A Diazirine–based Nucleoside Analogue for Efficient DNA Interstrand Photocross–Linking

Zhihai Qiu, Lianghua Lu, Xing Jian and Chuan He*

Department of Chemistry, The University of Chicago, 5735 South Ellis Avenue, Chicago, Ilinois 60637 E-mail: <u>chuanhe@uchicago.edu</u>

Supporting Information

Synthesis of Phosphoramidite 7

General Experiment Methods. Reagents were obtained from commercial sources and used as received unless otherwise noted. Water-sensitive reactions were performed in oven-dried glassware under an atmosphere of dry nitrogen. Anhydrous solvent CH_2Cl_2 was distilled over CaH_2 under an atmosphere of dry nitrogen prior to use. All other anhydrous solvents were used directly without further distillation. Thin-layer chromatography was performed on EM Science silica gel 60 F254 plates (250 μ m) and visualized by UV or KMnO₄ solution. Column chromatography was performed using silica gel (40-63 μ m). Et₃N was added into eluent to help purify acid sensitive compounds in some cases. NMR spectra were recorded on Bruker Avance 400 or 500 MHz instruments as solutions in CDCl₃ unless otherwise indicated. Chemical shifts (δ) are reported in parts per million (ppm), and are referenced to tetramethylsilane for ¹H and ¹³C NMR, H₃PO₄ for ³¹P NMR. Coupling constants (*J*) are reported in hertz (Hz). High-resolution mass spectroscopy was performed by Mass Spectrometry Facility at University of Notre Dame.

(2R,5R)-2-((tert-Butyldimethylsilyloxy)methyl)-5-(4-(3-(trifluoromethyl)-3H-diazirin-3-

yl)phenyl)dihydrofuran-3(2H)-one (3). To an oven-dried round-bottom flask is added a suspension of

1,4-anhydro-3,5-bis-O-(*tert*-butyldimethylsilyl)-2-deoxy-D-*erythro*-pent-1-enitol 1^[1] (0.488 g, 1.42 mmole), $3-(4-iodophenyl)-3-(trifluoromethyl)-3H-diazirine 2^{[2]}$ (0.368 g, 1.18 mmole), N,Ndicyclohexylmethylamine (0.30 mL, 1.4 mmole), palladium(II) acetate (0.026 g, 0.12 mmole), tetrabutylammonium chloride (0.328 g, 1.18 mmole) and crushed 4Å molecular sieves (0.4 g) in dry N,N-dimethylformamide (3.0 mL). The mixture is stirred at room temperature and degassed with nitrogen bubbling for 30 minutes, then heated to 60 °C for 40 hours under nitrogen atmosphere. The reaction mixture is then diluted with water (50 mL) and extracted with ethyl ether (1×50 mL). The organic layer is washed with saturated aqueous NaHCO₃ (50 mL) and brine (50 mL) sequentially, dried over anhydrous MgSO₄ and concentrated *in vacuo*. The residue is purified by flash column chromatography on silica gel (Hexanes : Ethyl ether = 4:1) to afford (2R,5R)-2-((tertbutyldimethylsilyloxy)methyl)-5-(4-(3-(trifluoromethyl)-3H-diazirin-3-yl)phenyl)dihydrofuran-3(2H)one **3** as an oil (0.225 g, 46% yield). ¹H NMR (500 MHz, CDCl₃): δ 7.54 (d, J = 8.4 Hz, 2H), 7.21 (d, J = 8.4 Hz, 2H), 5.19 (dd, J = 11.0, 5.9 Hz, 1H), 4.04 (t, J = 2.5 Hz, 1H), 3.97 (m, 2H), 2.82 (dd, J = 17.6, 5.9 Hz, 1H), 2.38 (dd, J = 17.6, 11.0 Hz, 1H), 0.87 (s, 9H, C(CH₃)₃), 0.08 (s, 3H, CH₃), 0.06 (s, 3H, CH₃). ¹³C NMR (126 MHz, CDCl₃): δ 213.5, 142.9, 129.1, 126.8 (q, J = 1.1 Hz), 126.7, 122.2 (q, J = 1.1 Hz) 274.7 Hz), 82.8, 76.9, 62.7, 46.3, 28.4 (q, J = 40.5 Hz), 25.9 (SiC(CH₃)₃), 18.3 (SiC(CH₃)₃), -5.4 (Si-CH₃), -5.6 (Si-CH₃). HRFABMS m/z: [M]⁺ calcd for C₁₉H₂₆F₃N₂O₃Si 415.1665, found 415.1674.

(2R,5R)-2-(Hydroxymethyl)-5-(4-(3-(trifluoromethyl)-3H-diazirin-3-yl)phenyl)dihydrofuran-

3(2*H***)-one (4).** To a solution of (2R,5R)-2-((*tert*-butyldimethylsilyloxy)methyl)-5-(4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)phenyl)dihydrofuran-3(2*H*)-one **3** (0.460 g, 1.11 mmole) in tetrahydrofuran (5 mL), which is chilled to 0 °C in an ice-water bath, are added glacial acetic acid (0.2 mL) and 1.0 M tetrabutylammonium fluoride solution in tetrahydrofuran (3.0 mL). The reaction

mixture is stirred at 0 °C for 1.5 hours then concentrated *in vacuo*. The residue is purified by flash column chromatography on silica gel (Hexanes : Ethyl ether = 1:4) to afford (2*R*,5*R*)-2-(hydroxymethyl)-5-(4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)phenyl)dihydrofuran-3(2*H*)-one **4** as an oil (0.308 g, 92% yield). ¹H NMR (500 MHz, CDCl₃): δ 7.49 (d, *J* = 8.1 Hz, 2H), 7.24 (d, *J* = 8.1 Hz, 2H), 5.24 (dd, *J* = 11.0, 5.9 Hz, 1H), 4.06 (t, *J* = 3.3 Hz, 1H), 3.96 (m, 2H), 2.90 (dd, *J* = 18.0, 5.9 Hz, 1H), 2.48 (dd, *J* = 18.0, 11.0 Hz, 1H). ¹³C NMR (126 MHz, CDCl₃): δ 213.4, 141.9, 129.2, 126.9 (q, *J* = 1.1 Hz), 126.7, 122.1 (q, *J* = 274.8 Hz), 82.6, 76.9, 61.4, 45.4, 28.3 (q, *J* = 40.5 Hz). HRFABMS *m/z*: [M]⁺ calcd for C₁₃H₁₂F₃N₂O₃ 301.0800, found 301.0756.

(1R)-1,4-Anhydro-2-deoxy-1-C-(4-(3-(trifluoromethyl)-3H-diazirin-3-yl)phenyl)-D-erythro-pentitol

(5). То solution of (2R,5R)-2-(hydroxymethyl)-5-(4-(3-(trifluoromethyl)-3H-diazirin-3a yl)phenyl)dihydrofuran-3(2H)-one 4 (0.300 g, 0.999 mmole) in acetonitrile (10 mL) and glacial acetic acid (2 mL), which is chilled to -10 °C in an ice-salt bath, is added sodium triacetoxyborohydride (0.529 g, 2.50 mmole) in one portion. The reaction mixture is stirred at -10 °C for 30 minutes under nitrogen atmosphere then concentrated in vacuo. The residue is purified by flash column chromatography on silica gel (CH₂Cl₂: MeOH = 7:1) to afford (1*R*)-1,4-anhydro-2-deoxy-1-C-(4-(3-(trifluoromethyl)-3*H*diazirin-3-yl)phenyl)-D-erythro-pentitol 5 as a white powder (0.170 g, 56% yield). ¹H NMR (500 MHz, CDCl₃): δ 7.37 (d, J = 8.4 Hz, 2H), 7.17 (d, J = 8.4 Hz, 2H), 5.16 (dd, J = 9.9, 5.5 Hz, 1H), 4.40 (m, 1H), 4.02 (m, 1H), 3.75 (m, 2H), 2.24 (m, 1H), 1.95 (m, 1H). ¹³C NMR (126 MHz, CDCl₃): δ143.3, 128.6, 126.7 (q, J = 1.1 Hz), 126.5, 122.2 (q, J = 274.7 Hz), 87.5, 79.5, 73.5, 63.3, 43.7, 28.4 (q, J = 1.1 Hz) 40.5 Hz). HRFABMS m/z: [M]⁺ calcd for C₁₃H₁₄F₃N₂O₃ 303.0957, found 303.0925.

Confirmation of anomeric geometry



β-anomer

H1' (5.18 ppm)	NOE observed at H4' (4.00 ppm), H2' α (2.10
	ppm) and H2' β (1.85 ppm)
H5' (3.70 ppm)	NOE observed at H3' (4.38 ppm), H2' β (1.85
	ppm)

For reference see J. Org. Chem. 2005, 70, 1132-1140.



(1R)-1,4-Anhydro-2-deoxy-5-O-(4,4'-dimethoxy)trityl-1-C-(4-(3-(trifluoromethyl)-3H-diazirin-3-

vl)phenvl)-D-erythro-pentitol (6). А solution of (1R)-1,4-anhydro-2-deoxy-1-C-(4-(3-(trifluoromethyl)-3H-diazirin-3-yl)phenyl)-D-erythro-pentitol 5 (0.170 g, 0.562 mmole) in dry pyridine (5 mL) is stirred at 0 °C for 10 minutes, followed by the addition of 4,4'-dimethoxytrityl chloride (0.286 g, 0.844 mmole) and 4-(dimethylamino)pyridine (0.004 g, 0.03 mmole) in one portion. The reaction solution is stirred at 0 °C under nitrogen atmosphere for 6 hours then diluted with water (50 mL) and extracted with ethyl ether $(2 \times 50 \text{ mL})$. The combined organic layer is washed with saturated aqueous NaHCO₃ (50 mL) and brine (50 mL) sequentially, dried over anhydrous MgSO₄ and concentrated in *vacuo*. The residue is purified by flash column chromatography on silica gel (gradiently from Hexanes : Ethyl ether = 3:1 to Ethyl ether with 4% TEA) to afford (1R)-1,4-anhydro-2-deoxy-5-O-(4,4'dimethoxy)trityl-1-C-(4-(3-(trifluoromethyl)-3H-diazirin-3-yl)phenyl)-D-erythro-pentitol 6 as a white powder (0.242 g, 71% yield). ¹H NMR (500 MHz, CDCl₃): δ 7.44 (m, 2H), 7.39 (d, J = 8.4 Hz, 2H), 7.33 (m, 4H), 7.26 (m, 2H), 7.21 (m, 1H), 7.14 (d, J = 8.4 Hz, 2H), 6.81 (m, 4H), 5.17 (dd, J = 9.9, 5.5 Hz, 1H), 4.40 (m, 1H), 4.06 (m, 1H), 3.78 (s, 6H), 3.33 (dd, J = 9.9, 4.4 Hz, 1H), 3.27 (dd, J = 9.9, 5.5 Hz, 1H), 2.24 (ddd, J = 13.2, 5.5, 1.8 Hz, 1H), 1.97 (ddd, J = 13.2, 9.9, 6.0 Hz, 1H). ¹³C NMR (126) MHz, CDCl₃): δ158.6, 144.9, 144.1, 136.1, 136.1, 130.2, 128.3, 128.0, 127.0, 126.6, 126.5, 122.3 (q, J = 274.8 Hz), 113.3, 86.6, 86.4, 79.4, 74.6, 64.5, 55.3, 43.9, 28.5 (q, J = 40.1 Hz). HRFABMS m/z: [M]⁺ calcd for C₃₄H₃₁F₃N₂O₅ 604.2185, found 604.2188.

(1*R*)-1,4-Anhydro-2-deoxy-5-*O*-(4,4'-dimethoxy)trityl-1-*C*-(4-(3-(trifluoromethyl)-3*H*-diazirin-3yl)phenyl)-D-*erythro*-pentitol 2-cyanoethyl diisopropylphosphoramidite (7). A solution of (1*R*)-1,4anhydro-2-deoxy-5-*O*-(4,4'-dimethoxy)trityl-1-*C*-(4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)phenyl)-D*erythro*-pentitol **6** (0.121 g, 0.200 mmole), N,N-diisopropylethylamine (0.10 mL, 0.57 mmole) and 2cyanoethyl N,N-diisopropylchlorophosphoramidite (0.07 mL, 0.3 mmole) in dry dichloromethane (2 mL) is stirred at 0 °C under nitrogen atmosphere for 1.5 hours. The reaction solution is directly purified by flash column chromatography on silica gel (Hexanes : Ethyl ether = 2:1 with 4% TEA) to afford (1*R*)-1,4-anhydro-2-deoxy-5-*O*-(4,4'-dimethoxy)trityl-1-*C*-(4-(3-(trifluoromethyl))-3*H*-diazirin-3-yl)phenyl)-D-*erythro*-pentitol 2-cyanoethyl diisopropylphosphoramidite **7** as a white foam (0.136 g, 84% yield). ¹H NMR (500 MHz, CDCl₃): δ 7.46-7.43 (m, 4H), 7.35-7.25 (m, 9H), 6.83-6.79 (m, 4H), 5.25-5.15 (m, 1H), 4.55-4.45 (m, 1H), 4.25-4.15 (m, 1H), 3.90-3.20 (m, 12H), 2.70-1.90 (m, 4H), 1.35-1.05 (m, 12H). ³¹P NMR (202 MHz, CDCl₃): δ 148.7, 148.6.

Oligonucleotide Synthesis and Characterization

Oligonucleotides were synthesized on Applied Biosystems Expedite Nucleic Acid synthesizer using standard β -cyanoethyl phosphoramidite chemistry. All oligonucleotides were synthesized in DMT-off mode on the 5'-end and deprotected from CPG (alkyl controlled pore glass) supports in concentrated ammonia overnight at room temperature. Trimer oligonucleotides were purified using Phenomenex Jupiter 4 μ Proteo 90Å column (250 × 10.00 mm) with a linear gradient of 5%-95% buffer B (50 mM ammonium acetate in 50:50 (v/v) CH₃CN/H₂O) in buffer A (50 mM ammonium acetate in H₂O). 15-mer oligonucleotides were purified by preparative 20% denaturing polyacrylamide gel electrophoresis (PAGE) and isolated by excision and elution from gel. The recovered material was subsequently quantified by UV absorption at 260 nm, assuming the molar extinction coefficient of the diazirine base analogue to be zero at 260 nm. All synthesized oligodeoxynucleotides were confirmed by MALDI mass spectroscopy.

Melting Temperature (T_m) Measurement

Melting temperatures of double-stranded DNA's were measured using MicroCal VP Differential Scanning Calorimeter (DSC) (Northampton, MA) with 4 μ M concentration of dsDNA in 10 mM Tris-HCl (pH 7.4) and 100 mM NaCl buffer. The samples were scanned relative to the reference buffer over the temperature range of 20-75 °C at a rate of 90 K•h⁻¹. The sample DSC curves were adjusted from the reference buffer and melting temperatures were determined directly from the DSC curves.

dsDNA Interstrand Photocross-link

The solutions of dsDNA samples (30 μ M) in 10 mM Tris-HCl (pH 7.4) and 100 mM NaCl were irradiated at 0 °C for varying periods of time using a 450-W mercury vapor lamp (Ace Glass Incorporated) with a glass filter to remove UV lights < 300 nm. The samples were then denatured and analyzed using 20% DNA polyacrylamide gel and visualized using SYBR[®] Gold nucleic acid gel stain (Invitrogen). The images were taken using BIO-RAD Molecular Imager FX.

Hydroxyl Radical Footprinting

To radioactively label the cross-linked product, the strand with DBN was prepared with a protrusion of G at the 5' end. After annealing, photocross-linking and purification of the dsDNA, 25 pmol product was labeled with α -P-32-dCTP by Klenow fragment (exo-) at 37 °C for 30 min. The reaction mixture was then extracted by Phenol-Chloroform and purified by spin column. Reagents of 2 µL of 0.1 M DTT, 2 µL of 1% H₂O₂, and 2 µL of Fe(II)-EDTA (equal volumes of 2 mM ammonium iron (II) sulfate hexahydrate and 4 mM EDTA) were added to 20 µL labeled product (0.2 pmol) and rapidly mixed. The reaction was incubated at 37 °C for 10-20 min and quenched by 5 µL of 100 mM thiourea. The mixture

was evaporated by speed-vacuum, dissolved in 10 μ L of DNA loading buffer and loaded to 20% urea PAGE gel with markers.





Figure S1. MALDI-TOF mass spectrum of the trimer TBT in 2',4',6'-trihydroxyacetophenone (THAP, MW 168) matrix. The peaks at 879, 907, and 1046 correspond to (M–N₂), M, and (M–N₂+THAP), respectively.

Voyager Spec #1[BP = 896.6, 6764]



Figure S2. MALDI-TOF mass spectrum of the UV-irradiated trimer TBT in 2',4',6'-trihydroxyacetophenone (THAP, MW 168) matrix. The peak at 896 corresponds to $(M-N_2+H_2O)$.

Voyager Spec #1[BP = 4660.9, 8320]



Figure S3. MALDI-TOF mass spectrum of the 15-mer oligonucleotide (5'-ATG AAC C<u>B</u>G GAA AAC-3') in 3hydroxypyridine-2-carboxylic acid (3-HPA, MW 139) matrix. The peaks at 4633, 4650, 4661, and 4772 correspond to $(M-N_2)$, $(M-N_2+H_2O)$, M, and $(M-N_2+HPA)$, respectively.



Figure S4: PAGE analysis of photocross–linked products of the diazirine base analogue paired with A, T, G, and C, respectively (30 μ M). B represents the 15-mer ssDNA that contains the diazirine base analogue (5'-ATG AAC C**B**G GAA AAC-3'); A, T, G, and C represent the complementary ssDNA strands shown in Table 1. Method I: Reaction mixture without separation. Method II: The cross-linked product purified by urea denature DNA gel.



Acquired: 16:50:00, August 17, 2007

Figure S5. MALDI-TOF mass spectrum of the cross-linked band of 15-mer oligonucleotide B (5'-ATG AAC C<u>B</u>G GAA AAC-3') with the complementary ssDNA strands A in 3-HPA matrix.



Figure S6. MALDI-TOF mass spectrum of the cross-linked band of 15-mer oligonucleotide B (5'-ATG AAC C<u>B</u>G GAA AAC-3') with the complementary ssDNA strands T in 3-HPA matrix.



Figure S7. MALDI-TOF mass spectrum of the cross-linked band of 15-mer oligonucleotide B (5'-ATG AAC C<u>B</u>G GAA AAC-3') with the complementary ssDNA strands G in 3-HPA matrix.



Figure S8. MALDI-TOF mass spectrum of the cross-linked band of 15-mer oligonucleotide B (5'-ATG AAC C<u>B</u>G GAA AAC-3') with the complementary ssDNA strands C in 3-HPA matrix.

Figure S9: Hydroxyl radical footprinting of the cross-linked products of 42-mer dsDNAs with the diazirine base analogue B paired with T. The cross-linked products were excised from the gel, eluted, ³²P-labelled at the 3'-end of the complementary strands, treated with $H_2O_2/Fe^{2+}(EDTA^4)$ and reanalyzed by PAGE. Sample A represents the product of a 42-mer ssDNA containing the diazirine base analogue (5'-GCC CTT TAC ATC TTA GGA TTC T**B**T ATG CGC GGG AGA GTG TAT-3') cross-linked with its complementary strand; Sample B represents the product of a different 42-mer ssDNA (5'-GCC CTT TAC ATC TTA GGA TTC C**B**G ATG CGC GGG AGA GTG TAT-3'') cross-linked with the complementary strand. Line 1: Marker; Line 2: control of sample A; Line 3: sample A was treated with $H_2O_2/Fe^{2+}(EDTA^4)$ for 10 min; Line 4: sample A was treated with $H_2O_2/Fe^{2+}(EDTA^4)$ for 20 min; Line 5: control of sample B; Line 6: sample B was treated with $H_2O_2/Fe^{2+}(EDTA^4)$ for 20 min; Line 7: sample B was treated with $H_2O_2/Fe^{2+}(EDTA^4)$ for 10 min. Based on the footprinting result, two relatively strong fragments of 21 mer and 23 mer were observed, which indicates that the cross-link occurs mostly at positions 22 and 24 (3' and 5' adjacent sites of T opposite of DBN) of the complementary strand. There was almost no longer fragment observed indicating that the cross-linking is confined to the 3 nearby bases opposite of DBN.

-

