Synthesis of a trans-syn thymine dimer building block. Solid phase synthesis of CGTAT[c,s]TATGC

John-Stephen Taylor* and Ian R. Brockie

Department of Chemistry, Washington University, St Louis, MO 63130, USA

Received January 26, 1988; Revised and Accepted May 6, 1988

ABSTRACT

The synthesis of a building block for the sequence specific introduction of the trans-syn thymine dimer into oligonucleotides via solid phase DNA synthesis technology is described. CGTAT[c,s]TATGC was synthesized in 48% overall yield by a partially automated procedure. The stepwise coupling yield for addition of the trans-syn thymine dimer building block was 58%. The dimer containing oligonucleotide was characterized by 500 MHz $^1$H COSY and NOESY spectroscopy and 202.5 MHz $^{31}$P NMR. The $^1$H chemical shifts for the trans-syn thymine dimer unit of the decamer were found to be quite similar to those found for the trans-syn thymine dimer of TpT. Upon photolysis at 254 nm, CGTAT[c,s]TATGC was converted to a major product which coeluted with authentic CGTATTATGC and a minor product which coeluted with authentic CGTAT[c,s]TATGC, further supporting the presence of an intact trans-syn thymine dimer unit.

INTRODUCTION

Exposure of DNA to ultraviolet light leads to the formation of cyclobutane, (6-4) (1) and Dewar (2) photoproducts at dipyrimidine sequences. Failure to repair such lesions has been linked with such genetic diseases as xeroderma pigmentosum and skin cancer (3). Correlations between photolesions and mutations have been hampered by the lack of general methods for the preparation of well-characterized photolesion-containing DNA for physical, enzymological and mutagenesis studies. Photolysis of oligonucleotides containing unique TpT sites has been used to site-specifically introduce cis-syn thymine dimers for enzymological, (4) biophysical (5) and NMR studies (6). Such an approach, however, is not practical if the sequence of interest contains multiple photoproduct sites, or if the desired photoproduct is a minor one, as is the trans-syn thymine dimer. Recently, we reported the synthesis of a building block for the sequence specific incorporation of cis-syn thymine dimers into oligonucleotides by solid phase phosphoramidite DNA synthesis technology (7). Herein we report the synthesis of a building block for the sequence specific introduction of trans-syn thymine dimers, the other member of the cyclobutane pyrimidine...
Figure 1. trans-syn thymine dimer and the building block 1a for its sequence-specific incorporation into DNA by solid phase DNA synthesis technology.

dimer class, into oligonucleotides (Figure 1). We also report the synthesis and spectroscopic characterization of a decamer containing this photoproduct.

The ratio of trans-syn to cis-syn thymine dimers is 1:7 in photolyzed denatured DNA and less than 1:50 in native DNA (1). This suggests that trans-syn dimer production might be enhanced in transcriptionally active genes. Though the stereochemistry of trans-syn photoproducts formed in genomic DNA is not known, its major stereoisomer is probably the same as that established for the trans-syn photoproduct of the dinucleotide TpT.

This photoproduct has the TpC6S, pTC6S absolute configuration resulting from dimerization of a Tp(syn), pT(anti) glycosyl conformation. Formation of this photoproduct within a B-DNA helix necessitates that the hydrogen bond between the Tp-N3 proton and the complementary N1 of adenine be broken. A more intriguing possibility is that the trans-syn photoproducts are formed at B/Z-DNA junctions. Z-DNA has the base pairs flipped by 180° about the roll axis relative to those of B-DNA (8). A TpT unit at a B/Z-DNA junction with the 5'-thymine in the Z-DNA section and the 3'-thymine in the B-DNA section would lead to a cyclobutane dimer of trans-syn stereochemistry. Whether or not the trans-syn photoproducts produced in duplex DNA result from local B/Z-DNA junctions is unknown.

Unlike cis-syn cyclobutane dipyrimidine dimers, trans-syn photolesions are not substrates for yeast photoreactivating enzyme (9). It is not known whether trans-syn dimers are substrates for the uvrABC excinuclease. In fact much less is known about structure-activity relationships of
trans-syn dimers than the cis-syn dimers precisely because they are formed in such low yield. Perhaps a repair enzyme specific for trans-syn lesions exists. Likewise, the mutagenicity of this class of photoproducts is unknown and it is possible that these minor isomers are highly mutagenic.

MATERIALS AND METHODS

Abbreviations

DMT, 4,4′-dimethoxytrityl; TBDMS, tert-butyldimethylsilyl; TpT, thymidylyl-(3′-5′)-thymidine; TMP, trimethylphosphate; TPS, 3-(trimethylsilyl)propionic acid, sodium salt; THF, tetrahydrofuran; DMAP, N,N-di-methylaminopyridine; TMS, tetramethylsilane.

Materials

Phosphoramidites were from American Bionetics, Inc. (Fisher). Moisture sensitive reactions were conducted in oven-dried apparatus under positive nitrogen pressure. Solid starting materials were rendered anhydrous by stripping with toluene and pumping overnight under high vacuum. Aldrich Gold Label tetrazole was sublimed at 100°C and 0.005 torr prior to use. Anhydrous acetonitrile for syntheses was Aldrich Gold Label, from Kilo-Lab metal cylinders; acetonitrile for HPLC was Fisher HPLC grade. THF was dried by distillation from sodium/benzophenone; pyridine by distillation from calcium hydride; methylene chloride by storage over 4 Å sieves. TLC plates were Kieselgel 60, F254 plates from Merck. Analytical reverse phase C-18 chromatography was performed on a Waters µ-bondapak C-18 column (4.6 mm X 25 cm). Preparative silica chromatography was performed on a Rainin Dynamax column (2 cm X 30 cm). Anion exchange chromatography was performed on a Nucleogen DEAE 60-7 column (10 mm X 125 mm). Gradient HPLC was performed on a system composed of two Waters Model 501 pumps, a Rheodyne 7125 sample injector, and an ISCO V-4 variable wavelength detector with a 5 mm pathlength heat exchanger HPLC flow cell. Gradient formation was controlled by an Apple IIe based system, incorporating an Adalab data acquisition card, Chromadapt interface module, and Chromatochart software. Doubly distilled water was departmental distilled water redistilled through a Corning MP-1 still. Flash chromatography was performed with 40-63 µm silica from Baker. Solid phase DNA synthesis was conducted on a Biosearch SAM I Series 2 automated DNA synthesis machine. 300 MHz $^1$H, 121.5 MHz $^{31}$P and 75 MHz $^{13}$C NMR experiments were performed on a Varian XL-300 instrument with a 5 mm probe. 500 MHz $^1$H and 202.5 MHz $^{31}$P NMR experiments were performed on a Varian VXR-500 instrument. UV spectral data were acquired on a Bausch and Lomb
Spectronic 1001 spectrophotometer interfaced to an Apple IIe microcomputer with Bausch and Lomb wavelength scanning software, version 2.02. Infrared spectra were taken on a Perkin-Elmer 283B and calibrated with polystyrene. Chemical ionization mass spectrometry (CI-MS) was performed on either a Finnigan 3300 GC-MS adapted for FAB-MS, or a VG-ZAB-SE Double Focusing Mass Spectrometer (ZAB-MS).

Chemistry

**Chemical stability of the trans-syn dimer of TpT.** 7. The ammonium salt of the trans-syn dimer of TpT, 7, was prepared by sensitized photolysis of TpT, 7, according to a published procedure (10) and purified by C-18 HPLC. The $^1$H and $^{31}$P NMR spectra before and after treatment of this compound with concentrated ammonium hydroxide at 55°C for 6 h were compared.

**NO-Tf(s)(OCMe)OTBDMS. 4a.** Compounds 2ab were photolyzed as previously described (7). After removal of the cis-syn isomer B (compound 8a) from the ABCD mixture by flash chromatography, the trans-syn isomer C (compound 4a) was separated by preparative scale silica gel HPLC on a 2.54 cm X 25 cm Rainin Dynamax column with 4% methanol/dichloromethane at a flow rate of 8 mL/min. Trans-syn isomer C eluted at 31 min whereas isomers A and D eluted together at 36 min. From 2.0 g of ACD mix, 730 mg of compound 4a was isolated (36% yield). mp 172°C dec; $^1$H NMR (CDC1$_3$, ppm from TMS) 5.98 (dd, J=10.3, 5.4 Hz, pTH$^1$), 5.32 (dd, J=10.7, 4.9 Hz, TpH$^1$), 5.09 (b, TpH$^3$), 4.13 (m, pTH$^3$), 4.08 (b, TpH$^4$), 3.94 (d, J=7.0 Hz, TpC6H), 3.87 (b, pTH$^4$), 3.81 (d, J=11 Hz, POCH$_3$), 3.66 (m, TpH$^5$/H$^5''$), 3.64 (d, J=7.2 Hz, TpC6H), 3.66 (m, TpH$^2$), 3.60 (m, pTH$^5$/H$^5''$), 2.39 (m, TpH$^2$), 2.03 (m, pTH$^2$), 1.82 (m, pTH$^2$), 1.53 (s, TpC5CH$_3$), 1.48 (s, pTC5CH$_3$) 0.88 (s, tert-butyl CH$_3$), 0.08 (s, 2 SiCH$_3$); $^{31}$P NMR (ppm from TMP) -3.46; $^{13}$C NMR (CDC1$_3$, referenced to CDCl$_3$ at 77.00 ppm) 169.43, 169.35, 152.14, 151.16 (carbonyls), 90.86 (TpC1$'$), 85.18 (d, J$_{POCC}=-10.9$ Hz, TpC4$'$), 83.12 (pTC1$'$), 82.71 (d, J$_{POCC}=6.8$ Hz, pTC4$'$), 80.38 (d, J$_{POC}=-3.8$ Hz, TpC3$'$), 69.84 (pTC3$'$), 65.88 (s, TpC6), 65.64 (d, J$_{POC}=-3.0$ Hz, TpC5$'$), 62.93, 59.05 (TpC5$'$ and TpC6), 54.60 (d, methyl phosphate, J$_{POC}=5.1$ Hz), 47.46, 44.99 (two C5), 37.07, 35.97 (two C2$'$) 25.57 (tert-butyl methyls), 22.62, (TpC5CH$_3$), 20.46 (pTC5CH$_3$), 17.80 (quaternary carbon of tert-butyl) 17.71 (SiCH$_3$); UV (methanol) no $\lambda_{max}$ > 200 nm; IR (KBr, cm$^{-1}$) 3440, 3089, 2951, 2851, 1712, 1461, 1407, 1355, 1254, 1228, 1101, 1027, 834, 779; FAB-MS (glycerol) 675 (M+1, 25), 577 (54), 451 (26), 333 (100), 319 (41); high resolution FAB-MS calcd for C$_{27}$H$_{44}$N$_4$O$_{12}$P$i^+$ 675.2463, found 675.2463; TLC (10% MeOH/CH$_2$Cl$_2$) R$_f$ 0.50.
HO-T(5,6)PO(OMe)T-OH. 5a. Compound 4a (730 mg, 1.08 mmol) was dissolved in 40 mL of 20% 0.1 N HCl/acetic acid and incubated at room temperature for 34 h. The solvent was stripped off by rotary evaporation in vacuo, and further stripped down twice from 50:50 (v:v) toluene/methanol to yield the crude product as a white foam. Flash chromatography was performed on a 2.5 cm X 18 cm silica gel column packed with 4% methanol/dichloromethane. The compound was eluted with 150 mL portions of 4, 5, 6, 8, 10, and 12% methanol/dichloromethane followed by 300 mL of 15% methanol/dichloromethane. The purified product 5a was obtained as a white solid in 91% yield (550 mg, 0.98 mmol). mp 150-154d; 1H NMR (CD3CN plus trace DMSO-D6, referenced to CD2HCN at 1.93 ppm) 5.86 (dd, J=10.3, 5.4 Hz, H1'), 5.19 (m, H1'), 3.71 (d, J=11.0 Hz, POCH3), 1.31, 1.28 (2s, C5CH3); 31P NMR (CD3CN plus trace DMSO-D6, ppm from TMP in CDCl3 (insert)) -1.880; 13C NMR (75 MHz, CD3CN plus trace DMSO-D6, referenced to CD3CN at 118.2 ppm) 171.10, 170.58, 153.27, 151.64 (carbonyls), 90.78, 85.24, 84.11, 82.85 (d, JPOC=7.1 Hz), 81.17 (d, JPOC=5.0 Hz), 69.73, 67.03, 65.89, 63.02, 58.39 (C1', C3', C4', C5', C6), 54.92 (d, POCH3, JPOC=5.0 Hz), 48.41, 45.15 (C5), 36.77, 36.39 (C2'), 22.89, 20.01 (C5CH3); UV (methanol) no Amax > 200 nm; IR (KBr, cm-1) 3400, 1710, 1690, 1470, 1360, 1230, 1025, 760; ZAB-MS (glycerol) Positive ions: 561 (M+1, 17), 463 (17), 337 (18), 225 (28), 211 (8), 127 (22), 113 (13); ZAB-MS (glycerol) Negative ions: 559 (M-1, 52), 461 (12), 335 (72), 209 (17), 125 (77), 111 (100); TLC (10% methanol/90% dichloromethane) Rf < 0.10.

DMTO-T(5,6)PO(OMe)T-OH. 6a. 4,4'-dimethoxytrityl chloride (400 mg, 1.18 mmol) was added to a solution of compound 5a (550 mg, 0.98 mmol) in 10 mL pyridine, and allowed to stir at room temperature for 3.5 h. The solution was then concentrated to a gum in vacuo, stripped twice from toluene, and once from dichloromethane. The product was purified by flash chromatography utilizing a 2.5 cm X 15 cm silica gel column packed in 0.5% pyridine/dichloromethane. Elution with 0-15% methanol/0.5% pyridine/dichloromethane in 150 mL 2% steps yielded the product 6a as a white solid in 55% yield (470 mg, 0.54 mmol). mp 160-170 d; 1H NMR (17.8°, CDCl3, ppm from TMS) 6.10 (t, J=14.2, 7.1 Hz, H1'), 5.42 (d, J=10.0 Hz, H1'), 3.68 (s, ArOCH3), 3.67 (d, J=10.8 Hz, POCH3), 1.38, 1.33 (2s, C5CH3); 31P NMR (17.8°, CDCl3, ppm from TMP in CDCl3) -5.37; 13C NMR (17.8°, CDCl3, referenced to CDCl3 at 77.0 ppm) 169.75, 169.26, 151.54, 150.61 (carbonyls), 158.27, 135.65, 135.46, 129.87, 129.80, 127.92, 127.69, 113.02 (aromatic carbons), 87.89, 86.19 (C1'), 82.38, 82.09, 81.80 (JPOC=7.0 Hz), 80.87
Nucleic Acids Research

$(J_{POC} = 6.0 \text{ Hz})$, 69.34, 66.47, 63.69, 63.42, 58.38 (C3', C4', C5', C6', trityl center), 55.17 (ArOCH$_3$), 54.63 (d, POCH$_3$, $J_{POC} = 6.0 \text{ Hz}$), 46.94, 46.56 (C5), 37.23, 36.21 (C2'), 20.86, 20.57 (C5CH$_3$); UV (methanol) $\lambda_{max} = 231 \text{ nm}$, $\epsilon_{max} = 2.2 \times 10^4 \text{ cm}^{-1} \text{M}^{-1}$; IR (KBr, cm$^{-1}$) 3400, 1705, 1610, 1510, 1465, 1450, 1180, 1035, 830; ZAB-MS (glycerol) Positive ions: M+1 not seen, 305 (100), 186 (74), 136 (37); ZAB-MS (glycerol) Negative ions: 861 (M-1, 17), 847 (M-CH$_3$, 82), 623 (12), 335 (20), 321 (16), 195 (17), 182 (98), 125 (97), 111 (64); TLC (8% methanol in dichloromethane) Rf 0.3

**DMTO-T(c.g)PO(OMe)T-P(OMe)(NPT's)**. To compound 6A (250 mg, 0.29 mmol) in 6.0 mL anhydrous dichloromethane was added diisopropyl-ethylamine (0.20 mL, 1.16 mmol) and freshly distilled N,N-diisopropylamino-methoxyphosphorochloridite (0.083 mL, 0.435 mmol) (11). The reaction was stirred at room temperature for 0.5 h, then diluted with 20 mL of ethyl acetate which had been washed with saturated sodium bicarbonate. The organic layer was extracted three times with saturated sodium bicarbonate, then the aqueous layers were back extracted with ethyl acetate. The combined organic layers were dried over sodium sulfate, filtered, and stripped in vacuo to a white foam. Silica gel chromatography in 45/45/10 (v:v:v) ethyl acetate/dichloromethane/triethylamine afforded the product 1a as a white solid in 94% yield (280 mg, 0.27 mmol). $^1$H NMR (300 MHz, CDCl$_3$, ppm from TMS) 6.10 (m, H1'), 5.42 (d, J=8.8 Hz, H1'), 3.78 (s, ArOCH$_3$), 3.66 (d, J=12.7 Hz, POCH$_3$), 1.48, 1.43 (2s, C5CH$_3$); $^{31}$P NMR (17.8°, CDCl$_3$, ppm from TMP in CDCl$_3$) 147.17, 146.58 (2s, phosphoramidite), -4.42, -4.47 (2s, phosphate).

**Synthesis of CGTAT[c,g]TATGC.** Solid phase synthesis procedures are as previously described (7). 1.1 g of 15.9 µmole/g HO-ATGC-CPG in a 50 mL thick walled centrifuge tube was washed with three 10 mL portions of anhydrous acetonitrile and dried in vacuo. The trans-syn building block (1a, 130 mg, 0.13 mmol) in 0.7 mL of anhydrous acetonitrile was then added, followed immediately by tetrazole (26.7 mg, 0.38 mmol) in 0.9 mL of anhydrous acetonitrile. After 30 min at room temperature with occasional swirling, the reaction solution was removed and the support washed five times with anhydrous acetonitrile, twice with diethyl ether and dried in vacuo. The solid support was reintroduced into the synthesizer, washed extensively with dry acetonitrile, capped and oxidized. The yield for the trans-syn thymine dimer coupling step was determined to be 58%. This was followed by successive coupling of A, T, G and C to complete the synthesis in 48% overall.
yield. The oligonucleotide was cleaved from the support with concentrated ammonium hydroxide for 1 h and fully deprotected with fresh concentrated ammonium hydroxide at 55°C in a sealed tube for 6 h. The oligonucleotide was purified by anion exchange chromatography on a Nucleogen DEAE 60-7 column, with a 50 minute 0-0.8 M KCl/20% acetonitrile/20 mM KH$_2$PO$_4$/$K_2$HPO$_4$, pH 7 gradient. The desired oligonucleotide eluted at 23.4 min at a flow rate of 2.0 mL/min. The combined HPLC fractions were dialyzed in MW 1000 cutoff tubing at 4°C twice against 20 mM sodium phosphate buffer, pH 7, for 24 h, to remove KCl and methanol and to exchange the ammonium ions for sodium ions. This was followed by dialysis at 4°C twice against doubly distilled water in the presence of Chelex for 24 h, to remove phosphate buffer and trace paramagnetic ions. The isolated yield of £ was determined spectrophotometrically to be 3.76 mg or 10.1%.

**Synthesis of CGTATTATGC, 10.** The synthesis and purification of 10 followed the general description outlined for £.

**Synthesis of CGTAT[c,s]TATGC, 11.** 11 was synthesized and purified as described for £ utilizing the cis-syn thymine dimer building block, 12.

**Analytical C-18 Chromatography of 9, 10, 11.** Analytical C-18 HPLC was performed utilizing a 24 min linear gradient of 1 mL/min 10-40% methanol/75 mM KH$_2$PO$_4$/$K_2$HPO$_4$ pH 7.0, and detecting at 260 nm. Compound £ eluted as one major peak with a retention time of 18.0 min. CGTATTATGC, 10, eluted as one major peak with a retention time of 20.4 min whereas CGTAT[c,s]TATGC, 11, eluted at 17.6 min.

**NMR Sample Preparation.** The sodium salts of £ and 10 were speed-vacced twice from D$_2$O followed by dissolution in 0.7 mL "100%" D$_2$O, 0.2 mM TPS. $^1$H NMR spectra were acquired at 500 MHz and $^{31}$P NMR spectra at 202.5 MHz.

**Photolysis of 9.** 0.06 mM aqueous solutions (50 µL) of decamer £ at 0°C was exposed to 254 nm light from a Model UVGL-25 Mineralight lamp for 15 min, at a distance of 1 cm. The reaction mixture was assayed by analytical C-18 HPLC, under the conditions described above.

**RESULTS AND DISCUSSION**

**Synthesis.** The route to the trans-syn building block 1a is quite similar to that previously described for the synthesis of the cis-syn building block (7). A number of steps have been optimized and a more reactive phosphoramidite functionality has been selected.

We have previously shown that compounds 2ah (Figure 2), a mixture of two diastereomers epimeric at phosphorus, leads to four diastereomeric products...
Figure 2. Structures of the synthetic intermediates and products.

upon acetophenone sensitized photolysis; two trans-syn, and two cis-syn cyclobutane dimers, compounds 4ab and 8ab (7). These were labelled A through D, in order of their elution on a C-18 column. The trans-syn compound C (48) could be obtained by preparative scale silica gel HPLC purification after separation of the cis-syn compound B (8A) from the ABCD mixture by flash silica chromatography. We have found that isolation of compound C by preparative silica gel chromatography is preferable to preparative C-18 chromatography originally described (7).

Another modification of the procedure used for the preparation of the cis-syn thymine dimer building block was the reversal of the tritylation and desilylation steps. This allowed the use of 0.02 N HCl in acetic acid for the desilylation step leading to a reproducibly higher overall yield of the target phosphoramidite 1a than could be obtained using the original sequence of steps. Originally a methylmorpholinylphosphonite was used as coupling unit based on its known stability to chromatography on silica gel in the presence of triethylamine (12). Since then we have determined that the more reactive diisopropylaminophosphoramidites are also stable to such chromatography conditions. The phosphoramidite, 1a, was found to be a mixture of diastereomers on the NMR time scale based on the appearance of two phosphoramidite and two phosphotriester 31P NMR signals.

The trans-syn thymine dimer was incorporated into a decanucleotide to insure duplex formation at concentrations required for future NMR experiments and enzymatic ligation procedures. In order to facilitate studies on the
photochemistry of the central TpT unit and its trans-syn and cis-syn
cyclobutane dimer photoproducts the sequence CGTATTATGC was chosen as it is
devoid of additional dipyrimidine sequences.

It was important to know prior to the synthesis of the target
oligonucleotide whether or not the trans-syn thymine dimer would be stable
to the conditions used to remove the commonly used acyl protecting groups for
the amino groups of the bases A, C and G. Treatment of the trans-syn
photoproduct of TpT, compound 2i, with concentrated ammonium hydroxide at 55°C
for 6 h led to no observable change in the $^1$H NMR spectrum, suggesting that
it would indeed survive the final deprotection step.

The synthesis of CGTAT[c,s]TATGC was conducted in a semi-automated
fashion in order to insure the highest possible coupling yield for the building
block. By conducting the critical coupling step in a reaction flask, the
phosphoramidite building block can be added quantitatively to the support,
thereby minimizing losses associated with introducing it into a machine. In
this way a 58% coupling yield was achieved utilizing 7.4 equivalents of the
building block.

The trans-syn thymine dimer containing decamer was cleaved off the
resin with concentrated ammonium hydroxide at room temperature followed by
complete removal of all the protecting groups with concentrated ammonium
hydroxide. The use of thiophenol was not required. The crude decamer was then
purified by Nucleogen anion exchange HPLC resulting in an oligonucleotide that
was one major peak by analytical C-18 chromatography (Figure 3a), and >95% pure
by NMR analysis (Figure 4b). The decamer without a photodimer, CGTATTATGC, 1Q,
and the cis-syn thymine dimer containing decamer, CGTAT[c,s]TATGC, 11,
were similarly prepared for comparison purposes.

**Photochemistry of CGTAT[c,s]TATGC.** It is well known that
cis-syn and trans-syn thymine dimers photorevert to thymine upon
exposure to 254 nm light. For example, the photoequilibrium between TpT and
cyclobutane dimers at at 254 nm lies in favor of TpT by a factor of
approximately three (11). Exposure of CGTAT[c,s]TATGC, 2, to 254 nm light
for 15 min (Figure 3b) resulted in its conversion to a major product which
coeluted with CGTATTATGC, 1Q (data not shown) and three minor products. One of
these three minor photoproducts coeluted with synthetic CGTAT[c,s]TATGC, 11
(Figure 3c). The other minor products may be due to the impurities present in
the sample of 2 or to other photoproducts of 1Q. The photochemistry of the
synthetic decamer 2 is, however, consistent with the presence of an intact
trans-syn thymine dimer subunit which first photoreverts to a TpT subunit
Figure 3. Analytical gradient reverse phase HPLC traces of (A) CGTAT[c,s]TATGC, 2. (B) the resulting mixture after photolysis of 2 at 254 nm for 15 min, (C) coinjection of photoreverted 2 with CGTAT[c,s]TATGC, 11, (D) CGTATTATGC, 10. Elution time in min, detection at 260 nm. Details in the Experimental Section.

upon exposure to 254 nm light. Upon continued exposure, the TpT subunit resulting from the photoreversion then dimerizes to form the cis-syn thymine dimer. The photochemistry of 2, 10 and 11 is currently under investigation.

**Spectroscopy.** $^1$H and $^{13}$C chemical shift assignments for compound 4a (see experimental and Table I) were made possible by use of $^1$H COSY, $^1$H-$^{13}$C HETCOR and truncated driven NOE difference experiments. The COSY spectrum gave the deoxyribose ring $H1'$-$H2'$,$H2''$-$H3'$-$H4'$ connectivities as
Figure 4. Details of the $^1$H NMR spectra in D$_2$O at 30°C in ppm from TPS (A) 300 MHz spectrum of *trans-syn* photoproduct of TpT, 2. (B) 500 MHz spectrum of CGTAT[t,s]TATGC, 2. (C) 500 MHz spectrum of CGTATTATGC, 10.

5.98-2.03,1.82-4.13-3.87 and 5.32-3.46,2.39-5.09-7.09. Assignment of these proton connectivities to the individual sugar rings was made possible by the $^1$H-$^1$C HETCOR experiment. The H3' resonances at 4.13 and 5.09 ppm correlate with C3' resonances at 69.84 and 80.38 ppm respectively. Since the 80.38 ppm signal is coupled to phosphorus it can be assigned to TpC3'. TpH4' could not be assigned based on the COSY spectrum and was determined to be 4.08 ppm by its correlation to a 85.18 ppm $^1$C resonance. This resonance could only have been due to a Cl' or C4' carbon but since the pTC4' and both Cl' carbons were assigned, the 85.18 ppm must be due to TpC4'. TpH6 could be assigned based on the observed NOE between it and TpH1'. In turn TpCH$_3$ could be assigned based on the observed NOE between it and the TpH6 proton. In confirmation of this assignment an NOE was observed between pTC$_3$ and pTH6. The proton assignments determined are consistent with the chemical shifts observed for the *trans-syn* photoproduct of TpT (10, 13). In particular the TpH2' signal is approximately 1 ppm downfield from typical values for H2' protons, which for the *trans-syn* dimer of TpT has been attributed to deshielding induced by the T5C2 carbonyl group which is poised overhead (10).

The assignment of the $^1$H NMR spectrum of the *trans-syn* thymine dimer containing decamer (Figure 4) is not complete. However, most of the proton
Table I. \(^1\)H and \(^{31}\)P NMR Data for the trans-syn Thymine Dimer Unit of CGTAT\([t,s]\)TATGC, T\([t,s]\)T and 4a. Shifts in ppm, Coupling Constants in Hz.

<table>
<thead>
<tr>
<th></th>
<th>CGTAT([t,s])TATGC(^a)</th>
<th>T([t,s])T(^b)</th>
<th>4a(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tp C(_5)CH(_3)</td>
<td>1.20</td>
<td>1.51</td>
<td>1.53</td>
</tr>
<tr>
<td>H(_1)'</td>
<td>5.14</td>
<td>5.28</td>
<td>5.32</td>
</tr>
<tr>
<td>H(_6)</td>
<td>4.05</td>
<td>4.27</td>
<td>3.94</td>
</tr>
<tr>
<td>H(_2)'</td>
<td>3.27</td>
<td>3.34</td>
<td>3.46</td>
</tr>
<tr>
<td>H(_2)'</td>
<td>2.40</td>
<td>2.62</td>
<td>2.39</td>
</tr>
<tr>
<td>H(_3)'</td>
<td>4.72</td>
<td>4.77</td>
<td>5.09</td>
</tr>
<tr>
<td>H(_4)'</td>
<td>nd</td>
<td>4.16</td>
<td>4.08</td>
</tr>
<tr>
<td>J(_{H1}'-H2)'</td>
<td>10.5</td>
<td>10.2</td>
<td>10.2</td>
</tr>
<tr>
<td>J(_{H1}'-H2)'</td>
<td>4.9</td>
<td>4.5</td>
<td>4.4</td>
</tr>
<tr>
<td>pT C(_5)CH(_3)</td>
<td>1.13</td>
<td>1.45</td>
<td>1.48</td>
</tr>
<tr>
<td>H(_1)'</td>
<td>5.50</td>
<td>5.71</td>
<td>5.98</td>
</tr>
<tr>
<td>H(_6)</td>
<td>3.70</td>
<td>4.27</td>
<td>3.64</td>
</tr>
<tr>
<td>H(_2)'</td>
<td>1.73</td>
<td>2.22</td>
<td>2.03</td>
</tr>
<tr>
<td>H(_2)'</td>
<td>1.88</td>
<td>2.03</td>
<td>1.82</td>
</tr>
<tr>
<td>H(_3)'</td>
<td>4.67</td>
<td>4.57</td>
<td>4.13</td>
</tr>
<tr>
<td>H(_4)'</td>
<td>3.88</td>
<td>3.92</td>
<td>3.87</td>
</tr>
<tr>
<td>J(_{H1}'-H2)'</td>
<td>10.8</td>
<td>10.2</td>
<td>10.1</td>
</tr>
<tr>
<td>J(_{H1}'-H2)'</td>
<td>4.0</td>
<td>4.5</td>
<td>5.2</td>
</tr>
<tr>
<td>J(_{H6-H6})</td>
<td>6.9</td>
<td>7.3</td>
<td>7.2</td>
</tr>
<tr>
<td>(^{31})P</td>
<td>-3.73 (1)</td>
<td>-3.98</td>
<td>-3.47</td>
</tr>
<tr>
<td></td>
<td>-3.92 (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-4.01 (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-4.07 (2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-4.18 (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-4.29 (3)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^{a}\) Sodium salt in D\(_2\)O at 30°C, \(^1\)H NMR relative to TPS, \(^{31}\)P NMR relative to external TMP in D\(_2\)O, numbers in parentheses are relative peak integrals. \(^b\) \(^1\)H NMR (10), \(^{31}\)P NMR (15). \(^c\) In CDCl\(_3\), \(^1\)H NMR relative TMS, \(^{31}\)P NMR in CDCl\(_3\) at 19°C relative to TMP in a CDCl\(_3\) insert.

Signals in the TpT unit of CGTAT\([t,s]\)TATGC have been assigned (Table 1). The thymine dimer methyl signals could be assigned to 1.20 and 1.13 ppm based on NOESY crosspeaks to their respective H6 proton signals at 4.05 ppm and 3.70 ppm (Figure 5). The H6 signal at 4.05 ppm has a strong NOE crosspeak to proton signals at 3.28 ppm and 5.14 ppm. In the COSY spectrum these two signals are part of the same J-coupled H1'-H2'-H2" network. Examination of a molecular model of the trans-syn dimer of TpT suggests that the 5.14 ppm signal is due to TSH1' since only the base in the syn conformation has a short T5H6-H1' distance. In an analogous fashion the signal at 3.28 ppm can be assigned TSH2'. T6H2' was assigned by first noting that T6H6 shows an NOE to signals at 1.88 and 1.73 ppm which are part of the same T6 H1'-H2'-H2"-H3'...
Figure 5. Schematic representation of the NOE crosspeaks observed for the trans-syn thymine dimer subunit of CGTAT[t,s]TATGC in a 2D phase-sensitive 500 MHz NOESY spectrum at 30° with a 200 msec mixing time.

The chemical shifts and coupling constants of the T5 and T6 H1' protons of the trans-syn containing decamer compare favorably with those of the trans-syn dimer of TpT as do most of the chemical shifts of the deoxyribose ring protons (10, 13). A strong crosspeak was observed between T6H1' and T6H4' indicating that the T6 sugar ring prefers a conformation, or set of rapidly equilibrating conformations, on the pseudorotation pathway between C2',-endo and C2',-endo (14). This is consistent with the C2',-endo conformation proposed for the pT sugar ring of the trans-syn thymine dimer of TpT (13). The overall similarity of the NMR spectral data of the trans-syn thymine dimer subunit of CGTAT[t,s]TATGC to that of the dinucleotide suggests that its local conformation is not greatly affected by substitution at either the 5'- or 3'-ends. How the conformation will change on going to a duplex decamer remains to be determined.

CONCLUSION

We have reported the synthesis of the phosphoramidite building block, compound 1a, for the sequence specific introduction of trans-syn thymine dimers into oligonucleotides by solid phase DNA synthesis methodology. The building block was obtained in three steps from the previously described...
compound 4a. The building block was used to synthesize the first trans-syn thymine dimer containing oligonucleotide, CGTAT[c,s]TATGC, which was characterized by $^1$H NMR spectroscopy and its photochemistry. The trans-syn and corresponding cis-syn thymine dimer containing oligonucleotides are now being used to construct unique photolesion containing duplexes and viruses for use in physical, enzymological and mutagenesis studies.

ACKNOWLEDGMENTS

This investigation was supported by PHS Grant number R01-CA40463, awarded by the National Cancer Institute, DHHS. The assistance of the Washington University High Resolution NMR Service Facility, funded in part through NIH Biomedical Research Support Shared Instrument Grant 1 S10 RR02004, and a gift from the Monsanto Company is gratefully acknowledged. We thank D. Garrett for the NMR spectrum of CGTATTATGC and D. Svoboda for the phosphorylating agent and a sample of CGTAT[c,s]TATGC. We thank Dr. Andrew Tyler for conducting fast atom bombardment mass spectrometry. Mass spectrometry data were obtained at the Washington University Mass Spectrometry Resource, which is supported by a grant from the NIH [RR00954].

*To whom correspondence should be addressed

REFERENCES