Synthesis of oligodeoxynucleotides containing 2-thiopyrimidine residues — a new protection scheme

Robert G. Kuimelis and Krishnan P. Nambiar*

Department of Chemistry, University of California, Davis, CA 95616, USA

Received December 23, 1993; Revised and Accepted March 7, 1994

ABSTRACT

A method is described for the incorporation of 2'-deoxy-2-thiouridine (d^{S2}U) and 2'-deoxy-2-thiothymidine (d^{S2}T) into oligodeoxynucleotides at predetermined positions. This requires N³ or O⁴-acylation of d^{S2}U and d^{S2}T with toluoyl chloride. These base-protected thiopyrimidines are completely stable toward the aqueous iodine oxidation reagent used in the phosphoramidite DNA synthesis method. The toluoyl protecting group is removed during the standard post-synthetic ammonia treatment. This novel protection strategy allows d^{S2}U and d^{S2}T to be efficiently incorporated into oligodeoxynucleotides at predetermined sites without the usual problem of desulfurization and decomposition. Several 14-mers containing the Eco-RI recognition site (dGGCGGAAXCCGCC and dGGCGGAAXXCGCGG, where X represents dT, d^{S2}U or d^{S2}T) have been synthesized and characterized by base composition, thermal denaturation, CD spectroscopy and endonuclease substrate activity.

INTRODUCTION

A detailed understanding of how proteins such as repressors, restriction endonucleases and methyltransferases recognize and bind unique DNA sequences and catalyze specific chemical reactions is of critical importance to structural and molecular biology. This knowledge is also necessary for the development of novel nucleic acid binding agents with designed properties. One approach that has proven fruitful for extracting such information is the use of nucleoside analogues. This strategy involves adding, deleting, or altering specific functional groups within the recognition site of the oligonucleotide and measuring the effect on the binding constant, reaction rate, or specificity. A related approach is to design nucleoside analogues that have unique chemical or physical properties, such as cross-linking ability¹ or spectroscopic probe features². Provided the modification to the nucleobase is conservative and does not significantly alter the substrate conformation, these experiments can provide considerable information at the molecular level about individual protein–DNA contacts. These approaches have already yielded important details concerning the interaction of nucleic acids with polymerases,³ repressors,⁴ restriction

endonucleases,⁵ and DNA modification enzymes⁶. Nucleoside analogues have also recently been used to study the specific nucleobase functional groups necessary for ribozyme structure and catalytic activity.⁷

2-Thiopyrimidine nucleosides possess several properties that make them especially attractive and potentially useful when incorporated into oligodeoxynucleotides.⁸ As photosensitizing probes, they can be used to study protein–DNA interactions via triplet state energy transfer to aromatic amino acid residues in protein–DNA complexes.⁹ This is the primary reason we wanted to incorporate these analogues into oligonucleotides. Also, because the thiocarbonyl is available to interact with ligands in the minor groove of duplex DNA and has altered hydrogen bonding characteristics, 2-thiopyrimidines can be used to study minor-groove interactions. Moreover, the large, polarizable sulfur atom of 2-thiopyrimidines is known to enhance the stability of duplex DNA by increasing the stacking interactions.¹⁰ The increased lipophilicity would also increase the permeability of these molecules through cell membranes. Both of these features could prove important for antisense type applications. 2-Thiouracil is also selectively taken-up into melanin-producing cells and is a specific tumor marker for malignant melanoma.¹¹ In addition, d^{S2}U possesses significant anti-viral and anti-leukemic activity.¹² Oligomeric forms may also prove to be medicinally useful.

The thiocarbonyl functionality of 2-thiopyrimidines is reactive toward the aqueous iodine oxidation reagent used in the phosphoramidite DNA synthesis method.^{13, 14a} This prevents the direct use of the unprotected phosphoramidite synthon. Some investigators have used the unprotected d^{S2}T synthon,¹⁴ but yields are poor and pure oligonucleotides are not obtained. Oxidative desulfurization to the corresponding 4-pyrimidinone is the main reaction.¹³ In the case of d^{S2}U, which has not been incorporated into an oligonucleotide prior to this report, the 4-pyrimidinone is less stable and is further degraded with eventual strand cleavage.

In this paper, we describe how 2-thiopyrimidines can be protected by acylation of N³ or O⁴ with toluoyl chloride to render these nucleoside analogues completely resistant to desulfurization or degradation by the oxidation reagent. The toluoyl protecting group is quantitatively removed during the standard post-synthetic ammonia treatment and no additional

*To whom correspondence should be addressed

deprotection steps are necessary. To illustrate the usefulness of this approach, we have used this novel protection scheme to synthesize several 14-mers that contain the *Eco*-RI recognition site: dGGCGGAAXXCCGCC, a self-complementary sequence, and dGGCGGAAXXC GCGG. In each case one or both of the thymidine residues in the enzyme's hexanucleotide recognition site has been replaced with d³²U or d³²T. We have characterized these oligonucleotides by base composition, thermal denaturation, CD spectroscopy and endonuclease substrate activity. The palindromic sequences containing d³²T exhibit increased melting temperatures and complex CD spectra. All of the sequences are cleaved by the *Eco*-RI endonuclease.

MATERIALS AND METHODS

General

¹H, ¹³C, and ³¹P NMR spectra were recorded on a General Electric QE-300 at 300.7, 75.5, and 121.7 MHz, respectively. Unless otherwise indicated, all spectra were recorded in CDCl₃ and referenced to the residual solvent peak. Assignments for most structures were made based in part up on COSY, NOE, edited DEPT, and carbon-proton correlation experiments. Melting points are uncorrected. All reactions were performed under dry nitrogen. Thin layer chromatography (TLC) was performed on Merck silica gel 60 F-254 aluminum-backed plates (0.2 mm layer thickness) and visualization was performed by UV illumination and by staining with phosphomolybdic acid in EtOH. Flash chromatography refers to column chromatography, performed under a slight N₂ pressure, with Merck silica gel 60 (230–400 mesh).

Enzymes were from United States Biochemical. Phosphoramidites were from Cruachem. Solvents were from Burdick & Jackson; CH₃CN and DMF were anhydrous, HPLC grade. All other reagents were from Aldrich Chemical Company and were used as supplied unless otherwise noted. Pyridine was distilled twice from KOH, once from BaO, and stored over 5 Å molecular sieves. CH₂Cl₂ was distilled twice from P₂O₅ and stored over 5 Å molecular sieves. THF was refluxed over Na /

benzophenone and distilled from the dark blue solution just prior to use.

HPLC was performed with an IBM Instruments, Inc. LC/9533 Ternary Gradient Liquid Chromatograph connected to a variable UV detector and a Shimadzu C-R6A Chromatopac integrator. Buffer solutions: A (CH₃CN), B [100 mM triethylammonium acetate (pH 7.0)], C (20 mM KH₂PO₄ (pH 7.0) + 20% CH₃CN), D (buffer C + 1.0 M KCl). All analyses were performed at room temperature.

UV measurements were performed with a Perkin-Elmer Lambda 6B UV/VIS Spectrophotometer connected to a Lambda Interface Module and a Digital Temperature Controller equipped with an aluminum cuvette heating block. The spectrophotometer was interfaced to a PC running Perkin-Elmer Computerized Spectroscopy Software (PECSS) for data acquisition and instrument control.

CD measurements were performed with a Jasco J-600 spectropolarimeter with the following settings: band width, 1.0 nm; sensitivity, 5 mdeg; accumulation, 35; resolution, 1.0 nm; time constant, 2 sec; scan speed, 50 nm/min.

1-(2'-Deoxy-β-D-erythro-ribofuranosyl)-2-thiouracil (1a)

The 3',5'-*O*-*p*-toluoyl diester (710 mg, 1.5 mmol), synthesized as previously described,^{13b,16} was dissolved in absolute MeOH (25 mL) and 0.1 M NaOMe in MeOH (60 mL) was added at 0 °C with stirring. After 3.0 hrs the reaction was judged to be complete by TLC. H₂O (3 mL) was added to quench the excess NaOMe, followed by portions of Dowex 50W-X8 cation exchange resin (H⁺ form) until the pH was neutral by litmus paper. The clear solution was filtered, the resin washed with 70% EtOH, and the solvents and 4-methyl-methylbenzoate were removed *in vacuo* to leave a white, crystalline solid (340 mg, 94%). This material was used without further purification, but can be recrystallized from MeOH (m.p. 140.6–141.6 °C). HRFABMS calcd for C₉H₁₃O₄N₂S, found 245.0602 (Δ -0.6 mmu). UV(MeOH) 215, 272, 290nm (sh). ¹H NMR (DMSO-*d*-6) δ 12.7(s,1H,NH) 8.1(d,1H,H₆) 6.8(dd,1H,H_{1'}) 6.0(d,1H,H₅) 5.3(s,1H,H_{3'}OH) 5.2(s,1H,H_{5'}OH) 4.3(d,1H,H_{4'}) 3.8(d,1H,H_{3'}) 3.6(s,2H,H_{5'}) 2.3(m,1H,H_{2'}) 2.0(m,1H,H_{2''}); ¹³C

Scheme I. Reaction conditions: i) DMTTrCl, DMAP, pyridine ii) TBDMSCl, imidazole, DMF iii) toluoyl chloride, diisopropylethyl amine, pyridine iv) TBAF, pyridinium acetate, THF v) (i-Pr)₂N⁺POCH₂CH₂CN, diisopropylammonium tetrazolide, CH₂Cl₂. Compounds 4a,b; 5a,b and 6a,b are N³, O⁴ regioisomeric mixtures, represented in the scheme by a wavy line.

NMR (DMSO, *d*-6) δ 175.5, 159.8, 140.6, 106.7, 89.3, 88.0, 69.9, 60.8, 40.1.

1-(2'-Deoxy- β -D-erythro-ribofuranosyl)-2-thiothymine (1b)

The 3'-*O*-*p*-toluoyl monoester (100 mg, 0.266 mmol), synthesized as previously described,^{13a} was transesterified with NaOMe in MeOH as described above for **1a** to afford a white solid (50 mg, 73%). ¹H NMR (CDCl₃ + CD₃OD) δ 8.0(s, 1H, H₆) 6.8(t, 1H, H₁) 4.3(dd, 1H, H₃) 3.8(ddd, 2H, H₅) 3.3(m, 1H, H₄) 2.4(m, 1H, H₂) 2.1(m, 1H, H₂) 1.8(s, 3H, CH₃). ¹³C NMR (CDCl₃ + CD₃OD) δ 173.9, 161.4, 137.2, 115.6, 89.4, 87.3, 69.3, 60.6, 40.5, 12.2.

1-(2'-Deoxy-5'-*O*-4,4'-dimethoxytrityl- β -D-erythro-ribofuranosyl)-2-thiouracil (2a)

Diol **1a** (500 mg, 2.1 mmol) was dried in a vacuum desiccator (P₂O₅ + KOH) for 24 hrs before being dissolved in anhydrous pyridine (15 mL). 4,4'-Dimethoxytrityl chloride (972 mg, 2.9 mmol) was added followed by 4-dimethylaminopyridine (12 mg, 0.10 mmol). After 5 hrs the reaction was judged to be complete by TLC. MeOH (1 mL) was added to the orange solution and after 10 min the mixture was poured into sat NaHCO₃ (100 mL) and extracted with CH₂Cl₂ (4 $\times$ 35 mL). Solvents were removed under reduced pressure after drying (Na₂SO₄). The resulting foam was purified by flash chromatography (0–2% MeOH/CHCl₃ + 0.5% pyridine) to give a slightly yellow foam (1.2 g, 90%). FABMS *m/z* 547 (MH⁺). ¹H NMR (DMSO, *d*-6) δ 12.5(s, 1H, NH) 7.75(d, 1H, H₆) 7.4, 7.2, 6.8(m, 13H, Ar) 6.75(dd, 1H, H₁) 5.6(d, 1H, H₅) 5.4(s, 1H, H₃-OH) 4.3(s, 1H, H₄) 3.9(s, 1H, H₃) 3.7(d, 6H, CH₃) 3.2(dd, 2H, H₅) 2.4(m, 1H, H₂) 2.2(ddd, 1H, H₂). ¹³C NMR (DMSO, *d*-6) δ 175.4, 159.5, 158.1, 144.5, 140.2, 136.1, 135.3, 129.7, 127.9, 127.6, 127.8, 113.3, 106.4, 88.9, 85.9, 69.3, 62.7, 55.0, 39.7.

1-(2'-Deoxy-5'-*O*-4,4'-dimethoxytrityl- β -D-erythro-ribofuranosyl)-2-thiothymine (2b)

The same tritylation procedure described above for uridine analogue **2a** was applied to diol **1b** (50 mg, 0.194 mmol), except the reaction time was increased to 18 hrs. The same work-up and purification afforded a slightly yellow foam (112 mg, 100%). ¹H NMR (CDCl₃) δ 7.9(s, 1H, H₆) 7.4–7.1(m, 9H, Ar) 7.0(dd, 1H, H₁) 6.8(dd, 4H, Ar) 5.1(br s, 1H, H₃-OH) 4.6(m, 1H, H₃) 4.2(dd, 1H, H₄) 3.8(s, 6H, OCH₃) 3.5(ddd, 2H, H₅) 2.7(m, 1H, H₂) 2.3(m, 1H, H₂) 1.4(s, 3H, CH₃). ¹³C NMR (CDCl₃) δ 174.1, 161.1, 158.5, 136.4, 135.2, 129.9, 127.9, 127.8, 126.9, 116.1, 113.1, 89.7, 86.7, 86.6, 70.8, 62.9, 55.1, 41.0, 11.9.

1-(2'-Deoxy-3'-*O*-*t*-butyldimethylsilyl-5'-*O*-4,4'-dimethoxytrityl- β -D-erythro-ribofuranosyl)-2-thiouracil (3a)

Trityl ether **2a** (856 mg, 1.57 mmol) was co-evaporated with anhydrous DMF (1 $\times$ 5 mL) and dissolved in anhydrous DMF (5 mL). Imidazole (490 mg, 7.21 mmol) and freshly distilled *t*-butyldimethylsilyl chloride (544 mg, 3.60 mmol) was added. After 18 hrs the reaction mixture was poured into 5% NaHCO₃ (20 mL) containing a small amount of pyridine. This mixture was extracted with CH₂Cl₂ (4 $\times$ 20 mL) and the combined organics were washed with H₂O (1 $\times$ 20 mL). Solvents were removed under reduced pressure after drying (Na₂SO₄). The resulting foam was co-evaporated with toluene and purified by flash chromatography (0–2% MeOH/CHCl₃ + 0.5% pyridine)

to give a white foam (932 mg, 90%). ¹H NMR (CDCl₃) δ 10.6(s, 1H, NH) 8.2(d, 1H, H₆) 7.4–7.1(m, 9H, Ar) 6.8(dd, 4H, Ar) 6.7(dd, 1H, H₁) 5.5(d, 1H, H₅) 4.6(dd, 1H, H₃) 3.9(m, 1H, H₄) 3.8(s, 6H, OCH₃) 3.5(ddd, 2H, H₅) 2.5(m, 1H, H₂) 2.2(m, 1H, H₂) 0.8(s, 9H, *t*-Bu) 0.1(s, 3H, CH₃) –0.1(s, 3H, CH₃). ¹³C NMR (CDCl₃) δ 175.0, 160.2, 158.7, 144.0, 140.6, 135.0, 130.1, 128.1, 127.9, 127.2, 113.2, 106.5, 89.7, 86.9, 86.6, 69.8, 61.2, 55.2, 41.6, 25.6, 17.8, –4.7, –5.0.

1-(2'-Deoxy-3'-*O*-*t*-butyldimethylsilyl-5'-*O*-4,4'-dimethoxytrityl- β -D-erythro-ribofuranosyl)-2-thiothymine (3b)

Trityl ether **2b** (80 mg, 0.143 mmol) was treated as described above for the preparation of uridine analogue **3a**. After 27 hrs the reaction mixture was worked-up as above and purified by flash chromatography (CHCl₃ + 0.5% pyridine) to give a white foam (82.7 mg, 86%). ¹H NMR (CDCl₃) δ 11.9(s, 1H, NH) 8.0(s, 1H, H₆) 7.5–7.2(m, 9H, Ar) 6.8(m, 5H, Ar + H₁) 4.6(dd, 1H, H₃) 4.1(dd, 1H, H₄) 3.8(s, 6H, OCH₃) 3.5(ddd, 2H, H₅) 2.6(m, 1H, H₂) 2.3(m, 1H, H₂) 1.5(s, 3H, CH₃) 0.8(s, 9H, *t*-Bu) 0.1(s, 3H, CH₃) –0.1(s, 3H, CH₃). ¹³C NMR (CDCl₃) δ 173.9, 160.9, 158.7, 144.1, 136.3, 135.2, 130.0, 129.1, 128.1, 127.9, 127.7, 127.1, 116.3, 113.2, 89.7, 86.8, 86.7, 70.8, 62.0, 55.1, 41.4, 25.6, 17.8, 12.0, –4.7, –5.0.

1-(2'-Deoxy-3'-*O*-*t*-butyldimethylsilyl-5'-*O*-4,4'-dimethoxytrityl- β -D-erythro-ribofuranosyl)-2-thio-(3-*N*/4-*O*)-toluoyluracil (4a)

Silyl ether **3a** (720 mg, 1.09 mmol) was dissolved in anhydrous pyridine (20 mL), cooled to 0 °C, and diisopropylethylamine (284 μ L, 1.63 mmol) was added followed by toluoyl chloride (432 μ L, 3.27 mmol). After 3 hrs, CH₂Cl₂ (50 mL) was added and the reaction mixture was poured into 5% NaHCO₃ (25 mL). The organic layer was separated and washed again with 5% NaHCO₃ (1 $\times$ 25 mL), H₂O (1 $\times$ 25 mL) and then dried (Na₂SO₄). The dark foam obtained after removal of the solvents was purified by flash chromatography (CH₂Cl₂ + 0.5% pyridine) to give a white foam (828 mg, 86%). ¹H NMR (CDCl₃) δ 8.4, 8.3(d, 1H, H₆) 7.9(d, 2H, Tol) 7.4–7.2(m, 11H, Ar) 6.9(dd, 4H, Ar) 6.8(m, 1H, H₁) 5.6(d, 1H, H₅) 4.6(m, 1H, H₃) 4.0(m, 1H, H₄) 3.8(s, 6H, OCH₃) 3.5(m, 2H, H₅) 2.6(m, 1H, H₂) 2.5(s, 3H, CH₃) 2.4(m, 1H, H₂) 0.8(s, 9H, *t*-Bu) 0.1(s, 3H, CH₃) –0.1(s, 3H, CH₃). ¹³C NMR (CDCl₃) δ 174.2, 169.0, 163.0, 159.5, 146.0, 144.3, 141.0, 135.1, 113.0, 106.6, 89.9, 86.7, 86.5, 70.2, 61.3, 56.0, 41.8, 25.5, 21.5, 17.8, –4.7, –5.0.

1-(2'-Deoxy-3'-*O*-*t*-butyldimethylsilyl-5'-*O*-4,4'-dimethoxytrityl- β -D-erythro-ribofuranosyl)-2-thio-(3-*N*/4-*O*)-toluoylthymine (4b)

Silyl ether **3b** (82 mg, 0.122 mmol) was treated with toluoyl chloride as described above for the preparation of uridine analogue **4a**, except the reaction was conducted at 25 °C. After 1 hr, H₂O (200 μ L) was added and the reaction mixture was worked-up and purified as described above to give a white foam (87.1 mg, 90%). ¹H NMR (CDCl₃) δ 8.2, 8.1(s, 1H, H₆) 7.9(d, 2H, Tol) 7.5–7.2(m, 11H, Ar) 6.8(m, 5H, Ar + H₁) 4.6(m, 1H, H₃) 4.1(m, 1H, H₄) 3.8(s, 6H, OCH₃) 3.5(ddd, 2H, H₅) 2.6(m, 1H, H₂) 2.4(s, 3H, CH₃) 2.3(m, 1H, H₂) 1.6, 1.5(s, 1H, CH₃) 0.8(s, 9H, *t*-Bu) 0.1(s, 3H, CH₃) –0.1(s, 3H, CH₃). ¹³C NMR (CDCl₃) δ 172.4, 168.4, 160.5, 158.7, 145.7, 144.0, 136.3, 135.1, 116.2, 113.0, 89.5, 87.0, 86.8, 70.6, 61.8, 55.1, 41.5, 25.5, 21.7, 17.7, 12.0, –4.7, –5.0.

1-(2'-Deoxy-5'-O-4,4'-dimethoxytrityl-β-D-erythro-ribofuranosyl)-2-thio-(3-N/4-O)-toluoyl-uracil (5a)

Toluoyl derivative **4a** (828 mg, 1.06 mmol) was dissolved in THF (75 mL) containing pyridine (912 μL, 11.29 mmol) and AcOH (539 μL, 9.41 mmol) and 1 M tetrabutylammonium fluoride (1.41 mL, 14.11 mmol) in THF was added. After 27 hrs the reaction mixture was concentrated to a low volume, diluted with Et₂O (100 mL) and washed with H₂O (1 × 50 mL). The separated aqueous layer was extracted with Et₂O and the combined organic layers were washed with H₂O (25 mL), dried (Na₂SO₄) and concentrated to a yellowish foam. Purification by flash chromatography (CH₂Cl₂ + 0.5% pyridine) afforded some unreacted **4a** (129 mg) and the desired compound, **5a** (498 mg, 86% accounting for unreacted starting material). ¹H NMR (CDCl₃) δ 8.3(d,1H,H₆) 7.9(d,2H,Ar) 7.5–7.2(m,11H,Ar) 6.9(dd,4H,Ar) 6.8(dd,1H,H_{1'}) 5.7(d,1H,H_{5'}) 4.6(m,1H,H_{3'}) 4.1(m,1H,H_{4'}) 3.8(s,6H,OCH₃) 3.6(m,2H,H_{5'}) 2.7(m,1H,H_{2'}) 2.4(s,3H,CH₃) 2.3(m,1H,H_{2'}); ¹³C NMR (CDCl₃) δ 173.5, 168.0, 159.5, 158.7, 146.1, 144.1, 140.6, 135.0, 130.5, 130.1, 128.6, 128.1, 113.3, 106.0, 89.6, 87.2, 86.4, 70.3, 62.0, 55.3, 41.4, 21.9.

1-(2'-Deoxy-5'-O-4,4'-dimethoxytrityl-β-D-erythro-ribofuranosyl)-2-thio-(3-N/4-O)-toluoyl-thymine (5b)

Toluoyl derivative **4b** (87 mg, 0.11 mmol) was dissolved in THF and treated with tetrabutylammonium fluoride and pyridinium acetate as described above for uridine analogue **5a**. After 72 hrs, the reaction mixture was worked-up as described above and purified by flash chromatography (0.5–0.75% MeOH/CH₂Cl₂ + 0.5% pyridine) to give a white foam (65 mg, 87%). ¹H NMR (CDCl₃) δ 8.1,8.0(s,1H,H₆) 7.8(d,2H,Tol) 7.5–7.2(m,11H,Ar) 6.9(dd,1H,H_{1'}) 6.8(dd,4H,Ar) 4.6(m,1H,H_{3'}) 4.5(br s,1H,H_{3'-OH}) 4.1(m,1H,H_{4'}) 3.8(s,6H,OCH₃) 3.5(ddd,2H,H_{5'}) 2.6(m,1H,H_{2'}) 2.4(s,3H,CH₃) 2.3(m,1H,H_{2'}) 1.5(s,3H,CH₃); ¹³C NMR (CDCl₃) δ 172.5, 168.0, 161.1, 158.8, 145.9, 144.8, 137.0, 135.8, 116.0, 113.5, 89.5, 86.5, 86.4, 71.0, 63.5, 55.5, 41.7, 22.0, 12.1.

1-(2'-Deoxy-3'-O-(2-cyanoethyl-N,N-diisopropyl)-phosphoramidite-5'-O-4,4'-dimethoxytrityl-β-D-erythro-ribofuranosyl)-2-thio-(3-N/4-O)-toluoyl-uracil (6a)

Toluoyl derivative **5a** (385 mg, 0.58 mmol) was stored in a vacuum desiccator (P₂O₅ + KOH) overnight, dissolved in anhydrous CH₂Cl₂ (15 mL) and diisopropylamine (29.3 mg, 0.29 mmol) was added followed by tetrazole (20.3 mg, 0.29 mmol). β-Cyanoethyl-N,N,N'-tetraisopropylphosphoramidite (192 mg, 0.64 mmol) was then added dropwise with stirring. After 5 hrs the reaction was judged to be complete by TLC and was diluted with CH₂Cl₂ (20 mL) and poured into 2.5% NaHCO₃ (10 mL). The layers were separated and the organic layer was washed with H₂O (10 mL), dried (Na₂SO₄) and concentrated. The residue was purified by flash chromatography (0.5% MeOH/CH₂Cl₂ + 0.5% pyridine) to give a white foam (386 mg, 77%). ³¹P NMR (CDCl₃) δ 151.45, 152.26, 152.56, 152.64.

1-(2'-Deoxy-3'-O-(2-cyanoethyl-N,N-diisopropyl)-phosphoramidite-5'-O-4,4'-dimethoxytrityl-β-D-erythro-ribofuranosyl)-2-thio-(3-N/4-O)-toluoyl-thymine (6b)

Toluoyl derivative **5b** (65 mg, 0.096 mmol) was phosphitylated, worked-up and purified as described above for uridine analogue **6a** to give a white foam (61 mg, 72%). ³¹P NMR (CDCl₃) δ 151.50, 151.87, 152.42.

Oligonucleotide synthesis

All oligonucleotides were synthesized at 0.2 μmol scale on a Pharmacia Gene Assembler Plus DNA synthesizer using conventional β-cyanoethyl phosphoramidite chemistry. The standard bases were dissolved in CH₃CN to 0.1 M according to the manufacturer's specifications; all modified bases were dissolved in CH₃CN:DMF (1:1) to the same concentration. The synthesis protocol was modified to increase the coupling time from 1.7 to 8 min for all bases. All oligomers were synthesized 'trityl off'. After synthesis, the solid supports were treated with fresh, concentrated NH₄OH (1.0 mL) for 12 hrs at 25 °C and 2 hrs at 50 °C, and the solution was then concentrated to dryness.

Table 1. Summary of oligodeoxynucleotide sequences and properties.

	sequence (5'→3') ^a	X	T _m (°C) ^b	-ΔH _{1/2} ^c (kcal·mol ⁻¹)	E ₉₅ °C E ₂₀ °C	Eco-RI substrate? ^d	composition dC; dG; dT; dA; dX	yield (%) ^e
odn-1	d(GGCGGAATTCGCGC)	---	64	48	1.13	+++	5.5; 4.8; 1.7; 2.0; 0.0	nd
odn-2	d(GGCGGAAXTCGCGC)	2ST	69	51	1.12	++	5.0; 4.9; 1.0; 1.8; 1.0	46
odn-3	d(GGCGGAATXCGCGC)	2ST	72	46	1.18	+++	5.5; 5.2; 0.9; 1.6; 0.9	30
odn-4	d(GGCGGAAXXCGCGC)	2ST	71	57	1.17	++	5.5; 4.7; 0.0; 1.7; 2.1	32
odn-5	d(GGCGGAATTCGCGG)	---	70	118	1.14	++	3.2; 7.1; 1.9; 1.8; 0.0	51
odn-6	d(GGCGGAAXTCGCGG)	2ST	66	109	1.11	+	3.6; 6.7; 1.0; 1.7; 1.0	31
odn-7	d(GGCGGAATXCGCGG)	2ST	67	111	1.13	++	3.3; 7.0; 0.9; 1.8; 1.0	36
odn-8	d(GGCGGAAXXCGCGG)	2ST	68	110	1.12	+	3.6; 6.6; 0.0; 1.7; 2.0	26
odn-9	d(GGCGGAAXTCGCGG)	2SU	67	128	1.11	+	3.5; 6.9; 0.9; 1.7; 0.9	15
odn-10	d(GGCGGAATXCGCGG)	2SU	66	116	1.12	+++	3.2; 7.0; 0.9; 1.8; 1.1	18
odn-11	d(GGCGGAAXXCGCGG)	2SU	77	111	1.13	++	3.0; 7.1; 0.0; 1.8; 2.0	10

^a Only the top strand is shown. Odn-1 through odn-4 are palindromic. Odn-5 through odn-11 were annealed with an equimolar amount of a second, complementary strand containing the normal four bases. ^b Determined in 100 mM NaCl, 25 mM KH₂PO₄ (pH 7.0). ^c Determined from the differentiated UV melting curve. ^d >90% of substrate cleaved by Eco-RI endonuclease after 4 hrs (+++); >90% after 16 hrs (++); >80% after 32 hrs with additional enzyme (+). ^e Percentage yield has been calculated as (optical density of the pure oligo / optical density of the crude material injected into HPLC) × 100.

The residue was then taken up in 0.3 M NaOAc (100 μ L, pH 5.2), 1 mL cold EtOH was added, the solution kept at -78°C for 12 hrs, and then centrifuged at 14,000 RPM at 0°C for 40 min. After decanting the supernatant, the pellet was dried under vacuum and stored at -20°C .

Oligonucleotide purification

The crude, detritylated oligonucleotides were purified by reversed-phase HPLC (Rainin Microsorb C18 column, 25 cm $\times$ 4.6 mm, 5 μ particle size, 60 $\text{\AA}$ pore) using a linear gradient of 6 to 20% A in buffer B over 25 min at a flow of 1.5 mL/min. Detection was at 260 nm for analytical runs (0.05 O.D.) and 290 nm for preparative runs (10–12 O.D.). Appropriate fractions were collected, lyophilized and stored at -20°C . The purity of some of the samples was confirmed with anion exchange HPLC (Pharmacia Mono-Q HR 5/5, 50 $\times$ 5 mm, 10 μ particle size) using a linear gradient of 10 to 70% buffer D in buffer C over 70 min. Gel electrophoresis was also performed on some of the samples using 20% acrylamide, 7M urea.

Deoxynucleoside composition analysis

The enzymatic digestions were done with snake venom phosphodiesterase (SVP, *Crotalus adamanteus*) and bacterial alkaline phosphatase (BAP).¹⁵ The purified oligo (0.05 O.D.) was incubated overnight at 37°C with SVP (1.0 U) and BAP

(2.0 U) in a total volume of 80 μ L containing Tris buffer (42 mM, pH 7.5) and MgCl_2 (15 mM). The nucleosides were recovered by adding 3.0 M NaOAc (10 μ L, pH 5.2), absolute EtOH (250 μ L), chilling at -78°C for 4 hrs, and centrifuging at 12,000 RPM for 20 min. The supernatant was collected, evaporated, dissolved in 50 μ L of H_2O , and analyzed by reversed-phase HPLC (Rainin Microsorb C18 column, 25 cm $\times$ 4.6 mm, 5 μ particle size, 60 $\text{\AA}$ pore) using a linear gradient of 5 to 6% A in buffer B over 15 min at a flow of 1.0 mL/min. Detection was at 260 nm. The following correction factors were applied to the integrated areas of the deoxynucleoside peaks: dC = 1.000; dG = 0.624; dT = 0.830; dA = 0.474; d³²U and d³²T = 0.903.

Thermal denaturation and CD spectroscopy

Approximately 0.7 O.D. units of each purified oligonucleotide was dissolved in 25 mM KH_2PO_4 (pH 7.0) containing 100 mM NaCl and heated at 95°C for five min. This solution was allowed to cool to 20°C over a 2.5 hour period just prior to analysis. The samples were then heated from 20°C to 95°C at a rate of $0.71^{\circ}\text{C}/\text{min}$ and the absorbance at 260 nm was measured at 2.2 sec intervals. The same samples were used for the CD measurements after re-annealing as described above. CD measurements were taken at room temperature with a 1 mm path-length cell.

Figure 1. (A) Reversed-phase HPLC chromatogram of crude odn-6. (B) Base composition analysis of odn-6. (C) Reversed-phase HPLC chromatogram of crude odn-9. (D) Base composition analysis of odn-9.

Eco-RI endonuclease digestions

6.4 μ g of each purified oligonucleotide in 100 μ L buffer [25 mM KH_2PO_4 (pH 7.0), 100 mM NaCl] was annealed as described above just prior to analysis. These solutions were adjusted to 25 mM Tris (pH 7.5), 10 mM MgCl_2 , 1 mM dithiothreitol and 40 units *Eco*-RI endonuclease was added. The digestions were conducted at 37 °C. 30 μ L aliquots were removed at four and twelve hour intervals and the enzyme was deactivated by heating at 60 °C for five min. An additional 40 units of the enzyme was added to the remaining sample and the digestion was allowed to proceed an additional 16 hours. Analysis was performed by reversed-phase HPLC using the conditions described above for the purification of the oligomers.

RESULTS AND DISCUSSION**Synthesis of the phosphoramidites**

Scheme 1 outlines our method for the synthesis of the protected 2-thiopyrimidine phosphoramidite monomers. 2'-Deoxy-2-thiouridine, **1a**, was prepared from the α -chlorosugar and silylated 2-thiouracil, as we described in an earlier paper.¹⁶ 2'-Deoxy-2-thiothymidine, **1b**, was prepared via a modification of Faerber and Scheit's procedure employing sulfhydrylase of the cyclic 2,5'-anhydro species.^{13a,17} Both free nucleosides were selectively tritylated to give the 5'-*O*-trityl ethers **2a** and **2b**, and then silylated with *t*-butyldimethylsilyl chloride and imidazole in DMF to give the 3'-*O*-silyl ethers **3a** and **3b**. Treatment of these diethers with toluoyl chloride and *N,N*-diisopropylethylamine in pyridine readily afforded the fully protected derivatives **4a** and **4b**. These were obtained as inseparable mixtures of the N³ and O⁴-acyl products (2:1), which is in contrast to thymidine or uridine, which generally only gives the N³ product.¹⁸ Attempts to alter the ratio by modifying the reaction conditions were unsuccessful. Fortunately, both the N³ and O⁴ product are unreactive toward the iodine reagent and both are readily cleaved by aqueous ammonia. All attempts to directly protect the sulfur atom of **3a**, for example with methylmethanethiosulfonate, were unsuccessful.^{14b} Derivatives **4a** and **4b** were desilylated with excess tetrabutylammonium fluoride in THF containing pyridinium acetate to give the key intermediates **5a** and **5b**. The pyridinium acetate is necessary to quench the alkoxide that is formed and thus prevent deprotection of the toluoyl group and/or transesterification. Phosphitylation with β -cyanoethyl-*N,N,N',N'*-tetraisopropylphosphoramidite¹⁹ and diisopropylammonium tetrazolide in CH_2Cl_2 afforded the phosphoramidite synthons **6a** and **6b** in good yield.²⁰

Synthesis and purification of the oligomers

Table 1 gives the sequences of the 11 oligodeoxynucleotides (odn) we have synthesized. All of the sequences are 14 bases in length and contain the *Eco*-RI hexanucleotide recognition site. The first set of sequences (odn-1–4) is self-complementary, or palindromic. The second set (odn-5–11) is non-palindromic and requires a separate, complementary strand to form the DNA duplex required for enzyme activity. All of the bases were quantitatively incorporated into the oligomers, as determined by a spectrophotometric analysis of the trityl fractions.

Figure 1A shows the crude HPLC chromatogram of a representative oligomer. In each case the desired oligomer was the major product. All of the sequences were purified by reversed-phase HPLC. Some of the palindromic sequences required the

Figure 2. UV melting curves of the DNA duplexes. From left to right at the inflection points (T_m): odn-1, 2, 4, 3 (A); odn-6, 7, 8, 5 (B); odn-10, 9, 5, 11 (C).

collection of two peaks, depending on the injection concentration.²¹ The earlier eluting peak is most likely the hairpin form. Both peaks gave the expected base composition results and identical retention times on an anion-exchange HPLC column.

Although both d^{32}U and d^{32}T were quantitatively coupled into the oligomer, the d^{32}U containing oligomers were consistently recovered in lower yields after purification. Perhaps this is due to the increased reactivity of d^{32}U toward some other reagent, such as the ammonia used for the deprotection.^{14b} However, without base protection of d^{32}U to prevent oxidation, it is not at all possible to incorporate this residue into oligonucleotides using conventional phosphoramidite methodologies. All our attempts revealed that, although the phosphoramidite is incorporated efficiently, no oligomer of the desired length remains intact after the ammonia deprotection, as determined by gel electrophoresis. Iodine promoted desulfurization to the 4-pyrimidinone during DNA synthesis is virtually quantitative,

Figure 3. CD spectra of odn-1–4 (A), odn-5–8 (B) and odn-5, 9–11 (C).

and strand-cleavage occurs at this pyrimidinone site when the oligomer is treated with aqueous ammonia.²²

Characterization of the oligomers

Base composition analysis. Figure 1B shows a representative HPLC base composition analysis. dS²U and dS²T, and all the standard bases, are completely base-line resolved under the analysis conditions. Peaks were identified by co-injection with authentic samples. These results show that the thionucleoside residues are intact, unmodified and are present in the correct ratios. Also, there are no unidentified peaks. We have synthesized the desulfurized nucleosides dH²U and dH²T, and these elute between dC and dG.

Thermal denaturation. Figure 2 shows the UV melting curves. All of the sequences exhibited a cooperative hyperchromic melting profile consistent with a helix-coil transition. All of the T_m values were above 64 °C, which implies that each oligonucleotide will be double-helical at 37 °C. The palindromic dS²T containing sequences (odn-1–4) showed increased T_m values relative to the control (+Δ 4–6 °C). This enhanced stability has previously been observed.^{14b} The increase in stability, however, is not a general effect of sulfur substitution and appears to be due to inter-strand interactions as opposed to intra-strand interactions. This is revealed by the T_m values for the non-palindromic dS²T containing oligonucleotides (odn-5–8), which are slightly lower than the control (–Δ 2–4 °C). This is the first report where non-palindromic oligonucleotides containing dS²T have been prepared. The CD results discussed below are also consistent with inter-strand interactions. Odn-11 exhibits an especially high T_m of 77 °C, possibly due to improved stacking of the thiocarbonyl groups.

The differentiated UV melting curves for the non-palindromic sequences were sharp and symmetrical, whereas those for the palindromic sequences were broader and more complex. The complexity of these first derivative curves implies a multi-state helix-coil transition or multiple intermediates along the melting pathway.²³ Similar results have been obtained for the Dickerson dodecamer, a palindromic sequence closely related to odn-1–4.²⁴ Analysis of the shape of the differentiated UV

melting curve allowed determination of the van't Hoff transition enthalpies²⁵ shown in Table 1.

Circular dichroism. The CD spectra for the non-palindromic modified sequences (odn-6–11) closely resemble the control sequence, odn-5, which has characteristics typical of B-type DNA.²⁶ The main difference between the control and modified sequences is a decrease in intensity of the negative 246nm band. This is possibly due to a widening of the minor-groove to accommodate the more sterically bulky sulfur atom. A-type DNA, which has a wider minor-groove, also has a decreased 246nm band. However, A-type DNA also has a more intense (ca. 2×) 282nm band — a feature *not* exhibited by odn-6–11 (see fig. 3B,C). Although CD spectroscopy is a sensitive probe of macromolecular conformation, it is influenced by nearest-neighbor interactions and is therefore somewhat sequence dependent.²⁷ The differences we observe could be due to nearest-neighbor interactions involving the thiopyrimidine chromophore(s).

The palindromic modified sequences (odn-2,3,4) show more complex CD spectra (see fig. 3A). Odn-4 has 3 distinct maxima: 282, 256 and 237nm. Odn-2 shares the 282 and 256nm band, while odn-3 has the 282 and 237nm band. Since the 256 and 237nm bands are unique to the palindromic sequences, these features can be correlated to inter-strand interactions.

Other workers have made the general conclusion, based on CD experiments with a limited number of short palindromic sequences, that oligonucleotides containing dS²T do not adopt a B-type DNA duplex.^{5c,14b} Our CD results show that if this is the correct interpretation of the CD data, it only holds for *palindromic* sequences where there are possible *inter-strand* sulfur–sulfur interactions. Moreover, the endonuclease digestion experiments described below suggest that the complex CD spectra shown in fig. 3A are not due to significant conformational deviations from the B-type duplex.

Eco-RI endonuclease substrate activity. Each of the oligodeoxynucleotides was tested for its ability to act as a substrate for the *Eco*-RI endonuclease enzyme. No effort was made to extract detailed kinetic information. We have attempted here to

use the enzyme as a qualitative probe of the overall conformation of the oligomer in light of the anomalous CD spectra. We postulated that the enzyme would not recognize a severely distorted duplex and therefore cleavage would not occur. In particular, we were interested in seeing if the palindromic sequences would be good substrates. It is known from X-ray crystallography of the *Eco*-RI endonuclease-DNA complex that the protein interacts only with the major groove of the DNA.²⁸ Also, dodecanucleotides containing dU in place of dT in the hexanucleotide recognition site are known to be efficiently hydrolyzed.^{5d,e} NMR studies of the Dickerson dodecamer, which contains the *Eco*-RI recognition site, have shown that it adopts qualitatively a B-type duplex in solution.²⁹ Since the thiocarbonyl of 2-thiopyrimidines is positioned in the minor groove, our sequences should be cleaved, provided they adopt a conformation suitable to the enzyme.

The results of the digestions are summarized in Table 1. Given enough time, the endonuclease completely cleaved all of the sequences at the correct position, as determined by HPLC analysis of the digestion mixtures (data not shown).^{5d} In addition, all of the sequences where the internal thymidine was replaced with d³²SU or d³²ST (odn-2,4,6,8,9,11) were cleaved less efficiently than the control. Replacement of the external thymidine with d³²SU or d³²ST (odn-3,7,10) had no significant effect, and in one case (odn-10) increased the cleavage rate relative to the control. Most importantly, the palindromic sequences that produced the anomalous CD spectra (odn-2–4) were very effective substrates. This suggests that these oligonucleotides form the standard B-type DNA duplex and that the complex CD spectra are a result of the new thiopyrimidine chromophore(s) and their nearest-neighbor interactions.

REFERENCES

- Nikiforov, T.T. and Connolly, B.A. (1992) *Nucleic Acids Res.*, **20**, 1209–1214.
- Tsao, D.H.H., Maki, A.H. and Chase, J.W. (1990) *FEBS Letters*, **261**, 389–391.
- Dubendorff, J.W., de Haseth, P.L., Rosendahl, M.S. and Caruthers, M.H. (1987) *J. Biol. Chem.*, **262**, 892–898.
- (a) Mazzarelli, J.M., Rajur, S.B., Iadarola, P.L. and McLaughlin, L.W. (1992) *Biochemistry*, **31**, 5925–5936. (b) Caruthers, M.H. (1980) *Acc. Chem. Res.*, **13**, 155–160.
- (a) Lesser, D.R., Kurpiewski, M.R., Waters, T., Connolly, B.A. and Jen-Jacobson, L. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 7548–7552. (b) Aiken, C.R., McLaughlin, L.W. and Gumpert, R.I. (1991) *J. Biol. Chem.*, **266**, 19070–19078. (c) Newman, P.C., Nwosu, V.U., Williams, D.M., Cosstick, R., Seela, F. and Connolly, B.A. (1990) *Biochemistry*, **29**, 9891–9901. (d) McLaughlin, L.W., Benseler, F., Graeser, E., Piel, N. and Scholtissek, S. (1987) *Biochemistry*, **26**, 7238–7245. (e) Brennan, C.A., Van Cleve, M.D. and Gumpert, R.I. (1986) *J. Biol. Chem.*, **161**, 7270–7278.
- (a) Reich, N.O. and Danzitz, M.J. Jr. (1992) *Biochemistry*, **31**, 1937–1945. (b) Brennan, C.A., Van Cleve, M.D. and Gumpert, R.I. (1986) *J. Biol. Chem.*, **161**, 7279–7286.
- (a) Grasby, J.A., Jonathan, P., Butler, G. and Gait, M.J. (1993) *Nucleic Acids Res.*, **21**, 4444–4450. (b) Fu, D.-J. and McLaughlin, L.W. (1993) *Biochemistry*, **32**, 10629–10637. (c) Slim, G. and Gait, M.J. (1992) *Biochem. Biophys. Res. Comm.*, **183**, 605–609. (d) Fu, D.-J. and McLaughlin, L.W. (1992) *Biochemistry*, **31**, 10941–10949. (e) Fu, D.-J. and McLaughlin, L.W. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 3985–3989.
- (a) Taherian, M.R. and Maki, A.H. (1981) *Chem. Phys.*, **55**, 85. (b) Smith, C.A. and Maki, A.H. (1991) *Chem. Phys. Lett.*, **179**, 497–502. (c) Khamis, M.I., Casas-Finet, J.R. and Maki, A.H. (1987) *J. Biol. Chem.*, **262**, 1725.
- (a) Tsao, D.H.H., Casas-Finet, J.R., Maki, A.H. and Chase, J.W. (1989) *Biophys. J.*, **55**, 927.
- Saenger, W. (1984) *Principles of Nucleic Acid Structure*. Springer-Verlag, New York.
- Palumbo, A., d'Ischia, M., Misuraca, G., Iannone, A. and Prota, G. (1990) *Biochim. Biophys. Acta*, **1036**, 221–227.
- Hunter, J.H. and Skulnick, H.I. (1976) U.S. Patent 3,975,374; *Chem. Abstr.*, **85**, 177895b.
- (a) Kuimelis, R.G. and Nambiar, K.P. (1993) *Tetrahedron Lett.*, **34**, 3813–3816. (b) Kuimelis, R.G. Presented in part at the 203rd National Meeting of the American Chemical Society, San Francisco, CA, April 1992; paper ORGN 489.
- (a) Rajur, S.B. and McLaughlin, L.W. (1992) *Tetrahedron Lett.*, **33**, 6081–6084. (b) Connolly, B.A. and Newman, P.C. (1989) *Nucleic Acids Res.*, **17**, 4957–4974.
- Eadie, J.S., McBride, L.J., Efcavitch, J.W., Hoff, L.B. and Cathcart, R. (1987) *Anal. Biochem.*, **165**, 442–447.
- Kuimelis, R.G., Hope, H. and Nambiar, K.P. (1993) *Nucleosides & Nucleotides*, **12**, 737–755.
- Faerber, P. and Scheit, K.H. (1970) *Chem. Ber.*, **103**, 1307–1311.
- Welch, J.C. and Chattopadhyaya, J. (1983) *Acta Chem. Scand. B*, **37**, 147–150.
- Nielsen, J. and Dahl, O. (1987) *Nucleic Acids Res.*, **15**, 3626.
- Barone, A.D., Tang, J.Y. and Caruthers, M.H. (1984) *Nucleic Acids Res.*, **12**, 4051–4061.
- Brown, T. and Brown, D.J.S. (1992) In Lilley, D.M.J. and Dahlberg, J.E., (eds.) *Methods in Enzymology*, Academic Press, New York, Vol. 211A, pp. 20–35.
- This laboratory, unpublished results.
- Cantor, C.R. and Schimmel, P.R. (1980) *Biophysical Chemistry, Part III: The behavior of biological macromolecules*. W.H. Freeman and Company, New York, pp. 1109–1181.
- Breslauer, K.J. (1986) In Hinz, H.J., (ed.), *Thermodynamic Data for Biochemistry and Biotechnology*. Springer-Verlag, Berlin, pp. 402–427.
- (a) Marky, L.A. and Breslauer, K.J. (1987) *Biopolymers*, **26**, 1601–1620. (b) Gralla, J. and Crothers, D.M. (1973) *J. Mol. Biol.*, **78**, 301.
- (a) Johnson, W.C., Jr. (1985) In Glick, D., (ed.), *Methods of Biochemical Analysis*. John Wiley & Sons, New York, pp. 61–163. (b) Ivinov, V.I., Minchenkova, L.E., Schyolkina, A.K. and Poletayev, A.I. (1973) *Biopolymers*, **12**, 89–110.
- Gray, D.M., Ratliff, R.L. and Vaughan, M.R. (1992) In Lilley, D.M.J. and Dahlberg, J.E., (eds.) *Methods in Enzymology*, Academic Press, New York, Vol. 211A, pp. 389–406.
- Frederick, C.A., Grable, J., Melia, M., Samudzi, C., Jen-Jacobson, L.J., Wang, B.-C., Green, P., Boyer, H.W. and Rosenberg, J.M. (1984) *Nature*, **309**, 327–331.
- Hare, D.R., Wemmer, D.E., Chou, S.-H., Drobny, G. and Reid, B.R. (1983) *J. Mol. Biol.*, **171**, 319–336.