Supporting Information: Highly Selective and Efficient Interstrand Cross-Linking Formation to Thymine Without Photo-Irradiation

Keiichi Hattori, Tomoya Hirohama, Shuhei Imoto, Shuhei Kusano, Fumi Nagatsugi*

Institute of Multidisciplinary Research for Advanced Materials, Tohoku University, 2-1-1 Katahira, Aoba-ku,

Sendai-shi, Miyagi, 980-8577, Japan

Contents

- 1. Experimental procedures (S1~S6)
- 2. Supporting Information Figure 1S. ¹H NMR of **6** (S7)
- 3. Supporting Information Figure 2S. ¹³C NMR of **6** (S8)
- 4. Supporting Information Figure 3S. ¹H NMR of 4-Phenoxyacetylamino-1-[2-(2'deoxy-3'-O-methoxymethyl-5'-trityl-b-D-ribofuranosyl)ethyl] -2-(2-octylthio ethyl)-6-oxopyrimidine. (**S9**)
- 5. Supporting Information Figure 4S. ¹H NMR of **7** (S10)
- 6. Supporting Information Figure 5S. ¹³C NMR of **7** (S11)
- 7. Supporting Information Figure 6S. ¹H NMR of 8 (S12)
- 8. Supporting Information Figure 7S. ³¹P NMR of **8** (S13)

9. Supporting Information Figure 8S. Fluor images of 20% denaturing PAGE analysis of ISC product with the target DNA (**12a**) or RNA (**12b**) (**S14**)

10. Supporting Information Figure 9S. Fluor images of 20% denaturing PAGE of Fe•EDTA treatment of cross-linked product obtained from **11**. (S15)

11. Supporting Information Figure 10S. MALDI-TOFMS of the cross-linking adducts (17). (S16)

12. Supporting Information Figure 11S. HRMS (ESI) of the product (**18**) from enzyme digestion of cross-linking adduct (**S17**)

13. Supporting Information Figure 12S-1. ¹H-NMR (600 MHz) spectrum of (**18**) in D₂O isolated from enzyme digest of cross-linked DNA (**S18**)

14. Supporting Information Figure 12S-2. ¹H-NMR (600 MHz) spectrum of (**18**) in D₂O isolated from enzyme digest of cross-linked DNA(**S19**)

15. Supporting Information Figure 13S. ${}^{1}\text{H}{}^{-1}\text{H} \cos y$ (600 MHz) spectrum of (18) in D₂O isolated from enzyme digest of cross-linked DNA(S20)

16. Supporting Information Figure 14S-1. ¹H-NMR (600 MHz) spectrum of (**18**) in DMSO isolated from enzyme digest of cross-linked DNA (**S21**)

17. Supporting Information Figure 14S-2. ¹H-NMR (600 MHz) spectrum of (**18**) in DMSO isolated from enzyme digest of cross-linked DNA (**S22**)

18. Supporting Information Figure 15S. ¹H-¹H cosy (600 MHz) spectrum of (**18**) in DMSO isolated from enzyme digest of cross-linked DNA (**S23**)

19. Supporting Information Table 1, Comparison of the chemical shifts in ¹H-NMR (in DMSO-d6) of adducts (S24)

General Methods. Unless otherwise noted, materials were obtained from commercial suppliers and used without further purification. The ¹H-NMR (400 MHz or 600 MHz) spectra were recorded on a JEOL LAMBDA 400 or 600. Chemical shift values (δ) are reported relative to H₃PO₄ (85%) for the ³¹P NMR (external

standard). ESIMSs were recorded using a BioTOF II mass spectrometer or APEX III (Bruker Daltonics) . MALDI-TOF mass spectra were measured by negative mode using the laser at 337 nm and 3-hydroxypicolinic acid as the matrix. The ultraviolet-visible (UV-vis) absorption spectra were recorded by a JASCO V-550 UV-VIS or GENESIS 10uv scanning system (Thermo Electron Corporation). ODN synthesis was carried out by the use of an automated DNA synthesizer (ABI, 392 DNA/RNA synthesizer) following the standard phosphoramidite chemistry. HPLC was performed on cosmosil 5C18AR and 5C18MS columns (nacalai tesque) monitoring at 254 nm. Densitometric analysis of the gel was carried out on the 20% denaturing polyacrylamide gel plates, and visualized, quantified with use of a FLA-5100 Fluor Imager.

4-Amino-2-(2-octylthioethyl)-6-oxopyrimidine То (4). solution of а 4-amino-2-bromo-6-oxo-pyrimidine (3) (250 mg, 1.3 mmol) and (PPh₃)₂PdCl₂ (91 mg, 0.13 mmol) in DMF (6.5 mL) was added tributylvinylstannane (0.78 mL, 2.7 mmol) under argon and heated to 100 °C for 2.5 hr. After cooled to room temperature, the reaction mixture was added octane thiol (0.45 mL, 2.6 mmol) and stirred at room temperature for overnight. The reaction mixture was concentrated under reduced pressure to give the crude product, which was chromatographed on a silica gel column (chloroform \rightarrow chloroform:methanol=19:1) to give 4 (285 mg, 1.0 mmol, 77 %) as a pale yellow crystals. mp 181 °C; ¹H-NMR (CD₃OD, 400 MHz, δ 0.90 (t, J = 6.8 Hz, 3H), 1.21-1.43 (m, 10H), 1.57 (quint, J = 7.2 Hz, 2H), 2.54 (t, J = 8.0 Hz, 2H), 2.75-2.80 (m, 2H), 2.85-2.90 (m, 2H), 5.20 (s, 1H); ¹³C NMR (CD₃OD, 100 MHz, δ) 13.9, 22.1, 28.2, 28.2, 28.6, 28.6, 29.0, 30.8, 31.2, 34.5, 83.0, 160.2, 162.8, 164.0; FTIR (cm⁻¹, KBr) 1456, 1605, 1702, 2848, 2923; HRMS (ESI) calcd for $C_{14}H_{25}N_3OS$, 306.1611 ($[M + Na]^+$); Found, 306.1612 ($[M + Na]^+$).

$\label{eq:2.1} \begin{array}{l} \mbox{4-Amino-1-[2-(2'deoxy-3'-O-methoxymethyl-5'-trityl-β-D-ribofuranosyl)ethyl]} \\ \mbox{2-(2-octylthioethyl)-6-oxopyrimidine (6)} \end{array}$

Lithium diisopropyloamide (LDA; 2.0 M in heptane, THF, ethylbenzene, 0.71 mL, 1.41 mmol) was added to the solution of **4** (400 mg, 1.4 mmol) in toluene (2.8 mL) under argon and the reaction mixture was stirred at room temperature for 1 hr. A solution of deoxy ribose derivative (**5**) (711 mg, 1.2 mmol) in toluene (3.5 mL) was added to the above mixture and heated to 100 °C for overnight. The reaction mixture was diluted with ethylacetate then washed successively with saturated aqueous NH₄Cl and brine. The organic layer was dried (Na₂SO₄), then evaporated to give the crude product, which was chromatographed on a silica gel column

S2 —

(chloroform:ethylacetate =6:4→ ethylacetate) to give **6** (238 mg, 0.33 mmol, 36 % based on recovery) as a colorless oil. ¹H NMR (CDCl₃, 400 MHz, ∂) 0.88 (t, *J* = 6.8 Hz, 3H), 1.15^{-1.40} (m, 10H), 1.50 (quint, *J* = 8.0 Hz, 2H), 1.60^{-1.80} (m, 2H), 1.95^{-2.15} (m, 2H), 2.44 (t, *J* = 6.8 Hz, 1H), 2.85 (t, *J* = 8.0 Hz, 2H), 2.89^{-3.10} (m, 3H), 3.19 (dd, *J* = 4.0, 10.0 Hz, 1H), 3.29 (s, 3H), 4.00^{-4.13} (m, 4H), 4.20 (d, *J* = 6.0 Hz, 1H), 4.45 (br s, 2H), 4.61 (s, 2H), 5.33 (s, 1H), 7.23 (t, *J* = 7.6 Hz, 3H), 7.30 (t, *J* = 7.6 Hz, 6H), 7.45 (d, *J* = 7.6 Hz, 6H); ¹³C NMR (CDCl₃, 100 MHz, ∂) 14.1, 22.6, 28.7, 28.8, 29.2, 29.3, 29.6, 31.8, 32.4, 34.3, 34.4, 34.6, 38.3, 40.7, 55.4, 64.3, 75.9, 78.8, 86.2, 86.6, 95.2, 127.0, 127.8, 128.7, 143.9, 160.4, 160.9, 163.2; HRMS (FAB) calcd for C₄₂H₅₅N₃O₅S, 714.3941 ([M + H]+); Found, 714.3948 ([M + H]+).

4-Phenoxyacetylamino-1-[2-(2'deoxy-3'-O-methoxymethyl-5'-trityl-β-D-ribofur anosyl)ethyl] -2-(2-octylthioethyl)-6-oxopyrimidine. To a solution of 6 (230 mg, 0.32 mmol) in dry CH₃CN (2.4 mL) was added 1-hydroxybenzotriazole (200 mg, 1.5 mmol) in dry pyridine solution (0.8 mL) in the presence of 4 Å molecular sieves at room temperature. After the resulting mixture was stirred for 1 h, freshly distilled phenoxyacetylchloride (0.262 mL, 1.6 mmol) was added slowly to the reaction The reaction mixture was stirred for overnight, diluted with mixture. dichloromethane, and washed with saturated aqueous NaHCO₃ and brine. The organic layer was dried over anhydrous Na₂SO₄, filtered, and concentrated under reduced pressure. The residue was purified by silica gel chromatography (dichloromethane:ethyl acetate 4:1, v/v) to give 4-amino protected derivative as a colorless foam (211 mg, 0.25 mmol, 78%): 1H NMR (CDCl₃, 400 MHz, *b*) 0.87 (t, J) = 6.8 Hz, 3H), 1.20-1.34 (m, 10H), 1.50 (quint, J = 8.0 Hz, 2H), 1.60-1.80 (m, 2H), 2.03 (dd, J= 6.0, 12.8 Hz, 1H), 2.09-2.19 (m,1H), 2.45 (t, J= 8.0 Hz, 2H), 2.94-3.22 (m, 4H), 3.29 (s, 3H), 4.00-4.16 (m, 4H), 4.20 (d, J = 6.8 Hz, 1H), 4.57 (s, 2H), 4.60(d, J = 2.0 Hz, 2H), 6.98 (d, J = 8.0 Hz, 2H), 7.07 (t, J = 8.0 Hz, 1H), 7.20-7.26 (m, 4H), 7.30 (t, J = 8.0 Hz, 6H), 7.35 (t, J = 8.0 Hz, 2H), 7.44 (d, J = 8.0 Hz, 6H), 8.43 (br s, 1H); ¹³C NMR (CDCl₃, 100 MHz, *ð*) 14.1, 22.6, 28.6, 28.8, 29.2, 29.3, 29.5, 31.8, 32.5, 33.9, 34.5, 38.3, 41.4, 55.4, 64.3, 67.5, 75.8, 78.8, 84.3, 86.6, 95.2, 97.3, 114.9, 122.6, 127.0, 127.8, 128.7, 129.9, 143.8, 152.4, 156.8, 160.6, 163.3, 167.1; HRMS (ESI) calcd for $C_{50}H_{61}N_3O_7S$, 870.4128 ([M + Na]+); Found, 870.4127 ([M + Na]+).

4-Phenoxyacetylamino-1-[2-(2'deoxy-β-D-ribofuranosyl)ethyl]-2-(2-octylthioeth yl)- 6-oxopyrimidine (7). A solution of 4-amino protected derivative (200 mg, 0.236 mmol) and dimethylsulfide (3.07 mL, 41.8 mmol) in dichloromethane (7.2 mL) was

S3 —

cooled with ice bath and BF_3 ·Et₂O (0.669 mL, 5.42 mmol) was added to the reaction mixture. After stirring for 1.5 h, the reaction mixture was diluted with dichloromethane and washed with saturated aqueous NaHCO₃ and brine. The organic layer was dried over anhydrous Na₂SO₄, filtered, and concentrated under reduced pressure. The residue was purified by silica gel chromatography (chloroform:methanol 24:1) to give 7 as a colorless foam (125 mg, 0.223 mmol, 94 %): ¹H NMR (CDCl₃, 400 MHz, *δ*) 0.88 (t, *J* = 8.0 Hz, 3H), 1.20-1.50 (m, 10H), 1.61 (quint, J = 6.8 Hz, 2H), 1.70-1.85 (m, 2H), 1.90 (br s, 1H), 2.00-2.15 (m, 2H), 2.42 (br s, 1H), 2.58 (t, J = 8.0 Hz, 2H), 2.93 (t, J = 6.8 Hz, 2H), 3.07 (t, J = 6.8 Hz, 2H), 3.59-3.68 (m, 1H), 3.73-3.81 (m, 1H), 3.86 (dt, J = 4.0, 4.8 Hz, 1H), 4.05-4.25 (m, 3H), 4.23-4.40 (m, 1H), 4.60 (s, 2H), 6.99 (d, J = 7.6 Hz, 2H), 7.08 (t, J = 7.6 Hz, 2H)1H), 7.22 (s, 1H), 7.36 (t, J = 7.6 Hz, 2H), 8.46 (br s, 1H); ¹³C NMR (CDCl₃, 100 MHz, *b* 14.1, 22.6, 28.7, 28.9, 29.2, 29.3, 29.6, 31.8, 32.7, 34.4, 34.6, 41.1, 41.4, 62.9, 67.5, 73.0, 75.6, 87.0, 97.3, 114.9, 122.7, 129.9, 152.6, 156.8, 160.3, 163.3, 167.1; HRMS (ESI) calcd for $C_{29}H_{43}N_3O_6S$, 584.2770 ([M + Na]+); Found, 584.2768 ([M + Na]+).

A solution of 7 (120 mg, 0.21 mmol) in dry pyridine (2.1 mL) was cooled with ice bath and dimethoxytrityl chloride (144 mg, 0.42 mmol) was added into the solution and the mixture was stirred for overnight. The resulting mixture was diluted with dichloromethane and washed with saturated aqueous NaHCO₃, H₂O, and brine. The organic layer was dried over anhydrous Na₂SO₄, filtered, and concentrated under reduced pressure. The residue was purified by silica gel chromatography (chloroform-MeOH=24:1) to afford 4-phenoxyacetylamino-1-[2-(2'deoxy-5'-O-(4,4'-dimethoxytrityl) -β-D-ribofuranosyl)ethyl]-2-(2-octylthioethyl)-6-oxopyrimidine as a colorless foam (166 mg, 0.192 mmol, 90 %). colorless foam; ¹H NMR (CDCl₃, 400 MHz, δ) 0.87 (t, J = 8.0 Hz, 3H), 1.20-1.50 (m, 10H), 1.52 (quint, J = 6.8 Hz, 2H), 1.70-1.85 (m, 2H), 1.94 (dd, J = 6.0, 12.4 Hz, 1H), 2.04-2.20 (m, 1H), 2.47 (t, J = 6.8 Hz, 2H), 2.88 (t, J = 6.8 Hz, 2H), 2.93-3.14 (m, 3H), 3.26 (dd, J = 4.8, 10.0 Hz, 1H), 3.79 (s, 6H), 3.86-3.92 (m, 1H), 4.06-4.20 (m, 3H), 4.24-4.32 (m, 1H), 4.58 (s, 2H), 6.83 (d, J = 8.8 Hz, 4H), 6.99 (d, J = 7.6 Hz, 2H), 7.08 (t, J = 7.6 Hz, 1H), 7.18-7.24 (m, 2H), 7.26-7.40 (m, 8H), 7.43 (d, J = 8.0 Hz, 2H), 8.43 (br s, 1H); ¹³C NMR (CDCl₃, 100 MHz, *ð*) 14.1, 22.6, 28.6, 28.9, 29.2, 29.3, 29.6, 31.8, 32.5, 34.1, 34.5, 40.6, 41.3, 55.2, 64.5, 67.5, 74.6, 75.5, 85.7, 86.2,

S4 -

97.3, 113.1, 114.9, 122.6, 126.8, 127.8, 128.1, 129.9, 130.0, 135.9, 144.8, 152.5, 156.8, 158.5, 160.5, 163.3, 167.1; HRMS (ESI) calcd for C₅₀H₆₁N₃O₈S, 886.4077 ([M + Na]+); Found, 886.4069 ([M + Na]+).

$\label{eq:2.1} 4-Phenoxyacetylamino-1-[2-(2'deoxy-3'-N,N-diixopropylcyanoethylphosphoram idyl-5'-O-(4,4'-dimethoxytrityl)-\beta-D-ribofuranosyl)ethyl]-2-(2-octylthioethyl)-$

6-oxopyrimidine (7). To a solution of the above product (81mg, 0.094 mmol) in dry dichloromethane (1.9 mL) was added diisopropylethylamine(0.09 mL, 0.56 mmol) at 0 °C and 2-cyanoethyl N,N-diisopropylchlorophosphoramidite(0.063 mL, 0.28 mmol) was added. The reaction mixture was stirred for 1 h at the same temperature. The resulting mixture was diluted with dichloromethane and washed with saturated aqueous NH₄Cl, H₂O, and brine. The organic layer was dried over anhydrous Na₂SO₄, filtered, and concentrated under reduced pressure. The residue was purified by silica gel chromatography (hexanes:ethyl acetate: $Et_3N = 57:38:5$) to afford 8 as a colorless foam (62 mg, 0.059 mmol, 63%). ¹H NMR (CDCl₃, 400 MHz, δ 0.87 (t, J = 6.8 Hz, 3H), 1.06 (d, J = 6.8 Hz, 3H), 1.10-1.20 (m, 9H), 1.20-1.35 (m, 10H), 1.51 (quint, J = 8.0 Hz, 2H), 1.70-1.85 (m, 2H), 2.00-2.20 (m, 2H), 2.40-2.50 (m, 3H), 2.60 (t, J = 6.8 Hz, 1H), 2.88 (t, J = 8.0 Hz, 2H), 2.88-3.20 (m, 4H), 3.40-3.75 (m, 4H), 3.79 (s, 6H), 4.00-4.10 (m, 1H), 4.10-4.20 (m, 3H), 4.33-4.43 (m, 1H), 4.57 (s, 2H), 6.81 (d, J = 5.2 Hz, 2H), 6.83 (d, J = 4.8 Hz, 2H), 6.98 (d, J = 8.8 Hz, 2H), 7.07 (t, J = 8.0 Hz, 1H), 7.15-7.24 (m, 2H), 7.24-7.38 (m, 8H), 7.44 (dd, J = 4.0, 6.8 Hz, 2H), 8.43 (br s, 1H); ³¹P NMR (CDCl₃, 162 MHz, δ) 145.2, 145.6.

Oligonucleotide Synthesis. All oligonucleotides weresynthesized at a 1 µmol scale on ABI 392 DNA/RNA synthesizer with standard β -cyanoethyl phosphoramiditechemistry. The 5'-terminal dimethoxytrityl-bearing ODNs were removed from the solid support by treatment with 28% NH₃ (2.5 mL) and evaporated under argon. The crude product was purified by reverse-phase HPLC with C-18 column (nacalai tesque: COSMOSIL 5C18-MS-II, 10 x 250 mm) by a linear gradient of 10-40%/20 min of acetonitrile in 0.1 M TEAA buffer at a flow rate of 4 mL/min. The dimethoxytrityl group of the purified ODN was cleaved with 10% AcOH and the mixture was additionally purified by ethanol precipitation to afford ODN (**9**) (433 nmol, 43%) MALDI-TOF MS (*m/z*) **9**: calcd 4914.44, found 4915.69.

Preparation of the ODN (11). To a solution of ODN **9** (1.0 nmol, 1.84 μ L) was added a solution of magnesium monoperphthalate (MMPP) (5.0 nmol) in carbonate buffer

adjusted to pH 10 at room temperature. After 1 h, NaOH (4 M, 1 μ L) was added, and the mixture was left for an additional 2 h to give ODN (**11**). MALDI-TOF MS (*m*/*z*) **11**: calcd 4768.14, found 4769.47.

Electrophoresis Assay of the Cross-Linking Reaction. The reaction was performed with 10 μ M of ODN (11) and 5 μ M of the target DNA (12a) or RNA (12b) labeled by fluorescein at 5'-end in a buffer of 100 mM NaCl and 50 mM MES buffer at pH 7.0. The reaction mixture was incubated at 30 °C for 1.5~12 h. The reaction was stopped by the addition of the loading dye (95% formamide, 20 mM EDTA, 0.05% xylene cyanol, and 0.05% bromophenol blue) and heating at 90 °C for 10 min. The cross-linked products were analyzed by a denaturing 20% polyacrylamide gel electrophoresiscontaining urea (7 M) with TBE buffer at 250 V for 2.5 h. The labeled bands were visualized and quantified with use of a FLA-5100 Fluor Imager.

Isolation of the Cross-Linking Adducts. The cross linking reaction was performed with the reactive ODN (**11**) (540 μ M) and the complementary ODN (**12a**) (N=dT) (785 μ M) in a buffer (100mM NaCl, 50 mM MES buffer, pH 5.0) at 30 °C. After 3 h, the cross-linking adduct was purified by HPLC with an ODS column (nacalai tesque: COSMOSIL 5C18-MS-II, 10 x 250 mm) by a linear gradient of 10-40%/20 min of acetonitrile in 0.1 M TEAA buffer at a flow rate of 4 mL/min to afford cross-linking adducts (**13**). MALDI-TOF MS (*m*/*z*) **11**: calcd 9789.548, found 9788.979.

Enzymatic digestion of **Cross-linking Adducts** (13). Snake venom phosphodiesterase (SVP, 0.4u/µL, 3 µL) and bacterial alkaline phosphatase (BAP, $0.4u/\mu L$, 3 μL) was added to the purified cross-linking adducts (41 nmol, 300 μL) dissolved in buffer (final concentration: 50 mM Tris-HCl, 1 mM MgCl₂). The reaction mixture was incubated at 37 °C for overnight. The reaction mixture was passed through a Microcon cellulose filter (30,000 molecular weight cutoff) by centrifugation at 14,000 RPM. The filtrate was washed with H₂O and the combined filtrate was lyophilized. The sample was dissolved in H₂O and purified with an ODS column (nacalai tesque: COSMOSIL 5C18-MS-II, 4.6 x 250 mm) by a linear gradient of 5-30%/20 min, 30-100%/30 min of acetonitrile in 50 mM ammonium formate buffer at a flow rate of 1 mL/min to give 14.



Supporting Information Figure 1S. ¹H NMR of **6**.

S7





Supporting Information Figure 3S. ¹H NMR of 4-Phenoxyacetylamino-1-[2-(2'deoxy-3'- O-methoxymethyl-5'-trityl- β -D-ribofuranosyl)ethyl] -2-(2-octylthioethyl)-6-oxopyrimidine.



Supporting Information Figure 4S. ¹H NMR of **7**.



Supporting Information Figure 5S. ¹³C NMR of **7**.





Supporting Information Figure 7S. ³¹P NMR of **8**.

Supporting Information Figure 8S. Fluor images of 20% denaturing PAGE analysis of ISC product with the target DNA (**12a**) or RNA (**12b**).



Supporting Information Figure 9S. (A) Fluor images of 20% denaturing PAGE of Fe•EDTA treatment of cross-linked product obtained from **11. (B)** Histogram showing Fe•EDTA cleavage of cross-linked product

Note: Cleavage ceases abruptly in cross-linked lanes (ISC). This represents the cross-linking site.



(B)





Supporting Information Figure 10S. MALDI-TOFMS of the cross-linking adducts (17).



cross-linking adduct







Supporting Information Figure 12S-2. ¹H-NMR (600 MHz) spectrum of (18) in D₂O isolated S19 from enzyme digest of cross-linked DNA



S20



of cross-linked DNA

Supporting Information Figure 14S-2. ¹H-NMR (600 MHz) spectrum of (**18**) in DMSO isolated from enzyme digest of cross-linked DNA S22

Supporting Information Figure 14S. ¹H-¹H cosy (600 MHz) spectrum of (**18**) in DMSO isolated from enzyme digest of cross-linked DNA

Supporting Information of Table 1S. Comparison of the chemical shifts in ¹H-NMR (in DMSO-d6) of adducts

	H8(Hf)	CH ₃	Sugar-1'	Hd
Т	7.69	1.77	6.17	-
T-2O-alkyl ^{a)}	7.81	1.78	6.08	4.29
T-4O-alkyl ^{a)}	7.99	1.87	6.12	4.25
T-3N-alkyl ^{a)}	7.75	1.81	6.20	3.80
Cross-link adduct (T1)	7.68	1.76	6.17	-
Cross-link adduct (T2)	7.87	1.82	6.10	4.28~4.13

a) These values are cited from reference (9).