Published in final edited form as:

Angew Chem Int Ed Engl. 2015 June 22; 54(26): 7666–7669. doi:10.1002/anie.201502566.

A Simple, High-Yield Synthesis of DNA Duplexes Containing a Covalent, Thermally-Reversible Interstrand Cross-link At a Defined Location**

Jacqueline Gamboa Varela and

Department of Chemistry, University of Missouri, Columbia, MO 65211

Kent S. Gates*

Departments of Chemistry and Biochemistry, University of Missouri, Columbia, MO 65211

Abstract

Interstrand DNA-DNA cross-links are highly toxic to cells because these lesions block the extraction of information from the genetic material. The pathways by which cells repair cross-links are important, but not well understood. The preparation of chemically well-defined cross-linked DNA substrates represents a significant challenge in the study of cross-link repair. Here we report a simple method that employs “post-synthetic” modifications of commercially available 2'-deoxyoligonucleotides to install a single cross-link in high yield at a specified location within a DNA duplex. The cross-linking process exploits hydrazone formation between a non-natural N⁴-amino-2'-deoxycytidine nucleobase and the aldehyde residue of an abasic site in duplex DNA. The resulting cross-link is stable under physiological conditions, but can be readily dissociated and re-formed through heating-cooling cycles.

Keywords

nucleic acids; DNA damage; cross-links; hydrazone; abasic site

The sequence of nucleobases on the interior of the DNA double helix constitutes the genetic code that guides all cellular operations.^[1,2] Interstrand cross-links (ICLs) are highly toxic to cells because these lesions prevent the strand separation that is required to extract genetic information from DNA.^[3] ICLs are important in diverse fields including medicinal chemistry, toxicology, and gerontology. For example, ICLs generated by some antineoplastic drugs selectively kill rapidly-dividing cancer cells.^[3a] ICL formation may contribute to the cytotoxic and carcinogenic effects of various pyrrolizidine alkaloids.^[3d] In addition, circumstantial evidence suggests that the formation of some as-yet-unidentified endogenous DNA cross-links may drive cellular senescence and aging.^[3b,4]

**We are grateful to the National Institutes of Health for support of this work (ES021007). MALDI mass spectrometry was conducted at the Charles W. Gehrke Proteomics Center at the University of Missouri and we thank Beverly DaGue for assistance with these experiments.

*gatesk@missouri.edu.

Supporting information for this article is available on the WWW under <http://www.angewandte.org> or from the author.

The complex multistep processes by which cells repair cross-links are important, but not well understood.^[5] The preparation of chemically well-defined cross-linked DNA substrates represents a significant challenge in the study of cross-link repair. Treatment of duplex DNA with cross-linking agents typically generates mixtures of structurally distinct ICLs in low yields.^[3b,6] As a result, laborious purification is required to obtain duplexes containing a single, structurally-defined ICL. Alternatively, elegant multistep chemical syntheses of cross-linked duplexes have been developed,^[7] but these methods may not be practicable in some biochemistry and biology laboratories where the study of cross-link repair takes place.

Here we report a very simple method for the generation of cross-linked DNA duplexes in high yields. Our approach employs inexpensive reagents to carry out “post-synthetic” modification of commercially available 2'-deoxyoligonucleotides and installs a single cross-link at a specified location within a DNA duplex. This work was inspired by our recent studies of imine-derived interstrand cross-links generated via the reaction of abasic sites with opposing guanine and adenine residues in duplex DNA.^[8] The cross-linking process described here exploits hydrazone formation between the non-natural nucleobase, *N*⁴-amino-2'-deoxycytidine (dC*), and the aldehyde residue of an abasic (Ap) site **1** located on opposing strands of an oligomeric DNA duplex (Scheme 1). Our previous work with the native nucleobases^[8] and examination of Ap-containing duplex structures^[9] suggested that these reaction partners would be well positioned when embedded within a 5'-ApA/5'-C*A sequence (Figure 1). We elected to mispair the dC* residue, with the expectation that greater dynamic freedom might facilitate cross-linking.

The 2'-deoxyoligonucleotide **B** containing a single dC* residue was prepared by treatment of the corresponding cytosine-containing oligonucleotide **A** with bisulfite (0.3 M) and hydrazine (4 M) in sodium phosphate (0.1 M, pH 5) for 3 h at 50 °C (Scheme 2A). This protocol mirrors the conditions developed by Hayatsu and coworkers for preparation of the *N*⁴-amino-2'-deoxycytidine nucleoside^[10] and adapted by Gao and Orgel for use with 2'-deoxyoligonucleotides.^[7c] The 5'-³²P-labeled, Ap-containing oligonucleotide **E** was prepared by treatment of the corresponding 5'-³²P-labeled 2'-deoxyuridine-containing oligonucleotide **D** with the enzyme uracil DNA glycosylase (UDG, Scheme 2B).^[11] Efficient formation of the Ap-containing oligonucleotide **E** was confirmed by treatment of the DNA with warm piperidine to generate the faster-migrating 3'-phosphate cleavage product **F** (Figure 3, lane 3).^[12]

The complementary oligonucleotides **B** and **E** were combined in sodium phosphate (50 mM, pH 5) containing NaCl (100 mM) at 50 °C, and allowed to stand at room temperature for 5 h. Electrophoretic analysis of the resulting labeled DNA on a 20% denaturing polyacrylamide gel revealed a very high yield (90 ± 2%) of a slow-moving band in the region of the gel where the cross-linked duplex **G** would be expected^[8a-c] to appear (Figure 3, lane 4). Inclusion of the aldehyde-capping reagent methoxyamine (2 mM)^[13] in the reaction significantly inhibited formation of the slow-moving band (85% decrease in yield), consistent with involvement of the Ap aldehyde residue in generation of this product (Figure 3, lane 5). No slow-moving band was generated when the ³²P-labeled Ap-containing oligonucleotide **E** was incubated with the native dC-containing oligonucleotide **A** (Figure 3, lane 6) or when hydrazine was omitted from the protocol used for preparation of the dC*-

containing oligonucleotide **B** (Figure 3, lane 7). When bisulfite was omitted from the dC* preparation protocol, a weak slowing-moving band (~8%) was generated (Figure 3, lane 8). Combination of the dC*-containing oligonucleotide **B** with the dU-containing oligonucleotide **D** did not produce a slow-moving band (Figure 3, lane 9). Likewise, treatment of a thymine-containing control oligonucleotide **C** (lacking the critical C residue) with bisulfite and hydrazine, followed by combination with the Ap-containing oligonucleotide **E**, did not produce a slow-moving band (Figure 3, lane 10).

Taken together, the results provided evidence for the central roles of both the dC* and Ap residues in generation of the slow-migrating band observed on the gel. Furthermore, combination of non-complementary oligonucleotides containing the dC* and Ap residues did not produce detectable amounts of a slow-moving band (Figure S1). This provided evidence that enforced proximity of the dC* and Ap residues within the DNA duplex was required for generation of the slow-moving band seen in Figure 3. Importantly, gel electrophoretic analysis revealed that the putative cross-linked duplex **G** was completely stable under physiological conditions (37 °C, 16 h, pH values of 5, 7, or 8, Figure S2).

We next conducted experiments establishing that the slow-migrating band seen in Figure 3 was, in fact, the cross-linked duplex **G**. First, addition of a seven-nucleotide overhang to the 3'-end of either oligonucleotide **B** or **E** resulted in further retardation in the gel mobility of the slow-moving band (Figure 4). These results provided evidence that the slow-moving bands observed in Figure 4 were, indeed, cross-linked duplexes **G**, **H**, and **I** containing two full-length strands. Second, ESI(-)TOF mass spectrometric analysis of the cross-linked duplex **J** (Figure S3) revealed strong signals matching the isotope envelope expected for the isobaric hydrazone cross-links **3** or **4** (molecular formula $C_{410}H_{517}N_{151}O_{245}P_{40}$).^[14]

We also examined the ability of the dC*-Ap cross-link to stabilize a 15 nucleotide DNA duplex against thermal melting. Toward this end, we first characterized melting of the Ap-containing duplex **K** using UV-vis spectrometry. This duplex displayed a melting temperature of 32 ± 0.6 °C (1 °C/min, 1 μ M duplex, in 50 mM sodium phosphate, pH 5, 100 mM NaCl, Figure 5A). The cooling curve for **K** closely matched the melting curve, as expected for a reversible DNA melting-hybridization process. Duplex **L**, containing the dC*-Ap cross-link, showed a marked increase in the melting transition, with an apparent T_m of 64 ± 1 °C at a heating rate of 1 °C/min (Figure 5A). Interestingly, the cooling curve for duplex **L** did *not* mirror the melting curve and instead closely resembled the annealing curve for the Ap-containing duplex **K**. This suggested that the hydrazone cross-link in duplex **L** hydrolyzed during the melting process (Figure 5B). Re-melting the sample after cooling produced a melting curve that closely matched the original melting curve, indicating that the cross-link in duplex **L** was regenerated during the cooling/annealing process. A second round of cooling gave a curve that again closely matched the annealing curve of the Ap-containing duplex **K**. Consistent with the notion that the heating curve observed for duplex **L** (Figure 5) reflects cross-link hydrolysis alongside duplex melting, the rate of heating exerted a significant effect on the apparent melting temperature, with T_m values of 56 ± 2 °C, 64 ± 1 °C, and 72 ± 2 °C observed at heating rates of 0.5 °C/min, 1 °C/min, and 3 °C/min, respectively. Taken together, the data are consistent with the presence of a covalent, but thermally-reversible, cross-link in duplex **L** (Figure 5B). The covalent-reversible nature

of the dC*-Ap cross-link is strongly reminiscent of the “post-it type” interstrand cross-link recently described by Carell and coworkers that employs reversible imine formation.^[7a]

In conclusion, we have developed a fast, simple, and inexpensive protocol for the installation of a single interstrand cross-link at a well-defined location in duplex DNA. The cross-link possesses sufficient chemical stability under physiologically-relevant conditions for use in biochemical assays and structural studies.^[15] In addition, we provided evidence that the dC*-Ap cross-link stabilizes duplex DNA against thermal melting and can reversibly traverse a thermal cycle involving hydrolysis of the hydrazone linkage during heating/melting, followed by regeneration of the cross-link during cooling/annealing. This type of covalent, thermally-reversible cross-link may find applications in both biochemistry and materials science.^[7a,15,16]

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Crick FH. C. *Nature*. 1970; 227:561–563.
2. Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. *Molecular Biology of the Cell*. 4th ed.. Garland Science; New York: 2002.
3. a Schärer OD. *ChemBioChem*. 2005; 6:27–32. [PubMed: 15637664] b Rajski SR, Williams RM. *Chem. Rev*. 1998; 98:2723–2795. [PubMed: 11848977] c Muniandy PA, Liu J, Majumdar A, Liu S.-t. Seidman MM. *Crit. Rev. Biochem. Mol. Biol*. 2010; 45:23–49. [PubMed: 20039786] d Woo J, Sigurdsson ST, Hopkins PB. *J. Am. Chem. Soc*. 1993; 115:3407–3415.
4. a Garinis GA, van der Horst GTJ, Vijg J, Hoeijmakers JH. *J. Nat. Cell Biol*. 2008; 10:1241–1247. b Niedernhofer LJ, Lalai AS, Hoeijmakers JH. *J. Cell*. 2005; 123:1191–1198. c Bergstrahl DT, Sekelsky J. *Trends in Genetics*. 2007; 24:70–76. d Grillari J, Katinger H, Voglauer R. *Nucleic Acids Res*. 2007; 35:7566–7576. [PubMed: 18083760]
5. Clauson C, Schärer OD, Niedernhofer LJ. *Cold Spring Harbor Perspectives in Biology*. 2013; 5:a012732/1–a012732/25. [PubMed: 24086043]
6. a Hearst JE. *Chem. Res. Toxicol*. 1989; 2:69–75. [PubMed: 2519712] b Rink SM, Solomon MS, Taylor MJ, Rajur SB, McLaughlin LW, Hopkins PB. *J. Am. Chem. Soc*. 1993; 115:2551–2557. c Takasugi M, Guendouz A, Chassignol M, Decout JL, Lhomme J, Thuong NT, Hélène C. *Proc. Nat. Acad. Sci. USA*. 1991; 88:5602–5606. [PubMed: 2062839]
7. a Tomás-Gamasa M, Serdjukow S, Su M, Müller M, Carell T. *Angew. Chem. Int. Ed. Eng*. 2014; 53:796–800. b Angelov T, Guainazzi A, Schärer OD. *Org. Lett*. 2009; 11:661–664. [PubMed: 19132933] c Gao K, Orgel LE. *Proc. Nat. Acad. Sci. USA*. 1999; 96:14837–14842. [PubMed: 10611299] d Ye M, Guillaume J, Liu Y, Sha R, Wang R, Seeman NC, Canary JW. *Chem. Sci*. 2013; 4:1319–1329. [PubMed: 23894693] e Pujari SS, Leonard P, Seela F. *J. Org. Chem*. 2014; 79:4423–4437. [PubMed: 24693949] f Stevens K, Madder A. *Nucleic Acids Res*. 2009; 37:1555–1565. [PubMed: 19151089] g Nishimoto A, Jitsuzaki D, Onizuka K, Tnaiguchi Y, Nagatsugi F, Sasaki S. *Nucleic Acids Res*. 2013; 41:6774–6781. [PubMed: 23778430] h Noll DM, Noronha AM, Miller PS. *J. Am. Chem. Soc*. 2001; 123:3405–3411. [PubMed: 11472110] i Balkrishnen B, Leonard NJ, Robinson H, Wang AH-J. *J. Am. Chem. Soc*. 1996; 118:10744–10751. j Manoharan M, Andrade LK, Cook PD. *Org. Lett*. 1999; 1:311–314. [PubMed: 10822568] k Nakatani K, Yoshida T, Saito I. *J. Am. Chem. Soc*. 2002; 124:2118–2119. [PubMed: 11878957] l Hong IS, Ding H, Greenberg MM. *J. Am. Chem. Soc*. 2006; 128:485–491. [PubMed: 16402835] m Hentshel S, Alzeer J, Angelov T, Schärer OD, Luedtke NW. *Angew. Chem. Int. Ed. Eng*. 2012; 51:3466–3469. n Qiu Z, Lu L, He C. *J. Am. Chem. Soc*. 2008; 130:14398–14399. [PubMed: 18842048] o Dohno C, Shibata T, Nakatani K. *Chem. Comm*. 2010; 46:5530–5532. [PubMed: 20458423]

- Saito I. J. Am. Chem. Soc. 2005; 127:16681–16684. [PubMed: 16305258] p Mukerjee S, Guainazzi A, Schärer OD. Nucleic Acids Res. 2014; 42:7429–7435. [PubMed: 24782532]
8. Johnson KM, Price NE, Wang J, Fekry MI, Dutta S, Seiner DR, Wang Y, Gates KS. J. Am. Chem. Soc. 2013; 135:1015–1025. [PubMed: 23215239] Dutta S, Chowdhury G, Gates KS. J. Am. Chem. Soc. 2007; 129:1852–1853. [PubMed: 17253689] Price NE, Johnson KM, Wang J, Fekry M,I, Wang Y, Gates KS. J. Am. Chem. Soc. 2014; 136:3483–3490. [PubMed: 24506784] For other cross-links derived from abasic sites, see: Ichikawa K, Kojima N, Hirano Y, Takebayashi T, Kowatab K, Komatsu Y. Chem. Comm. 2012; 48:2143–2145. [PubMed: 22246106] and reference 7j.
9. Chen J, Dupradeau F-Y, Case DA, Turner CJ, Stubbe J. Nucleic Acids Res. 2008; 36:253–262. [PubMed: 18025040]
10. Negishi K, Kawakami M, Kayasuga K, Odo J, Hayatsu H. Chem. Pharm. Bull. 1987; 35:3884–3887. [PubMed: 3435982]
11. a Lindahl T, Ljunquist S, Siebert W, Nyberg B, Sperens B. J. Biol. Chem. 1977; 252:3286–3294. [PubMed: 324994] b Varshney U, van de Sande JH. Biochemistry. 1991; 30:4055–4061. [PubMed: 2018771]
12. a Bailly V, Verly WG. Biochem. J. 1988; 253:553–559. [PubMed: 2460081] b Bayley CR, Brammer KW, Jones AS. J. Chem. Soc. 1961:1903–1907.c Gates KS. Chem. Res. Toxicol. 2009; 22:1747–1760. [PubMed: 19757819]
13. Rosa S, Fortini P, Karran P, Bignami M, Dogliotti E. Nucleic Acids Res. 1991; 19:5569–5574. [PubMed: 1719478]
14. Some N-aryl glycosyl hydrazones exist in the cyclic glycosyl N-aryl hydrazine form, see: Takeda Y. Carbohydrate Res. 1979; 77:9–23.
15. Synthetic interstrand cross-links have been used to stabilize short duplexes for biophysical analysis, structural studies, and for buttressing DNA-based materials, see ref 7a and: Ferentz AE, Keating TA, Verdine GL. J. Am. Chem. Soc. 1993; 115:9006–9014. McManus FP, Khaira A, Noronha AM, Wilds CJ. Bioconjugate Chem. 2013; 24:224–233. Chen W, Schuster GB. Org. Biomol. Chem. 2013; 11:35–40. [PubMed: 23086526] Altmann S, Labhardt AM, Bur D, Lehmann C, Bannwarth W, Billeter M, Wüthrich K, Leupin W. Nucleic Acids Res. 1995; 23:4827–4835. [PubMed: 8532525] Glick G. Biopolymers. 1998; 48:83–96. [PubMed: 9846126]
16. Liu Y-L, Chuo T-W. Biopolym. Chem. 2013; 4:2194–2205.

Figure 1.
Molecular model illustrating the juxtaposition of the dC* and Ap residues in a DNA duplex containing a 5'-ApA/5'-C*A sequence.

A 5' -GGTGTGAGGGAGTGAGTGA**C**ATTGGAGAGTGTGTT
B 5' -GGTGTGAGGGAGTGAGTGA**C***ATTGGAGAGTGTGTT
C 5' -GGTGTGAGGGAGTGAGTGA**T**ATTGGAGAGTGTGTT
D 5' ³²P-AACACACTCTCCAA**dU**ATCACTCACTCCCTCACACC
E 5' ³²P-AACACACTCTCCAA**Ap**ATCACTCACTCCCTCACACC
F 5' ³²P-AACACACTCTCCAA-OPO₃²⁻
G 5' ³²P-AACACACTCTCCAA**Ap** A ATCACTCACTCCCTCACACC
 3' -TTGTGTGAGAGGTT A **C***TAGTGAGTGAGGGAGTGTGG
H 5' ³²P-AACACACTCTCCAA**Ap** A ATCACTCACTCCCTCACACC
 3' -**AGGAGAA**TTGTGTGAGAGGTT A **C***TAGTGAGTGAGGGAGTGTGG
I 5' ³²P-AACACACTCTCCAA**Ap** A ATCACTCACTCCCTCACACC**AAGAGGA**
 3' -TTGTGTGAGAGGTT A **C***TAGTGAGTGAGGGAGTGTGG
J 5' -AACATCAAA**Ap** A TTAACACAAA
 3' -TTGTAGTTT A **C***AATTGTGTTT
K 5' -AACTCAA**Ap**AATCCT
 3' -TTGAGTT A **C**TAGGA **L** 5' -AACTCAA**Ap** A ATCCT
 3' -TTGAGTT A **C***TAGGA

Figure 2.
DNA sequences used in this study.

Figure 3.

Generation of an interstrand dC*-Ap cross-link in duplex **G**. Complementary oligonucleotides in sodium phosphate buffer (50 mM, pH 5) containing NaCl (100 mM) were briefly warmed to 50 °C and then incubated at room temperature (23 °C) for 5 h. The samples were mixed with formamide loading buffer and subjected to electrophoretic analysis on a 20% denaturing polyacrylamide gel (0.4 mm thick). The labeled fragments in the gel were visualized and quantitatively measured by phosphorimager analysis. Lane 1: size marker consisting of the ^{32}P -labeled dU-containing oligonucleotide **D**; Lane 2, the Ap-containing 2'-deoxyoligonucleotide **E**; Lane 3, the Ap-containing oligonucleotide **E** treated with piperidine (1 M, 95 °C, 30 min); Lane 4, generation of cross-linked duplex **G** from the complementary oligonucleotides **D**+**E**; Lane 5, **D**+**E** in the presence of $\text{CH}_3\text{ONH}_2\cdot\text{HCl}$ (2 mM); Lane 6, **A**+**E**; Lane 7, **A**+**E**, when hydrazine was omitted from the C* protocol; Lane 8, **A**+**E**, when bisulfite was omitted from C* the protocol; Lane 9, **B**+**D**; Lane 10, **C**+**E**, where oligo **C** was subjected to the C*-forming protocol.

Figure 4.

Evidence that the slow-moving band is an interstrand DNA cross-link containing both full-length oligonucleotides. Complementary oligonucleotides in sodium phosphate buffer (50 mM, pH 5) containing NaCl (100 mM) were briefly warmed to 50 °C and then incubated at room temperature (23 °C) for 5 h. The samples were mixed with formamide loading buffer and subjected to electrophoretic analysis on a 20% denaturing polyacrylamide gel (0.4 mm thick). The labeled fragments in the gel were visualized and quantitatively measured by phosphorimager analysis. Lane 1, dU-containing oligonucleotide **D**; Lane 2, Ap-containing oligonucleotide **E**; Lane 3, Ap-containing oligonucleotide **E** treated with piperidine (1 M, 95 °C, 30 min); Lane 4, cross-linked duplex **G**; Lane 5, cross-linked duplex **H**; Lane 6, size marker consisting of the ^{32}P -labeled, dU-containing oligonucleotide precursor component of duplex **I**; Lane 7, Ap-containing oligonucleotide component of duplex **I**; Lane 8, Ap-containing oligonucleotide component of duplex **I** treated with piperidine (1 M, 95 °C, 30 min); Lane 9, cross-linked duplex **I**.

Figure 5.

Thermal cycling of the dC*-Ap cross-link. Panel A: Melting-annealing cycles of the cross-linked duplex **L** and the Ap-containing duplex **K**. The DNA duplexes ($\sim 1 \mu\text{M}$) were heated in sodium phosphate buffer (50 mM, pH 5) containing NaCl (100 mM) at a rate of $1^\circ\text{C}/\text{min}$ and absorbance was monitored at 260 nm. Melting curve of the Ap-containing duplex **K** is shown in dark gray ($T_m 32 \pm 0.6^\circ\text{C}$) and the annealing curve for **K** is light gray ($T_m 28 \pm 0.6^\circ\text{C}$). The first heating curve for cross-linked duplex **L** is shown in purple ($T_m 64 \pm 1^\circ\text{C}$), the first cooling curve for **L** is red ($T_m 33 \pm 1^\circ\text{C}$), the second heating curve is orange, and the second cooling curve is green. Panel B: Schematic diagram of species proposed to be present at various times during the melting and annealing of duplex **L**.

Scheme 1.
Formation of dC*-Ap cross-link in duplex DNA.

Scheme 2. Synthesis of **dC*** and **Ap**-containing oligonucleotides.