Interstrand DNA-DNA cross-link formation between adenine residues and abasic sites in duplex DNA

Nathan E. Price‡, Kevin M. Johnson‡, Jin Wang†, Mostafa I. Fekry‡, Yinsheng Wang‡, and Kent S. Gates‡,§,*

‡Department of Chemistry, University of Missouri, Columbia, MO 65211
§Department of Biochemistry, University of Missouri, Columbia, MO 65211
†Department of Chemistry, University of California-Riverside, CA 92521-0403

Supporting Information

Table of Contents
Materials and Methods .................................................................3
Figure S1. Formation time course for duplex B .......................................8
Figure S2. Yields of cross-link in duplex B under various conditions.............9
Figure S3. Stability of duplex B under various conditions..........................10
Figure S4. Substitution of dC for dA at the cross-linking site abrogates cross-link formation...10
Figure S5. MALDI-TOF-MS of cross-linked duplex B ..................................11
Figure S6. LC-MS³ of cross-linked duplex B ...........................................12
Scheme S1. Pyranose-furanose forms of the cross-link remnant.....................13
Figure S7. Time course for cross-link formation in duplex I ..........................14
Figure S8. Yields of cross-link formation in duplex I under various conditions........15
Figure S9. Stability cross-link in duplex I under various post-formation work-up conditions….16
Figure S10. Hydroxyl radical footprinting of cross-linked duplex J ..................17
Figure S11. Nanospray QTOF MS of a cross-linked duplex containing the same core sequence found in duplex I .................................................................18
Figure S12. LC-ESI-MS and MS/MS for the analysis of the nuclease P1 digestion mixture of duplex I containing a dA-Ap cross-link.........................................................19
Figure S13. LC-ESI-MS/MS/MS analysis of the 4-enzyme digestion mixture of duplex I containing a dA-Ap cross-link.................................................................20

Figure S14. LC-ESI-MS/MS/MS analysis of the 4-enzyme digestion mixture of duplex I examining whether the duplex contains a dG-Ap cross-link.................................................................21

Figure S15. A molecular model showing the locations of the exocyclic amino groups of dG and dA relative to the aldehyde carbon of an Ap site in a 5’CApT/AAG sequence such as that found in duplex I.
Materials and Methods

Oligonucleotides were purchased from Integrated DNA Technologies. All enzymes were purchased from New England Biolabs (Ipswich, MA, USA). [γ-32P]-ATP (6000 Ci/mmol) was purchased from Perkin Elmer. C-18 sep-pak cartridges were from Waters. BS Polyprep columns were purchased from BioRad. Erythro-9-(2-hydroxy-3-nonyl)adenine (EHNA) hydrochloride was from Tocris Bioscience (Ellisville, MO, USA). Nuclease P1 and phosphodiesterases 1 and 2 were from Sigma-Aldrich (St. Louis, MO, USA). Alkaline phosphatase and proteinase K were from New England Biolabs (Ipswich, MA, USA). Quantification of radioactivity in polyacrylamide gels was carried out using a Personal Molecular Imager (BIORAD) with Quantity One software (v.4.6.5). All other reagents were purchased from Sigma-Aldrich.

Representative procedure for cross-link formation. Single-stranded uracil containing oligodeoxynucleotides were 5’-labeled using standard procedures. Labeled DNA was annealed with its complimentary strand and treated with the enzyme UDG (50 units/mL, final concentration) to generate Ap sites. UDG enzyme was removed by phenol-chloroform extraction. Individual DNA duplexes were incubated in a buffer composed of HEPES (50 mM, pH 7) containing NaCl (100 mM) at 37 °C for 120 h. The DNA was ethanol precipitated, resuspended in formamide loading buffer, loaded onto a 20% denaturing polyacrylamide gel, and the gel electrophoresed for 4 h at 1000 V. Direct loading of cross-linking reactions (without ethanol precipitation) showed that ethanol precipitation does not alter the yield of the slow-migrating, cross-link band. The amount of radiolabeled DNA in each band on the gel was measured by phosphorimager analysis. The time course for the formation of the Ap site cross-link was carried out by incubating a solution containing labeled DNA (approximately 300,000
cpm), in HEPES (50 mM, pH 7) and NaCl (100 mM) at 37 °C. At specified time points, aliquots (10 µL) were removed and frozen at –20 °C, after all time points were collected samples were loaded directly on the gel and analyzed as described above.

**Hydroxyl radical footprinting of cross-linked duplexes.** Standard literature protocols were used in the footprinting of the cross-linked duplexes. In these experiments, the strand opposing the uracil-containing oligonucleotide was 5’-labeled using standard procedures. Labeled DNA was annealed with the uracil-containing complement and treated with UDG to generate the abasic site as described above. The Ap-containing double-stranded DNA (~500,000 cpm) was incubated in HEPES (50 mM, pH 7) and NaCl (100 mM) at 37 °C for 120 h. The DNA was ethanol precipitated, suspended in formamide loading buffer, and the oligodeoxynucleotides resolved on a 0.4 mm thick 20% denaturing polyacrylamide gel. The slow-migrating cross-linked duplex band was visualized using X-ray film, the band cut out of the gel, the gel slice crushed, and the gel pieces vortexed in elution buffer (NaCl 200 mM; EDTA, 1 mM) at room temperature for at least 1 h. The mixture was filtered through a poly-prep column to remove gel fragments and the filtrate desalted using a C18 Sep-pak (100 mg size). The resulting solution was evaporated using a Speed-Vac concentrator, the residue redissolved in water (24 µL), split evenly into three microcentrifuge tubes, and diluted with 2x oxidation buffer (10 µL of a solution composed of sodium phosphate, 20 mM, pH 7.2; NaCl, 20 mM; sodium ascorbate, 2 mM; H2O2, 1 mM). To this mixture was added a solution of iron-EDTA (2 µL, EDTA, 70 mM; Fe(NH4)2(SO4)2•6H2O, 70 mM) to start the reaction, the mixture vortexed briefly, and incubated at room temperature for 1, 2, and 3 min before addition of thiourea stop solution (10 µL of a 100 mM solution in water). Hydroxyl radical footprinting reactions,
Maxam-Gilbert G, and Maxam-Gilbert A+G reactions were performed on the labeled single-strand to generate marker lanes. The resulting DNA fragments were analyzed using gel electrophoresis as described above.

**Static Nanospray QTOF MS.** The oligonucleotide sample was analyzed in a 40 mM dimethylbutylammonium acetate (pH 7.1) buffer. Negative ion MS spectra was taken for mass range of 280-3200 Da on an Agilent6520A QTOF MS with Chip Cube source (G4240A). Monoisotopic neutral masses were calculated from the multiply charged ion spectrum present in the 500-2000 Da mass range. Sample introduction was done with New Objective Econo12 N uncoated borosilicate glass emitters. Negative ion spectrum was acquired at a capillary potential sufficient to initiate spray of the sample. The nitrogen gas was heated to 290°C and introduced at a flow rate of 4L/min. The Fragmentor, Skimmer, and Octopole1 RF Vpp potentials were set to 200 V, 65 V, and 750 V, respectively. External Calibration was done with the Agilent ESI-Low calibration tuning mixture (cat. no. G1969-85000) and data analysis was performed with Agilent Mass Hunter Workstation QUalitative Analysis software v B.02.00, Build 2.0.197.0 with Bioconfirm Software (2008). Peptide isotope model was assumed and peak set height threshold for extraction was set to ≥ 500 counts. Deconvolution was done with a 0.1 Da step size with a result of 20 iterations of the algorithm calculation

**Enzymatic digestion.** Duplex I was digested using a 4-enzyme cocktail following the conditions described previously. Briefly, nuclease P1 (5 U), phosphodiesterase 2 (0.01 U), EHNA (20 nmol) and a 50-µL solution containing 300 mM sodium acetate (pH 5.6) and 10 mM zinc
chloride were added to 300 pmol of duplex I in a final volume of 500 µL. In this context, EHNA served as an inhibitor for deamination of 2’deoxyadenosine (dA) to 2’-deoxyinosine (dI) induced by adenine deaminase. The resulting mixture was incubated at 37˚ C for 48 h. To the digestion mixture were then added alkaline phosphatase (10 U), phosphodiesterase 1 (0.005 U) and 100 µL of 0.5 M Tris-HCl buffer (pH 8.9). The digestion was continued at 37˚ C for 2 h. Duplex B was digested using the 4-enzyme cocktail under similar conditions. Digestion with nuclease P1 alone was performed following previously published procedures, where 0.2 unit of nuclease P1 (0.2 U) was added to 300 pmol of duplex I in a 500 µL water solution. The resulting mixture was incubated at 37˚ C for 2 h. The above enzymatic digestion mixture was extracted with chloroform to remove enzymes, and the aqueous layer was dried in Speed-vac, reconstituted in water, and subjected to LC-MS/MS analyses as described previously.
1. Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular Cloning: A Lab Manual*; Cold Spring Harbor Press: Cold Spring Harbor, NY, 1989.

2. Sczepanski, J. T.; Jacobs, A. C.; Majumdar, A.; Greenberg, M. M. *J. Am. Chem. Soc.* 2009, *131*, 11132-11139.

3. Luce, R. A.; Hopkins, P. B. *Methods Enzymol.* 2001, *340*, 396-412.

4. Maxam, A. M.; Gilbert, W. *Methods Enzymol.* 1980, *65*, 499-560.

5. Johnson, K. M.; Price, N. E.; Wang, J.; Fekry, M. I.; Dutta, S.; Seiner, D. R.; Wang, Y.; Gates, K. S. *J. Am. Chem. Soc.* 2013, *135*, 1015-1025.

6. Wang, J.; Yuan, B.; Guerrero, C.; Bahde, R.; Gupta, S.; Wang, Y. *Anal. Chem.* 2011, *83*, 2201-2209.
**Figure S1.** Time course for the formation of the dA-Ap cross-link. Duplex B was incubated in HEPES buffer (50 mM, pH 7) containing NaCl (100 mM) at 37 °C. At each time point aliquots were removed and frozen at –20 °C. The $^{32}$P-labeled oligodeoxynucleotides were resolved on a denaturing polyacrylamide gel. **Left:** Lane 1, uracil-containing precursors of the oligodeoxynucleotide duplex. Lane 2, the abasic-site-containing duplexes cleaved by treatment with piperidine (1 M, 95 °C, for 25 min). Lanes 3-21 are formation time points in hours 0, 0.5, 1, 2, 6, 8, 10, 24, 30, 49, 55, 73, 78, 92, 101, 120, 144, 148, 168. The $^{32}$P-labeled oligodeoxynucleotides were resolved on a denaturing polyacrylamide gel A. The radioactivity in each band quantitatively measured by phosphorimager analysis, and the yield of slow-migrating (cross-linked DNA) band plotted versus time. **Right:** Time = 0 was defined as the end of the 1.5 h reaction of the uracil-containing oligodeoxynucleotide duplex with UDG to generate the Ap-containing oligodeoxynucleotide duplex B.
Figure S2. Cross-link formation in duplex B in the presence of various buffers, salt conditions, and additives. The Ap-containing duplex B was incubated for 120 h at 37 °C. The $^{32}$P-labeled oligodeoxynucleotides were resolved on a denaturing polyacrylamide gel and the radioactivity in each band quantitatively measured by phosphorimager analysis. In the graph above, the yield of slow-migrating (cross-linked DNA) band was normalized against a reaction carried out under standard conditions (50 mM HEPES pH 7.0, 100 mM NaCl, bar 1, far left on the graph). The conditions employed for the other reactions, from left to right were: bar 2, 50 mM cacodylate pH 7.0, 100 mM NaCl; bar 3, 16.5 mM NaPO$_4$, 1.8 mM citric acid pH 7.0, 100 mM NaCl; bar 4, 50 mM NaPO$_4$ pH 7.2, 50 mM NaCl; bar 5, 50 mM MOPS pH 7.0, 100 mM NaCl; bar 6, 50 mM Bis, 50 mM Tris pH 7.0, 100 mM NaCl; bar 7, 50 mM HEPES pH 7.0, 100 mM NaCl, 2 mM MgCl$_2$; bar 8, 50 mM HEPES pH 7.0, 300 mM NaCl; bar 9, 50 mM HEPES pH 7.0, 100 mM NaCl, 1 mM GSH; bar 10, 50 mM HEPES pH 7.0, 100 mM NaCl, 1 mM DTT; bar 11, 50 mM HEPES pH 7.0, 100 mM NaCl, 100 mM phenol.
Figure S3. Stability of the cross-linked duplex B under various conditions. Duplex B was incubated in HEPES buffer (50 mM, pH 7) containing NaCl (100 mM) at 37 °C for 120 h at 37 °C, ethanol precipitated, and then subjected to various conditions. The $^{32}$P-labeled oligodeoxynucleotides were resolved on a denaturing polyacrylamide gel and the radioactivity in each band quantitatively measured by phosphorimager analysis, and the yield of slow-migrating (cross-linked DNA) band was normalized against an aliquot that was precipitated and then resuspended in water (bar 1). Other aliquots were subjected to various work-up conditions: bar 2, 100 mM piperidine 60 °C for 15 min; bar 3, 90 °C for 15 min; bar 4, pH 3 (25 mM sodium acetate) 24 °C for 15 min; bar 5, pH 10 (25 mM potassium phthalate) 24 °C for 15 min, and bar 5, 200 mM methoxyamine at 60 °C for 30 min.
**Figure S4.** Substitution of dC for dA at the cross-linking site abrogates cross-link formation. Lanes 1-5, duplex B, lanes 6-10, duplex K. Lanes 1 and 6 contain ^32^P-labeled, uracil-containing 2'-deoxyoligonucleotide duplexes. Lanes 2 and 7 contain the abasic-site-containing duplexes subjected to piperidine work-up (1 M, 95 °C, 25 min) to cleave the Ap site. Lanes 3 and 8 contain the ^32^P-labeled UDG-treated (abasic-site-containing) duplexes without incubation. Lanes 4 and 9 contain the cross-linking reactions involving incubation of the abasic-site-containing duplexes in HEPES buffer (50 mM, pH 7) containing NaCl (100 mM) for 120 h at 37 °C. Lanes 5 and 10 contain the methoxyamine-capping reactions involving incubation of the abasic-site-containing duplexes incubated in HEPES buffer (50 mM, pH 7) containing NaCl (100 mM), and CH\textsubscript{3}ONH\textsubscript{2}•HCl (2 mM) at 37 °C. The ^32^P-labeled 2'-deoxyoligonucleotides were resolved by electrophoresis on a 20% denaturing polyacrylamide gel and the radioactivity in each band quantitatively measured by phosphorimager analysis.
MALDI TOF mass spectrometric analysis of an unlabeled version of cross-linked duplex B generated by incubation of duplex B in HEPES buffer (50 mM, pH 7.0) containing NaCl (100 mM) at 37 °C. The cross-link was purified by preparative gel electrophoresis, desalted using a G-25 Sephadex column and ethanol precipitated. The spectrum was externally calibrated using an oligonucleotide standard with m/z of 12,378.11. MALDI TOF analysis was performed as described previously (SI citation 5).
Figure S6. LC-ESI-MS/MS/MS analysis of the 4-enzyme digestion mixture of cross-linked duplex B. Shown here is the $252 \rightarrow 136 \, m/z$ transition in positive ion mode.
Scheme S1. Expected furanose (top) and pyranose (bottom) isomers of the cyclic hemiaminal arising from the reaction of deoxyadenosine with deoxyribose.
Figure S7. Time course for the formation of cross-link in duplex I. The abasic-site-containing duplex was incubated in HEPES buffer (50 mM, pH 7.0) and NaCl (100 mM) at 37 °C and aliquots were removed from the reaction and frozen prior to sequencing gel analysis (lanes 4-10). The lower bands correspond to the full length labeled oligodeoxynucleotides and the upper band cross-linked DNA. Lane 1 is the $^{32}$P-labeled uracil-containing precursor oligodeoxynucleotide, lane 2 is the $^{32}$P-labeled abasic-site-containing duplexes cleaved by treatment with piperidine (1 M, 95 °C, for 25 min), and lane 3 is the $^{32}$P-labeled abasic-site-containing duplex without incubation. The $^{32}$P-labeled oligodeoxynucleotides were resolved on a polyacrylamide gel and the radioactivity in each band quantitatively measured by phosphorimager analysis. The left panel shows a representative gel depicting the first 28 h of the cross-linking reaction. The right panel shows results from obtained from full time course experiments.
Figure S8. Relative yields of cross-linked DNA generated in duplex I in the presence of various buffers, salt conditions, and thiol additives. The “control” (or standard) cross-linking reaction (relative yield 100%, bar 1 on the graph) involved incubation of duplex I in HEPES buffer (50 mM, pH 7) containing NaCl (100 mM) at 37 °C for 120 h. These experiments examine the effects of different buffers, salts, or thiols present during the cross-linking reaction. Yields under other conditions were normalized against that obtained under standard conditions. The conditions examined were: bar 2, 50 mM cacodylate pH 7.0, 100 mM NaCl; bar 3, 50 mM NaPO₄ pH 7.2, 50 mM NaCl; bar 4, 50 mM MOPS pH 7.0, 100 mM NaCl; bar 5, 50 mM Bis, 50 mM Tris, 5 mM DTPA pH 7.0, 100 mM NaCl; bar 6, 50 mM HEPES pH 7.0, 100 mM NaCl, 2 mM MgCl₂; bar 7, 50 mM HEPES pH 7.0, 300 mM NaCl; bar 8, 50 mM HEPES pH 7.0, 100 mM NaCl, 1 mM GSH; bar 9, 50 mM HEPES pH 7.0, 100 mM NaCl, 1 mM DTT.
Figure S9. Stability of the cross-linked duplex I when subjected to various post-formation work-up conditions. Panel A displays the remaining amount of cross-linked DNA after various treatments compared with untreated cross-link. Lane 1 is the uracil-containing precursor of oligodeoxynucleotide duplex I. Lane 2 is the abasic-site-containing oligodeoxynucleotide duplex I without incubation. Lane 3 is the abasic-site-containing oligodeoxynucleotide duplex I cleaved by treatment with piperidine (1 M, 95 °C, 25 min). Lane 4 is duplex I incubated in HEPES buffer (50 mM, pH 7) containing NaCl (100 mM) at 37 °C, ethanol precipitated and stored in water for 15 min. Lanes 5-9 show cross-linked DNA generated as described for lane 4 except subjected to various work-up conditions: lane 5, 60 °C, 15 min; lane 6, 90 °C, 15 min; lane 7, pH adjusted to 3, 15 min; lane 8, pH adjusted to 10, 15 min; lane 9, 100 mM piperidine, 60 °C, 15 min. Panel B is a bar graph with a quantitative illustration of the data from the gels. The yields of the cross-linked DNA obtained remaining after various work-ups were normalized against the amount of cross-link present in the sample that was similarly precipitated, but not subjected to any work-up (lane 4).
Figure S10. Hydroxyl radical footprinting of duplex J to locate the site of cross-link attachment (Panel A). Duplex J contains the same core sequence as that in duplex I. The slightly longer duplex J shown above was employed to move the bands at the cross-linking site out of a “salt-flare” zone in the polyacrylamide gel and gives clear, sharp bands at the cross-link site. Lane 1 is a Maxam-Gilbert G-specific cleavage (sequencing) reaction of the labeled oligodeoxynucleotide strand in duplex J. Lane 2 is an A+G specific cleavage (sequencing) reaction of the labeled oligodeoxynucleotide strand in duplex J. Lanes 3-5 are the hydroxyl radical footprinting reactions of the labeled oligodeoxynucleotide strand in duplex J. Lanes 6-9 are the hydroxyl radical footprinting reactions of the slow-migrating, cross-link band generated by incubation of duplex J in HEPES buffer (50 mM, pH 7.0) and NaCl (100 mM) at 37 °C. The $^{32}$P-labeled oligodeoxynucleotides were resolved on a polyacrylamide gel visualized by phosphorimager analysis (Panel A) and used to develop densitometry traces (Panels B-D). Panel B is the Maxam-Gilbert G-specific cleavage (sequencing) reaction of the labeled oligodeoxynucleotide strand in duplex J. Panel C is the A+G specific cleavage (sequencing) reaction of the labeled oligodeoxynucleotide strand in duplex J. Panel D is the hydroxyl radical footprinting reaction of the slow-migrating cross-link band.
**Figure S11.** Nanospray QTOF MS of a cross-linked duplex containing the same core sequence found in duplex I.
Figure S12. LC-ESI-MS and MS/MS for the analysis of the nuclease P1 digestion mixture of duplex I with a dA-Ap cross-link. Shown in (a) is the selected-ion chromatogram (SIC) for monitoring the loss of a thymidine-5’-phosphate from the [M–2H]^{2-} ion of the tetramer shown (i.e., the m/z 571.3→821.4 transition). Inset gives the higher resolution “ultra-zoom” ESI-MS for the [M–2H]^{2-} ion of the tetramer. Displayed in (b) is the tandem mass spectrum for the [M–2H]^{3-} ion of the tetramer.
Figure S13. LC-ESI-MS/MS/MS for the analysis of the 4-enzyme digestion mixture of duplex I with a dA-Ap cross-link. Shown in (a) is the selected-ion chromatogram (SIC) for monitoring the m/z 368 → 252 → 136 transition. Displayed in (b) is the MS/MS from the cleavage of the [M + H]^+ ion of the completely digested cross-link remnant. Depicted in (c) is MS/MS/MS arising from the fragmentation of the ion of m/z 252 observed in (b).
Figure S14. LC-ESI-MS/MS/MS for the analysis of the 4-enzyme digestion mixture of duplex I, probing whether the cross-linked duplex contains a dG-Ap cross-link. Shown in (a) is the selected-ion chromatogram (SIC) for monitoring the $m/z$ 384 → 268 → 240 transition. Displayed in (b) is the MS/MS from the cleavage of the $[M + H]^+$ ion of the completely digested cross-link remnant. Depicted in (c) is MS/MS/MS arising from the fragmentation of the ion of $m/z$ 268 observed in (b).
Figure S15. A molecular model showing the locations of the exocyclic amino groups of dG and dA relative to the aldehyde carbon of an Ap site in a 5’-CApT/AAG sequence such as that found in duplex I. The model Ap site was generated in Pymol by deletion of a thymine nucleobase from B-DNA containing the relevant 5’-CApT/AAG sequence (pdb code 1CS2). DNA duplexes containing Ap sites often retain a B-DNA-like structure, especially when purine bases are located directly opposing the Ap site. For example, see: Chen, J.; Dupradaeu, F.-Y.; Case, D. A.; Turner, C. J.; Stubbe, J. Nucleic Acids Res. 2008, 36, 253-262. The amino group of dG (4.19 Å) is closer than that of dA (8.19 Å), yet cross-linking to the dA predominates.