A modified thymine for the synthesis of site-specific thymine-guanine DNA interstrand crosslinks

Jawad Alzeer¹ and Orlando D. Schärer¹,²,*

¹Institute of Molecular Cancer Research, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland and ²Department of Pharmacological Sciences, Chemistry Graduate Building 619, Stony Brook University, Stony Brook, NY11794-3400

Received June 22, 2006; Revised July 26, 2006; Accepted July 28, 2006

ABSTRACT

DNA interstrand crosslinks (ICLs) are highly cytotoxic lesions formed by a variety of important anti-tumor agents. Despite the clinical importance of ICLs, the mechanisms by which these lesions are repaired in mammalian cells have so far remained elusive. One of the obstacles in the study of ICL repair has been the limited availability of suitable methods for the synthesis of defined site-specific ICLs. We report here the synthesis of a site-specific ICL containing an ethylene-bridged G-T base pair based on the incorporation of a crosslink precursor containing a selectively reactive group on one strand using solid-phase synthesis. 3-(2-chloroethyl)thymidine was incorporated into oligonucleotides and underwent ICL formation upon annealing to a complementary strand by reacting with the base opposite to the modified T residue. A strong preference for ICL formation with a G residue opposite the reactive T was observed. Detailed characterization of the reaction product revealed that the alkylation reaction occurred with the O-6 group of G and a mechanism accounting for this preference is proposed. These G-T crosslinks introduced here will be useful for studies of ICL repair.

INTRODUCTION

DNA interstrand crosslinks (ICLs) are among the most cytotoxic lesions known and are produced as a result of a reaction of cellular DNA with compounds containing bifunctional electrophilic moieties. The extraordinary cytotoxicity of ICLs is due to the fact that they block all cellular processes that involve the separation of dsDNA such as transcription or replication (1). ICL-forming bifunctional compounds are frequently used in cancer chemotherapy, but are also generated endogenously as metabolic side products (2). To counteract the threats of ICLs, cells have developed mechanisms to repair these lesions, and the failure to repair such lesions is associated with inherited disorders such as Fanconi Anemia (3,4). In contrast, the repair of ICLs in tumor cells leads to resistance against anti-tumor agents (5,6). The understanding of the mechanisms of ICL repair should, therefore, not only significantly contribute to our understanding of the molecular basis of human disease, but also to the mechanisms of resistance of tumor cells to chemotherapeutic agents. This knowledge should thus eventually facilitate the development of new targets for cancer therapy. Although significant progress has been made in studying various DNA repair pathways (7,8), the mechanisms underlying the repair of ICLs remain poorly understood (1,2,5,9). Apart from the inherent complexity of ICL repair, in which two damaged strands need to be repaired, the difficulty of preparing well-defined and stable ICL-containing oligonucleotides as substrates for biochemical and cell biological studies presents an additional significant hurdle in enabling the investigation of ICL repair pathways. Traditionally ICLs have been prepared by treating dsDNA with bifunctional alkylating agents, but this approach offers little or no sequence specificity, produces complex mixtures of mono adducts, intra- and interstrand crosslinks and very low overall yield of the desired ICL (usually <5%) (10–14).

De novo chemical synthesis provides an alternative for the generation of site-specific ICLs. One approach has relied on the cross-linking of mono nucleosides outside of DNA and the incorporation of crosslinked dimers into DNA using solid-phase synthesis after appropriate functionalization and protection (15–20). This approach has been useful for the preparation of defined short oligonucleotides, but is somewhat limited in scope due to the need of up to three different protecting groups on the phosphoramidite of the dimer. A practically more straightforward approach to the synthesis of ICLs consists in the incorporation of crosslink precursor in complementary single-stranded oligonucleotides and installation of the crosslink using a suitable coupling reaction after annealing of the two oligonucleotides. This approach has been used with precursors in one (21–23) or both strands of DNA (24–26).

As part of our program to study the repair of ICLs formed by bifunctional alkylating agents used in cancer

*To whom correspondence should be addressed. Tel: +1 631 632-77545; Fax: +1 631 632-7546; Email: orlando@pharm.stonybrook.edu

© 2006 The Author(s).
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.0/uk/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
chemotherapy, we were interested in developing synthetic methods for the preparation of ICLs containing an ethylene bridge between two Watson–Crick paired bases to mimic the lesion formed by the chloro ethyl nitroso ureas (CENUs) (14,27). We describe a new approach for the preparation of ethylene bridge cross-linked DNA using a thymine containing a chloroethyl group at the N-3 position as a precursor. When paired to a complementary oligonucleotide with the modified T opposite to G, efficient ICL formation was observed. Interestingly, we found that of the two possible alkylation products with N-1 or O-6 of guanine, only reaction with the O-6 position was observed. Our studies thus provide a new approach to the synthesis of defined DNA interstrand crosslinks.

**MATERIALS AND METHODS**

All solvents and reagents were purchased from Sigma-Aldrich. Anhydrous solvents were prepared by distillation under an argon atmosphere in the presence of appropriate drying agents. Flash chromatography was performed using Silica gel 60 (220–430 mesh) as the adsorbent. Thin layer chromatography (TLC) was performed using precoated TLC plates (Merck, Silica gel 60 F254, 0.25 mm).

^1H-NMR and ^13C-NMR spectra were recorded on a Bruker 300 MHz NMR spectrometer operating at room temperature. Chemical shifts are reported in parts per million (δ) downfield from tetramethylsilane. Spin multiplicities are described as s (singlet), bs (broad singlet), d (doublet), dd (doublet of doublets), t (triplet) and m (multiplet). Mass spectra were recorded by ESI-MS (Finnigan MAT TSQ-700) and oligonucleotides were measured on Q-TOF Ultima Mass Spectrometer (Micromass, UK) with a nano-source in negative mode. The spectra were obtained by deconvolution of the multiply charged ions using the MaxEnt1 program (Micromass, UK).

HPLC was performed on a JASCO system equipped with a BGB Analytic column: Nucleosil 100 Å, C18, 5 mm 250 × 4.0 mm². The C18-SepPak cartridges were obtained from MicroPore. Snake venom phosphodiesterase I (SVPD) was purchased from Worthington Biochemical; calf intestine phosphatase, from New England Biolabs.

\[ N^3-(2-Chloroethyl)-5'-O-dimethoxytrityl-3'-O-(tert-butyl-dimethylsilyl)thymidine (2) \]

At room temperature, Cs₂CO₃ (850 mg, 86%) as a white foam. Rₛ (Hexane/EtOAc 3:1) 0.67; ^1H NMR (300 MHz, (CD₃)₂SO): δ 7.66 (d, 1H, J = 1.1 Hz), 7.41 (dd, 2H, J = 7.2 Hz), 7.28 (m, 7H), 6.82 (d, 4H, J = 8.77 Hz), 6.36 (t, 1H, J = 6.26 Hz), 4.52 (m, 1H), 4.31 (dt, 2H, J = 1.7, 6.9 Hz), 3.98 (dd, 1H, J = 2.73, 5.85 Hz), 3.77 (s, 6H), 3.71 (t, 2H, J = 6.8 Hz), 3.48 (dd, 1H, J = 2.76, 10.66 Hz), 3.20 (dd, 1H, J = 2.85, 10.64 Hz), 2.34 (m, 1H), 2.20 (m, 1H), 1.52 (s, 3H), 0.84 (s, 9H), 0.03 (s, 3H), 0.00 (s, 3H); ^13C NMR (75 MHz, (CD₃)₂SO): δ 163.12, 158.65, 150.62, 144.13, 135.32, 133.90, 128.95, 128.04, 127.86, 127.01, 113.17, 109.99, 86.70, 86.69, 85.51, 71.94, 62.78, 55.11, 42.06, 41.51, 39.93, 25.60, 17.81, 12.50, −0.47, −0.49; ESI MS (M + Na⁺): 743.3 (calcd 742.99).

\[ N^3-(2-Chloroethyl)-5'-O-dimethoxytritylthymidine (3) \]

At room temperature, tetrabutylammonium fluoride (164 mg, 0.52 mmol) was added to a solution of \( N^3-(2-Chloroethyl)-5'-O-dimethoxytrityl-3'-O-(tert-butyldimethylsilyl)thymidine \) (2) (250 mg, 0.34 mmol) in THF (3 mL). The mixture was stirred at room temperature for 1.5 h, treated with a saturated solution of NH₄Cl, extracted with EtOAc, washed with brine and H₂O and dried over Na₂SO₄. Evaporation and silica gel column chromatography (hexane/EtOAc 3:1) gave 3 (188 mg, 89%) as a colorless oil. Rₛ (Hexane/EtOAc 3:1) 0.30; ^1H NMR (300 MHz, (CD₃)₂SO): δ 7.58 (s, 1H), 7.41 (dd, 2H, J = 7.2 Hz), 7.33–7.23 (m, 7H), 6.88 (d, 4H, J = 8.7 Hz), 6.24 (t, 1H, J = 6.8 Hz), 5.33 (br s, OH), 4.33 (m, 1H), 4.13 (t, 2H, J = 6.8 Hz), 3.90 (m, 1H), 3.73 (s, 6H), 3.71 (m, 2H), 3.21 (m, 2H), 2.32–2.15 (m, 2H), 1.52 (s, 3H); ^13C NMR (75 MHz, (CD₃)₂SO): δ 162.34, 158.14, 158.03, 150.12, 144.54, 135.29, 135.15, 134.57, 129.77, 129.59, 127.75, 127.54, 126.65, 113.23, 113.11, 108.56, 85.74, 85.51, 84.73, 70.22, 63.53, 54.91, 41.53, 40.20, 39.67, 12.68; ESI MS (M + Na⁺): 629.1 (calcd 629.2).

\[ N^3-(2-Chloroethyl)-5'-O-dimethoxytrityl-3'-O-(2-Cyano-ethyl-N,N-diisopropylphosphor-amidite) thymidine (4) \]

At room temperature, diisopropylethylamine (240 ml, 1.4 ml) was added to a solution of \( N^3-(2-Chloroethyl)-5'-O-dimethoxytritylthymidine \) (3) (322 mg, 0.53 mmol) in CH₂Cl₂ (3 ml), followed by chloro(2-cyanoethoxy)(N,N-diisopropylamino)phosphine (190 µl, 0.84 mmol). After 15 min, the reaction mixture was filtered through a short silica gel column (hexane/EtOAc 2:1), and fractions containing the desired product were combined and evaporated to dryness to afford 4 (mixture of diastereomers) as a white foam (389 mg, 90%). Rₛ (Hexane/EtOAc 3:1) 0.55; ^1H NMR (300 MHz, CCl₄): δ 7.62 (s, 1/2H), 7.58 (s, 1/2H), 7.37 (d, 2H, J = 8.8 Hz), 7.25 (m, 7H), 6.81 (d, 4H, J = 8.2 Hz), 6.39 (m, 1H), 4.63 (br s, 1H), 4.27 (t, 2H, J = 6.8 Hz), 6.14 (d, 1H, J = 12.17 Hz), 3.76 (s, 6H), 3.70 (m, 3H), 3.54 (m, 4H), 3.30 (m, 1H), 2.30–2.60 (m, 4H), 1.44 (s, 3H), 1.15 (s, 9H, J = 5.9), 1.05 (d, 3H, J = 6.4 Hz); ^31P NMR (160 MHz, CCl₄): δ 149.65, 149.16; ESI MS (M⁺): 806.5 (calcd 806.32).

\[ N^3-(2-Iodoethyl)-5'-O-dimethoxytrityl-3'-O-(tert-butyldimethylsilyl)thymidine (11) (20) \]

At room temperature, NaI (83 mg, 0.56 mmol) was added to a solution of \( N^3-(2-Chloroethyl)-5'-O-dimethoxytrityl-3'-O-(tert-butyl-dimethylsilyl)thymidine \) (2) (213 mg, 0.29 mmol) in acetonitrile (3 ml). The reaction mixture was heated at 70°C for 18 h, cooled to room temperature, diluted with EtOAc, washed with sodium thiosulfate (10%, 3 × 5 ml), dried over Na₂SO₄ and evaporated to dryness to afford 11 (207 mg, 86%) as a light yellow foam. Rₛ (Hexane/EtOAc 3:1) 0.69. Analytical data were in accordance with previous reports for 11 (20).

\[ I-(N^3-(2')-5'-O-(bis(tert-butyl-dimethylsilyl))-2-N-(dihydroxydimethylsilane))-2'-N^3-(5'-O-(dimethoxytrityl))-3'-O-(tert-butyl-dimethylsilyl)thymidylinyl) ethane (13) \]

At room temperature, Ag₂O (147 mg,
0.64 mmol) was added to a solution of N\textsuperscript{3}-(2-idoethyl)-
5'-O-dimethoxytrityl-3'-O-(tert-butylidemethylisilyl)thymidine
\textbf{11} (346 mg, 0.42 mmol) and 3',5'-O-(bis(tert-butylidemethylisilyl))-2'-N-((dimethylaminomethylene)-2'-deoxyguanosine
\textbf{12} (228 mg, 0.42 mmol) in DMF (5 mL). The reaction mixture
was heated at 45°C for 17 h, cooled to room temperature,
treated with a saturated solution of NH\textsubscript{4}Cl, extracted with
EtOAc. The organic layer was washed with brine and H\textsubscript{2}O and
dried over Na\textsubscript{2}SO\textsubscript{4}. Evaporation and silica gel column
chromatography (hexane:EtOAc 2:1) gave \textbf{13} (180 mg, 71%)
as a foamy. \textsuperscript{1}H NMR (300 MHz, (CD\textsubscript{3})\textsubscript{2}SO): δ 7.59 (d, 1H, J = 1.1 Hz), 7.40
(dd, 2H, J = 7.13 Hz), 7.34–7.22 (m, 7H), 6.90 (d, 4H, J = 8.57 Hz),
6.21 (t, 1H, J = 6.4 Hz), 4.73 (t, OH, J = 5.8 Hz), 4.46 (m, 1H), 3.90 (d, 2H, J = 6.2 Hz), 3.84 (dd, 1H, J = 4.1, 7.2 Hz),
3.70 (s, 6H), 3.49 (dd, 2H, J = 6.5, 12.4 Hz), 3.29 (dd, 1H, J = 2.6, 10.6 Hz), 3.19 (dd, 1H, J = 4.3, 10.6 Hz), 2.32 (m, 1H),
2.18 (m, 1H), 1.55 (s, 3H), 0.79 (s, 9H), 0.01 (s, 3H), 0.00 (s, 3H); \textsuperscript{13}C NMR (75 MHz, (CD\textsubscript{3})\textsubscript{2}SO): δ 162.75, 158.09, 150.23, 144.41,
135.18, 135.08, 134.11, 129.57, 127.72, 127.53, 126.69,
113.09, 108.53, 85.83, 84.55, 71.22, 62.71, 57.29,
54.91, 42.48, 39.41, 25.43, 17.43, 12.38, −4.49; ESI MS (M + Na\textsuperscript{+}): 725.8 (calcld 725.9).

1-(O\textsuperscript{5'}-3',5'-O-(bis(tert-butylidemethylisilyl))-2'-deoxyguanosinyl)-2'-N\textsuperscript{3}-O-((dimethylamino)methylene)-3'-O-(tert-butylidemethylisilyl)thymidylinyl)ethane \textbf{17}.
At 0°C, NaH (9.1 mg, 0.38 mmol) was added to a solution of N\textsuperscript{3}-(2-hydroxyethyl)-
5'-O-dimethoxytrityl-3'-O-(tert-butylidemethylisilyl)thymidine \textbf{15} (65 mg, 0.12 mmol) and 2-amino-6-chloro-9-[2'-deoxy-3,5-
bis-O-(tert-butylidemethylisilyl)]-β-D-ribofuranosyluridine \textbf{16} (29.30) (91 mg, 0.12 mmol) in DME. The reaction mixture
was left at room temperature for 1 h, treated with a saturated
solution of NH\textsubscript{4}Cl, extracted with EtOAc, washed with brine
and H\textsubscript{2}O and dried over Na\textsubscript{2}SO\textsubscript{4}. Evaporation and silica gel column
chromatography (hexane:EtOAc 3:1) gave \textbf{17} (110 mg, 74%) as a white foamy. \textsuperscript{1}H NMR (500 MHz, (CD\textsubscript{3})\textsubscript{2}SO): δ 8.01 (s, 1H), 7.62
(s, 1H), 7.38 (d, 2H, J = 7.4 Hz), 7.32–7.21 (m, 7H), 6.89
(d, 4H, J = 8.5 Hz), 6.34 (br s, 2H), 6.19 (t, 1H, J = 6.7 Hz),
6.14 (t, 1H, J = 6.5 Hz), 4.60 (t, 2H, J = 5.7 Hz),
4.50 (m, 1H), 4.45 (m, 1H), 4.25 (t, 2H, J = 5.5 Hz), 3.82
(m, 2H), 3.72 (s, 6H), 3.71 (d, 1H, J = 7.0 Hz), 3.63 (dd,
1H, J = 4.3, 10.8 Hz), 3.27 (dd, 1H, J = 1.9, 10.5 Hz),
3.16 (dd, 1H, J = 4.9, 10.5 Hz), 2.71 (m, 1H), 2.32–2.21
(m, 2H), 2.13 (m, 1H), 1.51 (s, 3H), 0.89 (s, 9H), 0.86 (s, 9H),
0.78 (s, 9H), 0.10 (s, 6H), 0.00 (s, 3H), −0.05 (s, 3H); \textsuperscript{13}C NMR (125 MHz, (CD\textsubscript{3})\textsubscript{2}SO): δ 162.59,
159.98, 159.61, 158.18, 154.04, 150.24, 144.48, 137.37,
135.26, 135.15, 134.41, 129.68, 129.64, 127.64, 126.79,
113.89, 113.20, 108.58, 86.95, 85.94, 85.30, 84.81,
82.31, 72.20, 71.24, 62.73, 62.28, 59.68, 55.01, 39.43,
39.33, 38.68, 25.72, 25.64, 25.54, 20.70, 19.71, 17.66,
17.53, 14.03, 12.41, −4.82, −4.99, −5.17, −0.55, −5.57;
ESI MS (M + Na\textsuperscript{+}): 1202.6 (calcld 1203.6).

1-(O\textsuperscript{5'}-3',5'-O-(bis(tert-butylidemethylisilyl))-2'-deoxyguanosinyl)-
2'-N\textsuperscript{3}-O-((dimethylamino)methylene)-3'-O-(tert-butylidemethylisilyl)thymidylinyl)ethane \textbf{18}.
At room temperature, tetraethylammonium fluoride (48 mg, 0.15 mmol) was added to a solution of 1-[O\textsuperscript{5'}-3',5'-
O-(bis(tert-butylidemethylisilyl))-2'-deoxyguanosinyl]-2'-N\textsuperscript{3}-O-((dimethylamino)methylene)-3'-O-(tert-butylidemethylisilyl)
thymidylinyl)ethane \textbf{17} (36 mg, 0.03 mmol) in THF (1 mL). The reaction mixture
was left at room temperature for 2 h. After evaporation of the solvent and purification by silica gel
column chromatography (EtOAc:MeOH 15:1) the product was
dissolved in MeOH (320 μl) and water (80 μl) followed by
acetic acid (1.6 ml). The reaction mixture was left at room
temperature for 3 h and the solvent was then evaporated. The crude product was dissolved in H2O, washed with EtOAc (3 × 2 ml). Evaporation of the water gave 18 (14 mg, 85%) as a white foam. \(R_f\) (CH3Cl2/MethOH 5:1) 0.40; \(^1H\) NMR (300 MHz, (CD3)2SO): \(\delta\) 8.06 (s, 1H), 7.77 (s, 1H), 6.34 (br s, 2H), 6.19 (dd, 1H, J = 6.2, 7.6 Hz), 6.16 (t, 1H, J = 7.0 Hz), 5.26 (br s, OH, J = 3.8 Hz), 5.23 (br s, OH), 5.00 (br s, OH), 4.99 (t, OH, J = 4.9 Hz), 4.61 (m, 1H), 4.35 (m, 1H), 4.29 (dd, 1H, J = 5.5, 13 Hz), 4.25 (dd, 1H, J = 5.4, 11 Hz), 4.22 (m, 1H), 3.82 (m, 1H), 3.76 (m, 1H), 3.59 (m, 1H), 3.55 (m, 2H), 3.50 (m, 1H), 2.58 (m, 1H), 2.20 (m, 1H), 2.1–2.0 (m, 2H), 1.80 (s, 3H); \(^{13}C\) NMR (75 MHz, (CD3)2SO): \(\delta\) 162.65, 159.99, 159.50, 153.91, 150.43, 137.73, 134.89, 113.91, 108.41, 87.58, 87.37, 84.79, 82.74, 70.74, 70.21, 62.31, 61.72, 61.16, 39.55, 39.47, 39.34, 12.87; ESI MS (M + Na\(^+\)): 558.2 (calcd 558.5).

Oligonucleotides synthesis and purification

All oligonucleotides were synthesized on a PerSpective Bio-systems Expedite 8909 DNA synthesizer on a 1 μmol scale using the standard ultra-mild protocol. Ultra-mild phosphoramidites and Q support were purchased from Glen Research. 6-carboxy fluorescein (6-FAM) phosphoramidite was purchased from Biosearch Technologies. All phosphoramidites were dissolved to a concentration of 0.1 M in anhydrous CH3CN. Phenoxyacetic anhydride (0.25 M) in THF was chased from Biosearch Technologies. All phosphoramidites were used as a capping reagent. For the modified phosphoramidite the coupling time was extended to 15 min. The coupling yield of the modified phosphoramidite was >97%. The oligonucleotide 5 was removed from the solid support and deprotected by treatment with diisopropylamine/MeOH (1:10) for 17 h at room temperature, and purified by reverse phase HPLC using the following gradient: linear 0–10% B over 10 min, linear 10–20% B until 22 min, linear 20–25% B until 29 min, linear 25–35% B until 33 min, isocratic 100% B until 39 min, isocratic 0% B until 45 min; buffer A: 0.1 M TEAA (pH 7); and buffer B: CH3CN. The resulting peaks were identified by co-injection with the corresponding standards and eluted at the following times: dC (6.1 min), dG (13.0 min), T (14.3 min), dA (20.2 min), 6-FAM (28.5 min), cross-link dimer 18 (37.2 min), 6-FAM (41.7 min).

RESULTS AND DISCUSSION

Design of a sequence-specific alkylation reaction for the generation of DNA interstrand crosslinks

We set out to design a strategy that would allow for the incorporation of a precursor of an ethylene-bridged ICL into one strand of DNA using solid phase synthesis and specific alkylation reaction to a complementary strand following annealing. In our design we first considered the reaction of the CENUs with DNA itself. CENUs initially react with the oxygen of guanine base to form \(O^6\)-chloroethyguanine, which then undergoes intramolecular cyclization to give rise to \(N^1, O^6\)-ethanoguanine (Figure 1A). This positively charged intermediate is highly reactive and can undergo nucleophilic attack by the N-3 of cytosine located on the opposite strand of DNA (14). We reasoned that although the \(O^6\)-chloroethyguanine or \(N^1, O^6\)-ethanoguanine intermediates would likely be too unstable to be incorporated into DNA by solid phase synthesis (31), it might be possible to mimic this process in an alkylation reaction with a different reactive precursor in one strand. In a related study at the (mono)nucleoside level, Miller and coworkers synthesized a thymidine dimer crosslinked at their N-3 positions through an ethylene bridge using a 3-iodoethylthymidine derivative as a precursor, which they subsequently introduced into DNA as a dimer (20).

We reasoned that it might be possible to use a similar reaction in the context of dsDNA. Preliminary studies using 1-(2-chloroethyl)-guanosine and 3-(2-chloroethyl)-cytosine (Figure 1B) revealed that these molecules undergo rapid cyclization to the \(N^1, N^6\)-ethanoguanine and \(N^3, N^6\)-ethanocytosine, respectively, and would be unsuitable precursors for a C-G crosslink. In contrast, it appeared that the corresponding 3-(2-chloroethyl)thymidine would be a suitable ICL precursor with a stability profile that might allow its incorporation into DNA using solid phase synthesis (see below). We therefore set out to incorporate 3-(2-chloroethyl)thymidine into DNA and test its ability to mediate site-specific ICL formation.

Preparation of N-3-chloroethylthymidine and its incorporation into DNA

The synthesis of phosphoramidite 4 containing the chloroethyl thymidine moiety was accomplished as shown in Figure 2. Alkylation of 5'-O-dimethoxytrityl-3'-O-tert-butylidemethylsilylthymidine 1 (28) at N-3 position using Cs2CO3 and bromochloroethane gave N-3-chloroethylthymidine product 2 (32). The silyl group was removed from 2 using...
tetrabutylammonium fluoride to give alcohol 3, which was converted to the phosphoramidite 4. The building block 4 was thus obtained in a straightforward manner in 68% overall yield from 1.

Since we were concerned whether the chloroethyl group would be stable toward base induced hydrolysis, we tested its stability toward very mild basic deprotection conditions that we had previously developed (33). Chloroethyl 3 proved to be stable toward hydrolysis upon extended treatment in diisopropylamine/methanol (1:10) (data not shown). We therefore proceeded to incorporate phosphoramidite 4 into oligonucleotides using ‘ultramild’ phosphoramidites (iPrPac-dG, Pac-dA and Ac-dC) and Q support (34) and subsequently removed the oligonucleotide from solid support and removed base and phosphodiester bond-protecting groups by treatment with a 10% solution of diisopropylamine in methanol at room temperature for 17 h. The main product was purified by HPLC and analyzed by nano-ESI to confirm its identity and purity. The observed molecular weight (5366.2) was found to be consistent with the calculated molecular weight (5367.0) for oligonucleotide 5.

Site-specific ICL formation

With the chloroethyl thymine-containing oligonucleotide 5 in hand, we tested its ability to form interstrand crosslinks with complementary oligonucleotides. We annealed 5 to a complementary oligonucleotide 6 containing a 6-carboxyfluorescein (6-FAM) at the 5’-end and a G opposite to the reactive group. Following incubation of the two oligonucleotides at room temperature, we observed the formation of a band with slower migration in a 20% denaturing PAGE gel, indicative of ICL formation. The band increased over the course of 2 weeks, eventually yielding ~20% of the FAM-labeled oligonucleotide in crosslinked form (Figure 3B, lanes 2–4). The reaction was of similar efficiency in a variety of buffers (ammonium acetate, sodium acetate, Hepes·KOH, Tris–HCl, ammonium chloride) and over a wide pH range (5–12) (data not shown). We carried out our studies in Hepes·KOH at pH 6, which appeared to afford the cleanest reaction. Although chloride 5 was stable in aqueous solution at room temperature for several days, after 15 days we observed an increasing amount of the corresponding alcohol by mass spectrometry (data not shown). We then studied the influence of the concentrations of oligonucleotides 5 and 6 on the efficiency of the cross-linking reaction. We observed an increase in reaction efficiency up to about five equivalents of oligonucleotide 5 over 6 (Figure 3B, lanes 5–9). Under these conditions, we consistently achieved a cross-linking yield of 25–30% based on the unmodified strand (Figures 3B and 4). We then compared the ability of 5 to react with the four DNA bases. Oligonucleotide 5 exhibited the highest yield of crosslink formation if paired with guanine with the reactivity decreasing in the order G > C > A > T (Figure 3B, lanes 10–13).

Characterization of cross-linked oligonucleotide 10

The main crosslinked product between oligonucleotides 5 and 6 was purified by HPLC (Figure 4B) and its mass determined by nano-ESI-MS. The observed molecular weight (10957.0) was found to be consistent with the calculated mass (10957.4) of the cross-linked product. Since two products of identical mass can be formed by reaction of the chloride in 5 with either the N-1 or O-6 position of the guanine base (Figure 1C), we wished to determine the structure of the product.
crosslink by nucleoside composition analysis and comparison of the digested product with an authentic standard. We therefore needed to synthesize the cross-linked dimers 14 and 18, containing dG alkylated at N-1 and O-6, respectively (Figure 5). Initial attempts to alkylate the N-1 position of the protected-2'-deoxyguanosine 12 (32) with chloroethyl compound 2 using different alkylating conditions were not successful and only the starting materials were isolated. We reasoned that improving the reactivity of the alkylating agent might lead to success and replaced the chloride in the alkylation of iodide 11 and protected 6-chloropurine 16 (29,30) in DME using NaH as a base (Figure 5). Other attempts, using DMF as a solvent gave a complex mixture of products, whereas a Mitsunobu reaction between hydroxyethyl 15 and 3',5'-O-(bis(tert-butyldimethylsilyl))-2'-deoxyguanosine yielded no detectable product. The fully deprotected O6-linked dimer 18 was obtained by desilylation followed by acid hydrolysis of the trityl group.

Although the two dimers 14 and 18 have identical signals for their molecular weight ions in MS analysis, they can be distinguished by 13C-NMR spectroscopy based on their signals for the methylene groups. The 13C-NMR spectrum of dimer 14 shows signals at 61.66 and 60.92 p.p.m., corresponding to the two O-CH2 groups of the C-5' deoxyribose, and signals at 39.62, 39.43, 38.38, 38.28 p.p.m., corresponding to N-CH2CH2-N and C-2' of the two deoxyribose residues, respectively. The 13C-NMR spectrum of compound 18 showed signals at 62.31, 61.72, 61.16 p.p.m., indicating the presence of three O-CH2 groups: two for C-5' of the deoxyribose, and one for the O-ethylene bridge; the other three signals, at 39.55, 39.47, 33.34 p.p.m., correspond to N-CH2 and the C-2' of the two deoxyribose moieties.

The availability of the authentic standards 14 and 18 allowed us to identify the structure of the ICL in oligonucleotide 10. Cross-link oligonucleotide 10 was enzymatically digested, phosphorylated and analyzed by HPLC (Figure 6B). In addition to the four natural nucleosides dC, dG, T and dA, we observed two additional peaks; the first one corresponded to the fluorescent label at the 5' end of the G containing oligonucleotide. The second one co-eluted with the O6-linked dimer 18 with a retention time of 37.2 minutes, clearly distinct from 14, which eluted at 35.6 min. The nucleoside composition analysis therefore establishes that the crosslink is formed exclusively between N-3 of T and O-6 of dG.

**Mechanism of ICL formation between N-3 of T and O-6 of G**

The exclusive reaction of the chloroethyl group of T with O-6 of dG and absence of reaction with N-1 of dG deserves some comment. It appears that the chlorine in 3-(2-chloroethyl)thymidine is very inert due to the two electron-withdrawing groups in the pyrimidine ring, rendering it unreactive toward direct substitution by oxygen or nitrogen nucleophiles. This is illustrated by the extremely low reactivity of 2 toward nucleophilic substitution under aprotic conditions (Figure 5). Likewise, a direct S-N2 substitution of the chloride in the context of dsDNA is not a likely event. However, under solvolytic conditions, the formation of the ionic intermediate 19 (Figure 7) may occur at low frequency due to the intramolecular nature of this cyclization reaction. The ionic character of 19 causes it to react exclusively with the oxygen of the amide bond in dG, due to the latter’s partial anionic character. This mechanism accounts for the very slow ICL formation, because the frequency of the formation of 19 is very low. Once it is formed it is likely to react very quickly with the closest nucleophile with anionic character, in this case O-6 of dG.

This preferential reactivity of oxygen-centered rather than nitrogen-centered nucleophiles with electrophiles with ionic...
character is also observed in the reaction of alkylating agents with the DNA bases. Alkylating agents such as N-methyl-N-nitrosourea (NMU), which react with an SN1-like mechanism with cationic character in the transition state, are highly oxophilic and show high reactivity with O-6 of dG (35,36). Less reactive alkylating agents that react by an SN2-type mechanisms without significant positive charge in the transition state instead prefer to react with nitrogen-based nucleophile in DNA such as N-7 of G.

CONCLUSIONS
We have developed a simple and concise approach for the synthesis of an ethylene bridged DNA interstrand crosslink. For this purpose we have synthesized an oligonucleotide carrying a chloroethyl moiety at a defined T base poised to...
react selectively with a base on a complementary DNA strand. To our knowledge this constitutes the first example of an oligonucleotide with an alkyl chloride side chain and one of a few examples of a specific ICL formation using a reactive group on one strand of DNA (21–23). Since oligonucleotides containing 2-chloroethylthymidine have in principle the ability to react with different types of DNA in a sequence-specific manner, they might also be generally useful for the sequence-specific covalent modification of DNA. Given the slow kinetics of the cross-linking reaction, it remains to be seen whether it will be useful in the context of gene-modification using antisense or triple-helix approaches (37–39).

SUPPLEMENTARY DATA
Supplementary data are available at NAR online.

ACKNOWLEDGEMENTS
The authors are grateful to Thinh Ho, Angelo Guainazzi and in particular Prof. Francis Johnson for comments on the manuscript. This work was supported by the European Commission (LSHB-CT-2004-005204), the Roche Research Foundation (126-2003), the Swiss National Science Foundation (3130-054873.98) and the New York State Office of Science and Technology and Academic Research (NYSTAR) grant No. C040069. Funding to pay the Open Access publication charges for this article was provided by was provided by the European Commission (LSHB-CT-2004-005204).

Conflict of interest statement. None declared.

REFERENCES


