Synthesis and properties of oligodeoxynucleotides containing the analogue 2'-deoxy-4'-thiothymidine

E.Louise Hancox, Bernard A.Connolly¹ and Richard T.Walker*

School of Chemistry, University of Birmingham, Birmingham B15 2TT and ¹Department of Biochemistry and Genetics, University of Newcastle-upon-Tyne, Newcastle-upon-Tyne NE2 4HH, UK

Received April 8, 1993; Revised and Accepted June 9, 1993

ABSTRACT

The 2'-deoxythymidine analogue 2'-deoxy-4'-thiothymidine has been incorporated, using standard methodology, into a series of dodecadeoxynucleotides containing the *EcoRV* restriction endonuclease recognition site (GATATC). The stability of these oligodeoxynucleotides and their ability to act as substrates for the restriction endonuclease and associated methylase have been compared with a normal unmodified oligodeoxynucleotide. No problems were encountered in the synthesis despite the presence of a potentially oxidisable sulfur atom in the sugar ring. The analogue had very little effect on the melting temperature of the self-complementary oligodeoxynucleotides so synthesised and all had a CD spectrum compatible with a B-DNA structure. The oligodeoxynucleotide containing one analogue in each strand within the recognition site, adjacent to the bond to be cleaved (i.e. GAXATC, where X is 2'-deoxy-4'-thiothymidine), was neither a substrate for the endonuclease nor was recognized by the associated methylase. When still within the recognition hexanucleotide but two further residues removed from the site of cleavage (i.e. GATAXC), the oligodeoxynucleotide was a poor substrate for both the endonuclease and methylase. Binding of the oligodeoxynucleotide to the endonuclease was unaffected but the k_{cat} value was only 0.03% of the value obtained for the parent oligodeoxynucleotide. These results show that the incorporation of 2'-deoxy-4'-thionucleosides into synthetic oligodeoxynucleotides may shed light on subtle interactions between proteins and their normal substrates and may also show why 2'-deoxy-4'-thiothymidine itself is so toxic in cell culture.

INTRODUCTION

An understanding of the way in which proteins can recognise, bind and catalyse the reactions of nucleic acids is of great importance when considering the mechanisms of certain cellular processes. Many DNA-binding proteins (e.g. repressor proteins

and restriction endonucleases) show an extremely high specificity for a particular DNA sequence. The mechanisms by which this high selectivity is achieved are complex and depend upon a co-ordinated set of protein–DNA interactions which include:

1. Interaction of the protein with the unique functional group pattern presented in the major groove by a given base sequence (direct readout).

2. Interaction of the protein with an unusual altered sugar–phosphate backbone characteristic of a particular base sequence (indirect readout).

3. DNA distortion on binding to the protein with only the target site being able to support the distorted bound conformation.

Almost all specific DNA-binding proteins bind to the sugar–phosphate backbone and although this can engender selectivity (2 above), it often serves simply to increase the overall binding affinity.

The use of nucleobases which have been modified, either by the sequential deletion of a particular functional group or by the reversal of the relative positions of the bases within a sequence, has provided a useful method for studying the direct readout mechanism. This approach has yielded information concerning the interaction between RNA polymerase and its promoter site (1), and also the interaction of nucleic acids with repressor proteins (2–3), restriction endonucleases (4–7) and modification enzymes (7–8).

There has been very little investigation of the role played by the sugar moiety in such recognition processes usually for the very good reason that any significant change in the sugar moiety leads to a structure incompatible with naturally occurring polymers and hence makes their study less interesting. Any modifications to the sugar unit have been made primarily to increase the stability of the phosphodiester bond toward exonucleases thus enhancing the ability of the oligodeoxynucleotides to act as antisense inhibitors of gene expression (9–14). A large variety of nucleoside analogues containing modified glycones have been prepared as potential chemotherapeutic agents (15) but most are unsuitable for incorporation into oligodeoxynucleotides for the reason described above.

Recently however we reported the synthesis of a nucleoside analogue 1-(2'-deoxy-4'-thio- β -D-ribofuranosyl)thymine,

* To whom correspondence should be addressed

(2'-deoxy-4'-thiothymidine) (16), in which the furanose ring oxygen is replaced by a sulfur atom, thereby making 2'-deoxy-4'-thiothymidine an isoelectronic analogue of natural 2'-deoxythymidine. This nucleoside analogue is known to possess potent anti-CMV properties (16), although the cytotoxicity of the analogue has limited its potential therapeutic applications and its incorporation into polynucleotides may be expected to be accommodated within a flexible secondary and tertiary structure.

This analogue has been incorporated into a set of dodecadeoxynucleotides and a comparative study of the properties of the modified oligodeoxynucleotides and the parent oligomer is reported here.

RESULTS AND DISCUSSION

Synthesis of the 2'-deoxy-4'-thiothymidine monomer

There are currently four synthetic strategies by which the nucleoside analogue 2'-deoxy-4'-thiothymidine can be prepared. Dyson *et al.* and Secrist *et al.* have independently published synthetic routes (16–17), Uenishi and co-workers subsequently described a synthesis from acyclic precursors (18) and more recently Huang and Hui have reported a revised synthetic route (19).

The route chosen to 2'-deoxy-4'-thiothymidine, the details of which will be published elsewhere, closely follows that originally described by Dyson *et al.*

The 5'-O-DMT-3'-O-[2-cyanoethyl(diisopropylamino)-phosphinoyloxy]-2'-deoxy-4'-thiothymidine (3) was obtained as a diastereoisomeric mixture following the procedure described in the literature (20). The dimethoxytritylation of 2'-deoxy-4'-thiothymidine (1) proceeded smoothly in a yield of 82% to give compound (2) and the subsequent phosphitylation reaction gave a quantitative yield of the phosphoramidite (3) (Figure 1).

Synthesis and purification of the oligodeoxynucleotides

The monomer (3) was used to synthesise three modified dodecadeoxynucleotides containing the *EcoRV* recognition sequence.

The *EcoRV* restriction/modification system involves two enzymes which recognise the same hexameric base sequence GATATC in duplex DNA but catalyse different reactions at this site.

The *EcoRV* sequence was selected on the basis of the substantial amount of data that has already been published on the cleavage of both an unmodified oligodeoxynucleotide and a variety of derivatives in which either the base or the phosphate moiety has been modified. Another attraction of the *EcoRV* system is that the genes for both proteins have been cloned and sequenced (21) and that overproducing strains for both the endonuclease (22) and the methylase (23) are available.

Figure 1. Synthesis of the monomer.

A sequence containing at least twelve base-pairs was required to make a suitable substrate for the restriction enzyme. The sequence for which most data was available was the self-complementary, palindromic dodecamer shown below:

d(G-A-C-G-A-T-A-T-C-G-T-C)

(C-T-G-C-T-A-T-A-G-C-A-G)d—Complementary Strand

This contains the GATATC *EcoRV* recognition hexamer as the central six bases. The restriction endonuclease cuts between the central T and dA residues and the modification methylase adds a CH₃ group to the 6-amino position of the first dA.

Four of these self-complementary dodecanucleotides containing the *EcoRV* recognition sequence were prepared: the unmodified sequence (odn-1), two sequences (odn-2 and odn-3) in which a single 2'-deoxythymidine residue within the recognition site was replaced by 2'-deoxy-4'-thiothymidine (X), and one in which both 2'-deoxythymidines in the recognition site were altered (odn-4):

d(GACGATATCGTC) (odn-1)

d(GACGAXATCGTC) (odn-2)

d(GACGATAXCGTC) (odn-3)

d(GACGAXAXCGTC) (odn-4)

All oligodeoxynucleotides were synthesised on a 1 μmol scale using an automated solid phase technique. The couplings were successful and proceeded quantitatively, as estimated by the release of the orange dimethoxytrityl cation.

The removal of the base- and phosphate-protecting groups was achieved in one step by treatment with aqueous ammonia at 50°C overnight and the DMT-protected oligodeoxynucleotides were purified by reverse-phase HPLC. The DMT group was removed with 80% acetic acid and the fully deprotected oligodeoxynucleotides were desalted by gel filtration. The purities and yields of the three oligodeoxynucleotides containing 2'-deoxy-4'-thiothymidine were very similar to those of the control. This indicates that 2'-deoxy-4'-thiothymidine is perfectly compatible with the DNA synthesis and deblocking conditions of the standard phosphoramidite method.

Characterisation of the oligodeoxynucleotides

Base composition analysis. The composition of all of the oligodeoxynucleotides was determined by digestion with snake venom phosphodiesterase (from *Crotalus durissus*) and calf intestinal alkaline phosphatase. A mixture of standard solutions of the five deoxynucleosides was analysed by reverse-phase HPLC and gave the order of elution as dC, dG, dT, dA and dX. The relative amounts of each of the deoxynucleosides in the digestion mixtures were estimated from the peak areas with the use of the deoxynucleoside extinction coefficients (24). The

Table 1. Summary of properties of oligodeoxynucleotides.

OLIGODEOXY-NUCLEOTIDE	BASE COMPOSITION (THEORY)	T _m /°C	CD	<i>EcoRV</i> SUBSTRATE?	METHYLASE SUBSTRATE?
odn-1	Not Determined	52	B-DNA	YES	YES
odn-2	3.5C 2.7G 2.1T 3.2A 1.0X (3C 3G 2T 3A 1X)	49	B-DNA	NO	NO
odn-3	3.0C 3.1G 1.8T 2.9A 1.0X (3C 3G 2T 3A 1X)	47	B-DNA	YES	YES
odn-4	3.2C 3.1G 1.0T 3.3A 1.8X (3C 3G 1T 3A 2X)	55	B-DNA	NO	Not Determined

nucleoside compositions were shown to be consistent with the proposed sequences (Table 1) and confirmed that the analogue was stable to the reaction conditions and in particular no sulfone or sulfoxide was produced during the iodine treatment.

Melting curves. Melting curves for the four oligodeoxynucleotides were determined using a programmable temperature unit with Peltier element control. The melting profiles for odn-1, odn-2, odn-3 and odn-4 are shown in Figure 2.

All four oligodeoxynucleotides exhibited a double-stranded to single-coil transition between 47°C and 55°C. The value of 52°C obtained for odn-1 is in good agreement with the literature value of 53°C (25). In all of these melting profiles, the absorbance

Figure 2. Melting curves for d(GACGATATCGTC) (odn-1) (52°C), d(GATGAXATCGTC) (odn-2) (49°C), d(GACGATAXCGTC) (odn-3) (47°C) and d(GACGAXAXCGTC) (odn-4) (55°C).

begins to increase at around 30°C and reaches a maximum at about 80°C.

These results imply that the entire set of oligodeoxynucleotides will be predominantly in a double-helical form at the temperatures which were used for the enzyme assays. The self-complementary nature of the base sequence would theoretically allow the formation of a hairpin oligomer but no evidence was found to indicate the presence of such a structure.

The smooth nature of the curves is indicative of a single melting transition, although the exact shape of each profile varied.

To provide an explanation for the minor differences in melting temperature, one must consider two consequences which will arise from the presence of the 2'-deoxy-4'-thiothymidine residue in the oligodeoxynucleotides. Firstly, the bulky nature of the sulfur will probably serve to destabilise the modified oligomers and secondly, the lipophilicity of 2'-deoxy-4'-thiothymidine (as demonstrated by its retention on reverse-phase HPLC) will provide an opportunity for a hydrophobic interaction between either the sulfur atoms in neighbouring chains or the sugar and/or

Figure 3. CD Spectra for d(GACGATATCGTC) (odn-1), d(GATGAXATCGTC) (odn-2), d(GACGATAXCGTC) (odn-3) and d(GACGAXAXCGTC) (odn-4).

base and the sulfur atom of the neighbouring chain, thereby stabilising the duplex. There must obviously be a balance between these two factors and the temperature at which the modified duplexes will melt will be determined by which factor is predominant.

The slightly higher T_m value obtained for odn-4 may reflect the attraction involving the sulfur atoms which is not possible in odn-1, whereas the lower T_m values for odn-2 and odn-3 may indicate that the destabilising effect of the sulfur is the dominant feature here. The relative retention times using reverse-phase HPLC observed for the four oligodeoxynucleotides support this speculation in that the order of elution is odn-1, odn-2 and odn-3 (unresolved) and finally odn-4. This implies that odn-4 is the most lipophilic, and that the decreased lipophilicity observed for odn-2 and odn-3 may make the steric bulk the important factor in the melting of these oligomers.

CD spectra. The circular dichroism (CD) spectra of the dodecamers were measured and the results are presented in Figure 3.

As expected, the parent oligomer produces a spectrum typical of B-DNA (25) and the oligodeoxynucleotides containing the 2'-deoxy-4'-thiothymidine residue(s) gave almost identical spectra, suggesting that these analogues also possess B-DNA structures. (The anomalous spectrum obtained for odn-4 in the region below 240 nm was due to the spectrophotometer and this local difference should be disregarded.)

The data obtained from the measurement of the T_m and CD spectra indicate that the parent and modified oligodeoxynucleotides have the same double-helical B-DNA structures with very similar stabilities.

Interaction of the oligodeoxynucleotides with endonuclease

The cleavage experiments with *EcoRV* endonuclease were performed at 25°C using oligodeoxynucleotide present at a concentration of 10 μ M. Under these conditions the oligodeoxynucleotide substrates are double-stranded. Odn-1 was cleaved completely within fifteen minutes using the enzyme at a concentration of 42.3 μ M. The reaction was monitored by reverse-phase HPLC which showed complete conversion to two faster eluting products (d(GACGAT) and d(pATCGTC)).

Odn-3 was subjected to the same conditions but cleavage was complete only after the addition of further enzyme and overnight incubation. Odn-2 and odn-4 showed no cleavage even after overnight incubation with extra enzyme.

Figure 4. A model demonstrating the deformation required of 2'-deoxythymidine residues within the *EcoRV* recognition sequence when bound.

The position of the cleavage was confirmed by isolating the products of the cleavage reaction and subjecting them to digestion with snake venom phosphodiesterase and alkaline phosphatase. The analysis of the digestion mixture by HPLC gave deoxynucleoside compositions which were consistent with the expected position of cleavage.

Determination of the initial rates of cleavage for odn-1 and odn-3.

The rate of cleavage of the oligodeoxynucleotides by *EcoRV* was determined using the continuous spectrophotometric assay developed by Waters and Connolly (26).

The oligodeoxynucleotide (0.4 O.D₂₅₄ units) was hydrolysed by a known concentration of the enzyme. In the case of the parent oligodeoxynucleotide the enzyme was diluted 100-fold in order to obtain reliable data, whereas the modified oligodeoxynucleotide required undiluted enzyme to produce a cleavage rate which could be monitored over a reasonable time scale.

The assays revealed that the modified oligodeoxynucleotide is cleaved at about 1% of the rate at which the parent oligomer is cleaved.

Determination of the kinetic parameters — K_m and k_{cat} . The assays were performed on the phosphorylated oligodeoxynucleotide d(pGACGATAXCGTC) using substrate concentrations (S_0) in the range 1 μ M to 5 μ M.

It has been reported (27) that the *EcoRV* enzyme binds to all DNA sequences, including its recognition sequence, with equal affinity. It was revealed that this may be extended to include sequences which incorporate sugar-modified nucleosides since a value of 1.0 μ M was obtained for K_m indicating that the strength of binding of odn-3 to the enzyme is identical to that observed for the parent oligodeoxynucleotide (26). The value for k_{cat} of 0.01 min⁻¹ is about 0.03% of that obtained for the parent (26). This result is in reasonable agreement with the data obtained from the initial rate experiments which indicated an approximate 100-fold reduction in the rate of cleavage of odn-3 relative to odn-1.

The *EcoRV* enzyme is a dimer that possesses two-fold symmetry and extensions that wrap around the DNA helix. The molecule is roughly U-shaped with the DNA-binding site situated in the bowl of the U. The loops that extend inwards from the sides of the U include the recognition loops which become ordered when the enzyme binds its target sequence (28–29). The crystal structures show that DNA containing the cognate GATATC sequence is extensively distorted on binding. It has been proposed that this distortion places the scissile phosphodiester bond at the active site and simultaneously assembles the Mg^{2+} -binding site (Mg^{2+} is an essential cofactor for hydrolysis and its binding is thought to be composed of an oxygen ligand from the scissile phosphate and two oxygen ligands from aspartic acid side chains). The distortion positions these phosphate and aspartate groups such that they form a Mg^{2+} -binding pocket. With DNA that lacks the recognition sequence, no distortion occurs on binding, thus the Mg^{2+} -binding site is not formed, the cation cannot bind and hydrolysis cannot occur. This differential affinity for Mg^{2+} has been proposed to account for all of the rate differences seen ($\geq 10^6$) in the cleavage of cognate and non-cognate sites. Experimental evidence in support of this postulate is the high Mg^{2+} -binding affinity demonstrated by the endonuclease-cognate DNA complex whereas enzyme-non-specific DNA species bind Mg^{2+} very poorly (30–32).

A possible explanation for the results obtained for the interaction of the *EcoRV* enzyme with the modified oligodeoxynucleotides involves the ability of the modified oligodeoxynucleotide to distort. If one considers a model where C⁴-O-C¹ of the furanose ring lie in a plane, when the DNA is deformed, each of the sugar moieties in the constituent nucleosides must also undergo conformational changes (Figure 4).

It has been proposed that in the case of the enzyme-bound specific sequence dGGGATATCCC, the sugar moiety in the first 2'-deoxythymidine in the recognition site adopts a conformation more reminiscent of A-form DNA with the C-3' in an endo-conformation. The distortion of the second 2'-deoxythymidine is much less drastic and resembles the usual C2'-endo conformation (Figure 4). It can be postulated that 2'-deoxy-4'-thiothymidine is less able to pucker into the A-form conformation required of the first 2'-deoxythymidine residue in the oligodeoxynucleotide in order to contact the *EcoRV* effectively. If this is the case, it would explain why odn-2 and odn-4 are not substrates for *EcoRV* endonuclease. However, in the case of odn-3, a very minor distortion of the sugar is required and if this is achieved—albeit by overcoming a greater energy barrier than that required for the distortion of 2'-deoxythymidine in this position—then the cleavage would proceed, presumably at a considerably reduced rate.

Interaction of the oligodeoxynucleotides with methylase

The parent oligodeoxynucleotide odn-1 and the modified oligodeoxynucleotides odn-2 and odn-3 were incubated with the *EcoRV* modification methylase and the methyl donor (AdoMet) at 37°C. The reactions were monitored by HPLC—the methylated product eluted directly after the substrate as confirmed by a coinjection of the reaction mixture with starting oligodeoxynucleotide—which revealed that odn-1 was methylated within 30 minutes. The methylation of odn-3 was complete after the addition of more enzyme and overnight incubation, but in contrast, odn-2 appeared completely resistant to methylation. The product oligodeoxynucleotide resulting from the methylation of odn-3 was collected and analysed by base composition analysis which showed the production of one equivalent of d⁶MeA and the loss of one equivalent of dA.

The single peak resulting from the incubation of odn-2 with the enzyme and AdoMet was also collected and analysed to confirm that the oligodeoxynucleotide had indeed not been methylated.

Determination of the rates of methylation for odn-1 and odn-3.

The rates of methylation for known concentrations of the oligodeoxynucleotides were determined by HPLC using known quantities of enzyme. The value of 9.66 nmol min⁻¹ mg enzyme⁻¹ obtained for odn-1 agreed with the literature value of 10 nmol min⁻¹ mg enzyme⁻¹. The methylation of odn-3 was shown to proceed at a rate of 6.49 nmol min⁻¹ mg enzyme⁻¹ (approximately 67% of the rate of methylation observed for the parent oligomer).

Due to the lack of crystallographic data for *EcoRV* methylase, one can only conclude that the first 2'-deoxythymidine in the *EcoRV* restriction site is very important in the recognition of the oligodeoxynucleotide by the enzyme, and that the second 2'-deoxythymidine—while less important—is also required for the recognition process.

CONCLUSIONS

The analogue 2'-deoxy-4'-thiothymidine can be incorporated into oligodeoxynucleotides using the conventional methodology applicable to synthesis on a solid phase support. Previous work on the conformation of 2'-deoxy-4'-thiothymidine in solution and in the crystal form has shown that despite significant changes in bond lengths (2'-deoxythymidine values given in parentheses) of C1'-S, 1.84 (C1'-O, 1.43); C4'-S, 1.83 (C4'-O, 1.46), the effect is counteracted by corresponding changes in bond angles C4'-S-C1', 94° (C4'-O-C1', 110°), which leaves the remainder of the structure largely undisturbed (33). This gives a predominant conformation in solution of South (thiofuranose ring), anti (base orientation) which is precisely as is found for 2'-deoxythymidine itself. Analysis of the NMR data indicates that for 2'-deoxy-4'-thiothymidine, a second conformer with a North-type puckered thiofuranose ring is present in solution to the extent of 27% but whether this has any significance when the molecule is present in an oligodeoxynucleotide has yet to be determined.

What is clear from these results is that when so present in an oligodeoxynucleotide, the analogue can be accommodated with minimal disruption to the stability of the double helix. However the consequent effect upon protein-nucleic acid recognition has been shown, in some cases of endonuclease and methylase recognition, to be absolute. The ability to inhibit the cleavage by modification of a site which is remote from the cleavage position should find application in the molecular study of protein-nucleic acid recognition especially as further specific chemical reactions to produce the sulfone or sulfoxide should be possible either at the monomer or polymer level.

The toxicity shown by 2'-deoxy-4'-thiothymidine in cell culture (17) could well be due to the ability of some enzymes such as kinases (2'-deoxy-4'-thiothymidine is a good substrate for both cytosol and VZV thymidine kinase) to recognise the analogue and enable its incorporation into DNA to be achieved with catastrophic consequences upon subsequent metabolism where more subtle interactions are required.

EXPERIMENTAL

¹H NMR spectra were recorded on Bruker AC300 and AMX400 spectrometers, relative to an internal tetramethylsilane standard. ³¹P NMR spectra were obtained on a Jeol GX270 (270 MHz) spectrometer, relative to an internal phosphoric acid reference. Samples for elemental analyses were dried at 60°C under vacuum for 4 hours (prolonged heating resulted in degradation). The presence of the cited solvents is supported by appropriate NMR evidence.

Pyridine, triethylamine and diisopropylethylamine were heated under reflux over calcium hydride, distilled and stored over potassium hydroxide pellets. Tetrahydrofuran was heated under reflux over sodium in the presence of benzophenone until a deep blue colour had developed. The solvent was then distilled and used immediately.

HPLC for oligodeoxynucleotides was performed using a Varian 5000 liquid chromatograph fitted with an Apex-I 250 mm × 4.5 mm octadecylsilyl (5μ particle size) column. Detection was at 254 nm using a Varian UV-50 variable detector and chart recorder. Columns were run at room temperature for deoxynucleoside composition analysis and at 55°C for oligodeoxynucleotide purification and kinetic studies.

The following HPLC buffers were used: Buffer A—0.1 M triethylammonium acetate pH 6.5, containing 5% acetonitrile; Buffer B—0.1 M triethylammonium acetate pH 6.5, containing 65% acetonitrile.

HPLC Gradients were prepared using the above buffers at a flow rate of 1 ml min⁻¹. Gradient I: A linear gradient of 25–85% buffer B over 20 minutes; Gradient II: 100% buffer A for 5 minutes, followed by a linear gradient of 0–25% buffer B over 25 minutes; Gradient III: A linear gradient of 5–45% buffer B over 25 minutes; Gradient IV: 5–20% buffer B over 30 minutes. The following reaction buffers were used: Buffer α : 50 mM HEPES–NaOH pH 7.5, containing 55 mM NaCl and 25 mM MgCl₂; Buffer β : 25 mM KH₂PO₄ pH 7.5, containing 100 mM NaCl; Buffer γ : 50 mM HEPES–NaOH pH 7.5, containing 1 mg/ml bovine serum albumen; Buffer δ : 50 mM Tris pH 7.6, containing 10 mM MgCl₂, 5 mM dithioerythritol, 0.1 mM EDTA and 0.1 mM spermidine; Buffer ξ : 50 mM HEPES–NaOH pH 7.5, containing 100 mM NaCl; Buffer τ : 100 mM Tris pH 7.5.

T_m Determinations were obtained with a Perkin-Elmer PTP-1 (single cell holder) UV spectrophotometer.

Spectrophotometric assays were performed using a Kontron Instruments Uvikon 930 spectrophotometer.

Alkaline phosphatase (molecular biology grade from calf intestine, 1 unit/ μ L) was purchased from Northumbria Biologicals, and snake venom phosphodiesterase (*Crotalus durissus*) was obtained from Boehringer-Mannheim. *EcoRV* restriction endonuclease (42.3 μ M dimer) and *EcoRV* modification methylase (0.7 mg/ml) were purified as described (23, 34). AdoMet was purchased from Sigma as its chloride salt, and stock solutions were made up in 1 mM HCl immediately prior to use.

5'-O-(4,4'-dimethoxytriphenylmethyl)-2'-deoxy-4'-thiothymidine (2)

To a solution of 2'-deoxy-4'-thiothymidine (1 g, 3.87 mmol) in pyridine was added DMAP (0.023 g, 0.19 mmol) and triethylamine (0.7 ml, 5.4 mmol) followed by dimethoxytritylchloride (1.6 g, 4.7 mmol). The reaction mixture was stirred overnight at room temperature. The solvent was evaporated *in vacuo*, co-evaporated from toluene and then methanol. Eluting from a silica (7734) column with 90:10 chloroform–ethanol containing 2% triethylamine yielded the title compound as a cream powder (1.76 g, 82%). δ (¹H NMR; 270 MHz, DMSO-d₆) 11.35(1H, s, NH), 8.33(1H, s, H-6), 7.45–6.83(13H, m, aromatic), 6.28–6.20 (1H, t, H-1'), 5.31–5.35(1H, m, OH-3'), 4.36–4.29(1H, m, H-3'), 3.75–(6H, s, OCH₃), 3.43–3.36(1H, m, H-4'), 3.28–3.23-(2H, m, H-5'), 2.22–2.03(2H, m, H-2'), 1.63(3H, s, CH₃); m/z: 583(M+Na)⁺, 560(M)⁺, 303(DMT)⁺. (FOUND: C;65.1 H;5.9 N;4.5; C₃₁H₃₂N₂O₆S.0.08EtOH requires: C;65.4 H;6.1 N;4.6).

5'-O-(4,4'-dimethoxytriphenylmethyl)-3'-O-[2-cyanoethyl (diisopropylamino)phosphinoyloxy]-2'-deoxy-4'-thiothymidine (3)

5'-O-(4,4'-dimethoxytriphenylmethyl)-2'-deoxy-4'-thiothymidine (0.46 g, 0.82 mmol) was dissolved in THF (30 ml) and stirred under a stream of nitrogen at 0°C. Diisopropylethylamine (0.94 ml, 5.34 mmol) was added *via* syringe, followed by cyanoethyl N,N-diisopropylchlorophosphoramidite (0.43 ml, 1.61 mmol). After warming to room temperature over 2 hours the solvent was

removed *in vacuo* and the residue taken up in ethyl acetate. This was washed (saturated sodium bicarbonate followed by brine), dried (MgSO₄), filtered and concentrated *in vacuo* to reveal a colourless oil. The residue was purified by eluting from a silica (7734) column with 70:30 dichloromethane–ethyl acetate containing 2% triethylamine to give the title compounds as a mixture of diastereoisomers (0.62 g, 99%). δ (³¹P NMR; 300 MHz, CDCl₃) 148.81, 148.64; m/z: 760(M)⁺, 635(M-Thymine)⁺. (FOUND: C;61.1 H;6.1 N;6.9; C₄₀H₄₉N₄O₇ PS.0.037CHCl₃ requires: C;61.2 H;6.3 N;7.1).

Synthesis and characterisation of oligodeoxynucleotides

The oligodeoxynucleotides were synthesised on a 1 μ mol scale using a fully automated Applied Biosystems 381A DNA synthesizer and standard phosphoramidite chemistry. After removal of base-labile protecting groups (35% aqueous ammonia, 50°C overnight) the crude DMT-containing oligodeoxynucleotide was purified by reverse-phase HPLC using gradient I⁷. Removal of the DMT group with 80% acetic acid (1 hour, room temperature) gave oligodeoxynucleotides that appeared pure as judged by reverse-phase HPLC using gradient III. Oligodeoxynucleotides were desalted by passage over Pharmacia NAP-25 disposable gel filtration columns. Where necessary 5'-phosphorylation was carried out using polynucleotide kinase and ATP as described⁷.

The relative hydrophobicity of the oligodeoxynucleotides was determined by reverse-phase HPLC using gradient IV. The order of elution was as follows: odn-1 (8.05 minutes), odn-2 and odn-3 (9.7 minutes) and odn-4 (11.4 minutes).

Base composition analysis was performed in buffer α by digestion of the appropriate oligodeoxynucleotide (10 μ l) with snake venom phosphodiesterase (2 μ l) and alkaline phosphatase (2 units, 2 μ l). The product deoxynucleosides were analysed by reverse-phase HPLC using gradient II, and the peaks were identified by comparison with authentic nucleoside samples. The nucleoside compositions were calculated from the peak areas and the following extinction coefficients (M⁻¹cm⁻¹) at 254 nm: dC (6.1 $\times 10^3$), dG (11.8 $\times 10^3$), T (8.7 $\times 10^3$), dA (14.7 $\times 10^3$).

Melting temperature determination

Melting temperature determinations were performed on solutions containing 0.5–1.0 O.D₂₅₄ unit of oligodeoxynucleotide duplex in buffer β (1 ml). The temperature was raised from 15°C to 90°C using temperature increments of 1°C per minute.

Measurement of CD spectra

The appropriate oligodeoxynucleotide (1 O.D₂₅₄ unit) was dissolved in buffer β (1 ml).

EcoRV endonuclease cleavage reaction

A solution of the appropriate oligodeoxynucleotide (odn-1: 346 μ M, 6 μ l; odn-2: 334 μ M, 6 μ l; odn-3: 346 μ M, 6 μ l) in buffer α (25 μ l) was diluted to a total volume of 50 μ l with distilled water. The endonuclease (2.5 μ l) was added and the progress of the reaction was monitored by reverse-phase HPLC using gradient II. A further 2.5 μ l of enzyme was added to those solutions which showed no or partial reaction, and the mixtures were incubated at 25°C for an additional 24 hours.

Those oligodeoxynucleotides which showed substrate activity were purified using the above HPLC system and the identity of the hexameric products was determined by base-composition analysis.

Initial rate data for endonuclease cleavage reaction

Stock solutions of the endonuclease were prepared by dilution with buffer γ .

To a 1 ml volume of buffer α containing sufficient odn-3 to give an initial absorbance of approximately 0.4 O.D.₂₅₄ unit in a 1 cm path length quartz cuvette was added *EcoRV* endonuclease (20 μ l). The assay was performed at 30.7°C by thermostating the cell compartment and the absorbance at 254 nm was recorded 20 times a minute for 30 minutes.

The initial rate of hydrolysis for odn-1 was determined using a 100-fold diluted enzyme.

K_m and k_{cat} values for odn-3

To a 1 ml volume of buffer α containing sufficient phosphorylated odn-3 to give an oligodeoxynucleotide concentration in the range 1 μ M to 5 μ M (as determined by the initial absorbance reading) in a 1 cm path length quartz cuvette was added *EcoRV* endonuclease (20 μ l). The d(pGACGATAXCGTC) concentrations used were 1.1, 2.0, 3.7 and 5.0 μ M. The assay was performed as described in the literature (26) and the initial rates were determined from the linear portion of the plot of absorbance at 254 nm vs. time. K_m and k_{cat} values were calculated from a plot of substrate concentration/initial rate vs. substrate concentration.

EcoRV modification methylase reaction

The appropriate oligodeoxynucleotide (odn-1: 1730 μ M, 1.5 μ l; odn-2: 536 μ M, 6 μ l; odn-3: 1145 μ M, 2.8 μ l) was diluted into buffer ξ (40 μ l). S-Adenosyl methionine (AdoMet) (1 mM, 10 μ l) was added, followed by the enzyme (40 μ g) and the mixture was incubated at 37°C. The progress of the reaction was monitored by reverse-phase HPLC using gradient III. If, after 2 hours, the mixtures showed no or partial reaction, further enzyme (40 μ g) and an additional aliquot of AdoMet (10 μ l) were added and the solutions were incubated for 24 hours. The products of the methylation were purified using the same HPLC system and the identities of the products were determined by base-composition analysis.

Initial rate data for methylation of odn-1 and odn-3

A solution of the odn-3 (1145 μ M, 3 μ l) was diluted to a total volume of 80 μ l with buffer α . AdoMet (1 mM, 20 μ L) was added followed by *EcoRV* modification methylase (25 μ g).

The initial rate of methylation for odn-1 was determined in an analogous manner using 5 μ g of the enzyme.

The reactions were monitored over 2 hours by the removal at regular intervals of 15 μ l aliquots which were quenched into NaOH (0.5 M, 5 μ l). When the aliquots were analysed by HPLC, the proportion of starting oligodeoxynucleotide that had been methylated could be determined by estimating the relative areas of the appropriate peaks (7).

ACKNOWLEDGEMENTS

The authors wish to thank the Wellcome Foundation for financial support (to E.L.H.), Dr M.J. Gait and his colleagues at the MRC Laboratory, Cambridge for their help with synthesising the oligodeoxynucleotides and to Dr D.M. Brown F.R.S. at the same laboratory for the use of his T_m apparatus. Thanks are also due to Dr S. Halford (University of Bristol) for making unpublished data available and for helpful discussions. B.A.C. is a Research Fellow of the Lister Institute of Preventive Medicine.

REFERENCES

1. Dubendorff, J.W., de Haseth, P.L., Rosendahl, M.S. and Caruthers, M.H. (1987) *J. Biol. Chem.* **262**, 892–898.
2. Yansura, D.G., Goeddel, D.V., Cribbs, D.L. and Caruthers, M.H. (1977) *Nucleic Acids Res.* **4**, 723–737.
3. Caruthers, M.H. (1980) *Acc. Chem. Res.* **13**, 155–160.
4. Dwyer-Hallquist, P., Kezdy, F.J. and Agarwal, K.L. (1982) *Biochemistry* **21**, 4693–4700.
5. Seela, F. and Driller, H. (1986) *Nucleic Acids Res.* **14**, 2319–2332.
6. McLaughlin, L.W., Benseler, F., Graeser, E., Piel, N. and Scholtissek, S. (1987) *Biochemistry* **26**, 7238–7245.
7. Newman, P.C., Nwosu, V.U., Williams, D.M., Cosstick, R., Seela, F. and Connolly, B.A. (1990) *Biochemistry* **29**, 9891–9901.
8. Brennan, C.A., van Cleve, M.D. and Gumpert, R.I. (1986) *J. Biol. Chem.* **261**, 7278–7286.
9. Augustyns, K., Vandendriessche, F., Van Aerschot, A., Busson, R., Urbanke, C. and Herdewijn, P. (1992) *Nucleic Acids Res.* **20**, 4711–4716.
10. Richardson, F.C., Richardson, K.K., Kroin, J.S. and Hertel, L.W. (1992) *Nucleic Acids Res.* **20**, 1763–1768.
11. Morvan, F., Genu, C., Rayner, B., Gosselin, G. and Imbach, J.-L. (1990) *Biochem. Biophys. Res. Commun.* **172**, 537–543.
12. Sagi, J., Szemzo, A., Szecsi, J. and Orvos, L. (1990) *Nucleic Acids Res.* **18**, 2133–2140.
13. Perbost, M., Lucas, M., Chavis, C., Pompon, A., Baumgartner, H., Rayner, B., Griengl, H. and Imbach, J.-L. (1989) *Biochem. Biophys. Res. Commun.* **165**, 742–747.
14. Bellon, L., Barascut, J.-L., Maury, G., Divita, G., Goody, R. and Imbach, J.-L. (1993) *Nucleic Acids Res.* **21**, 1587–1593.
15. Perigaud, C., Gosselin, G. and Imbach, J.-L. (1992) *Nucleosides Nucleotides* **11**, 903–945.
16. Dyson, M.R., Coe, P.L. and Walker, R.T. (1991) *J. Med. Chem.* **34**, 2782–2786.
17. Secrist, J.A., Tiwari, K.N., Riordan, J.M. and Montgomery, J.A. (1991) *J. Med. Chem.* **34**, 2361–2366.
18. Uenishi, J., Motoyama, M., Nishiyama, T. and Wakabayashi, S. (1991) *J. C.S. Chem. Commun.* ??? 1421–1422.
19. Huang, B. and Hui, Y. (1993) *Nucleosides Nucleotides* **12**, 139–147.
20. Sinha, N.D., Biernat, J., McManus, J. and Koster, H. (1984) *Nucleic Acids Res.* **12**, 4539–4557.
21. Bougueleret, L., Schwartzstein, M., Tsugita, A. and Zabeau, M. (1984) *Nucleic Acids Res.* **12**, 3654–3676.
22. Bougueleret, L., Schwartzstein, M., Tsugita, A. and Zabeau, M. (1985) *Nucleic Acids Res.* **13**, 3823–3839.
23. Nwosu, V.U., Connolly, B.A., Halford, S.E. and Garnett, J. (1988) *Nucleic Acids Res.* **16**, 3705–3720.
24. Fassman, G.D. (1975) *CRC Handbook of Biochemistry and Molecular Biology*, 3rd Edition. *Nucleic Acids* Vol. I, pp. 65–213.
25. Connolly, B.A. and Newman, P.C. (1989) *Nucleic Acids Res.* **13**, 4957–4974.
26. Waters, T.R. and Connolly, B.A. (1992) *Anal. Biochem.* **204**, 204–209.
27. Taylor, J.D., Badcoe, I.G., Clarke, A.R. and Halford, S.E. (1991) *Biochemistry* **30**, 8743–8753.
28. Winkler, F.K. (1992) *Curr. Opin. Struct. Biol.* **2**, 235–238.
29. Winkler, F.K., Banner, D.W., Oefner, C., Tsernoglou, D., Brown, R.S., Heathman, S.P., Bryan, R.K., Martin, P.D., Petratos, K. and Wilson, K.S. (1993) *EMBO J.* **12**, 1781–1795.
30. Taylor, J.D. and Halford, S.E. (1989) *Biochemistry* **28**, 6198–6207.
31. Taylor, J.D. and Halford, S.E. (1992) *Biochemistry* **31**, 90–97.
32. Halford, S.E., Taylor, J.D., Vermote, C.L.V. and Vipond, I.B. (1993) In Eckstein, F. and Lilley, D.M.J. (eds), *Nucleic Acids and Molecular Biology*. Springer-Verlag, Vol 7.
33. Koole, L.H., Plavec, J., Liu, H., Vincent, B.R., Dyson, M.R., Coe, P.L., Walker, R.T., Hardy, G.W., Rahim, S.G. and Chattopadhyaya, J. (1992) *J. Am. Chem. Soc.* **114**, 9936–9943.
34. D'Arcy, A., Brown, R.S., Zabeau, M., Van Resandt, R.W. and Winkler, F.K. (1985) *J. Biol. Chem.* **260**, 1987–1990.