NMR. *Biochemistry*, **35**, 6212–6224.
30. Zegar,I.S., Chary,P., Jabil,R.J., Tamura,P.J., Johansen,T.N., Lloyd,R.S., Harris,C.M., Harris,T.M. and Stone,M.P. (1998) Multiple conformations of an intercalated (-)-(7S,8R,9S, 10R)-N6-[10-(7,8,9,10-tetrahydrobenzo[a]pyrenyl)]-2′-deoxyadenosyl adduct in the N-ras codon 61 sequence. *Biochemistry*, **37**, 16516–16528.
31. Volk,D.E., Rice,J.S., Luxon,B.A., Yeh,H.J.C., Liang,C., Xie,G., Sayer,J.M., Jerina,D.M. and Gorenstein,D.G. (2000) NMR evidence for syn-anti interconversion of a trans opened (10R)-dA adduct of benzo[a]pyrene (7S,8R)-diol (9R,10S)-epoxide in a DNA duplex. *Biochemistry*, **39**, 14040–14053.
32. Li,Z., Mao,H., Kim,H.-Y., Tamura,P.J., Harris,C.M., Harris,T.M. and Stone,M.P. (1999) Intercalation of the (-)-(1R,2S,3R, 4S)-N6-[1-benz[a]anthracenyl]-2′-deoxyadenosyl adduct in an oligodeoxynucleotide containing the human N-ras codon 61 sequence. *Biochemistry*, **38**, 2969.
33. Li,Z., Kim,H.-Y., Tamura,P.J., Harris,C.M., Harris,T.M. and Stone,M.P. (1999) Intercalation of the (1S,2R,3S,4R)-N6-[1-(1,2,3,4-tetrahydro-2,3, 4-trihydroxybenz[a]anthracenyl)]-2′-deoxyadenosyl adduct in an oligodeoxynucleotide containing the human N-ras codon 61 sequence. *Biochemistry*, **38**, 16045–16057.
34. Li,Z., Tamura,P.J., Wilkinson,A.S., Harris,C.M., Harris,T.M. and Stone,M.P. (2001) Intercalation of the (1R,2S,3R,4S)-N6-[1-(1,2,3,4-tetrahydro-2,3,4-trihydroxybenz[a]anthracenyl)]-2′-deoxyadenosyl adduct in the N-ras codon 61 sequence: DNA sequence effects. *Biochemistry*, **40**, 6743–6755.
35. Cosman,M., Fiala,R., Hingerty,B., Laryea,A., Lee,H., Harvey,R.G., Amin,S., Geacintov,N.E., Broyde,S. and Patel,D. (1993) Solution conformation of the (+)-trans-anti-[BPh]dA adduct opposite dT in a DNA duplex: intercalation of the covalently attached benzo[c]phenanthrene to the 5′-side of the adduct site without disruption of the modified base pair. *Biochemistry*, **32**, 12488–12497.
36. Cosman,M., Laryea,A., Fiala,R., Hingerty,B.E., Amin,S., Geacintov,N.E., Broyde,S. and Patel,D.J. (1995) Solution conformation of the (-)-trans-anti-benzo[c]phenanthrene-dA ([BPh]dA) adduct opposite dT in a DNA duplex: intercalation of the covalently attached benzo[c]phenanthrenyl ring to the 3′-side of the adduct site and comparison with the (+)-trans-anti-[BPh]dA opposite dT stereoisomer. *Biochemistry*, **34**, 1295–1307.
37. Yan,S., Shapiro,R., Geacintov,N.E. and Broyde,S. (2001) Stereochemical, structural, and thermodynamic origins of stability differences between stereoisomeric benzo[a]pyrene diol epoxide deoxyadenosine adducts in a DNA mutational hot spot sequence. *J. Am. Chem. Soc.*, **123**, 7054–7066.
38. Ruan,Q., Kolbanovskiy,A., Zhuang,P., Chen,J., Krzeminski,J., Amin,S. and Geacintov,N. (2002) Synthesis and characterization of site-specific and stereoisomeric fjord dibenzo[a,l]pyrene diol epoxide-N^6-adenine adducts: Unusual thermal stabilization of modified DNA duplexes. *Chem. Res. Toxicol.*, **15**, 249–261.
39. Gillet,C.J. and Scharer,O.D. (2006) Molecular mechanisms of mammalian global genome nucleotide excision repair. *Chem. Rev.*, **106**, 253–276.
40. Min,J.H. and Pavletich,N.P. (2007) Recognition of DNA damage by the Rad4 nucleotide excision repair protein. *Nature*, **449**, 570–575.
41. Lukin,M. and de los Santos,C. (2006) NMR structures of damaged DNA. *Chem. Rev.*, **106**, 607–686.
42. Sugasawa,K., Okamoto,T., Shimizum,Y., Masutanim,C., Iwaim,S. and Hanaokam,F.A. (2001) A multistep damage recognition mechanism for global genomic nucleotide excision repair. *Genes Dev.*, **15**, 507–521.
43. Schärer,O.D. (2007) Achieving broad substrate specificity in damage recognition by binding accessible nondamaged DNA. *Mol. Cell*, **28**, 184–186.
44. Gunz,D., Hess,M.T. and Naegeli,H. (1996) Recognition of DNA adducts by human nucleotide excision repair. Evidence of a thermodynamic probing mechanism. *J. Biol. Chem.*, **271**, 25089–25098.
45. Zheng,H., Cai,Y., Ding,S., Tang,Y., Kropachev,K., Zhou,Y., Wang,L., Wang,S., Geacintov,N.E., Zhang,Y. *et al.* (2010) Base flipping free energy profiles for damaged and undamaged DNA. *Chem. Res. Toxicol*, **23**, 1868–1870.
46. Zaliznyak,T., Bonala,R., Johnson,F. and de los Santos,C. (2006) Structure and stability of duplex DNA containing the 3-(Deoxyguanosin-N^2-yl)-2-acetylaminofluorene (dG(N^2)-AAF) lesion: a bulky adduct that persists in cellular DNA. *Chem. Res. Toxicol.*, **19**, 745–752.
47. Cui,X.S., Eriksson,L.C. and Moller,L. (1999) Formation and persistence of DNA adducts during and after a long-term administration of 2-nitrofluorene. *Mutat. Res.*, **442**, 9–18.
48. Reeves,D.A., Mu,H., Kropachev,K., Cai,Y., Ding,S., Kolbanovskiy,A., Kolbanovskiy,M., Chen,Y., Krzeminski,J., Amin,S. *et al.* (2011) Resistance of bulky DNA lesions to nucleotide excision repair can result from extensive aromatic lesion-base stacking interactions. *Nucleic Acids Res.*, **39**, 8752–8764.
49. Isaacs,R.J. and Spielmann,H.P. (2004) A model for initial DNA lesion recognition by NER and MMR based on local conformational flexibility. *DNA Repair*, **3**, 455–464.
50. Maillard,O., Camenisch,U., Clement,F.C., Blagoev,K.B. and Naegeli,H. (2007) DNA repair triggered by sensors of helical dynamics. *Trends Biochem. Sci.*, **32**, 494–499.
51. Yang,W. (2006) Poor base stacking at DNA lesions may initiate recognition by many repair proteins. *DNA Repair*, **5**, 654–666.
52. Cai,Y., Ding,S., Geacintov,N.E. and Broyde,S. (2011) Intercalative conformations of the 14R (Þ)- and 14S (-)-trans-anti-DB[a,l]P-N6-dA adducts: molecular modeling and MD simulations. *Chem. Res. Toxicol.*, **24**, 522–531.
53. Dreij,K., Seidel,A. and Jernström,B. (2005) Differential removal of DNA adducts derived from *anti*-diol epoxides of dibenzo[*a,l*]pyrene and benzo[*a*]pyrene in human cells. *Chem. Res. Toxicol.*, **18**, 655–664.