Rapid chemical synthesis of d(CACGTG) in milligram amounts by solution-phase phosphotriester chemistry

G V RAJENDRAKUMAR, N SIVAKAMISUNDARI and
K N GANESH*†

Centre for Cellular and Molecular Biology, Hyderabad 500 007, India
†*Present address*: Division of Organic Chemistry, National Chemical Laboratory,
Pune 411 008, India

Abstract. This paper reports an optimised method for chemical synthesis of short oligodeoxyribonucleotides in milligram amounts by solution-phase phosphotriester chemistry. The procedure is a slight modification of one reported earlier in the literature. Detailed experimental procedures for condensation, deprotection and purification are described. The efficiency of the method is demonstrated by the synthesis of a hexamer, d(CACGTG), in good yields and high purity. The homogeneity of the product and the sequence composition are shown by 270 MHz ^{1}H NMR spectrum.

Keywords. Oligodeoxyribonucleotide synthesis; chemical synthesis of d(CACGTG); ^{1}H NMR of oligonucleotides.

1. Introduction

In spite of rapid developments in synthetic methodologies for oligodeoxynucleotides (Gait 1984; Sonveaux 1986) the preparation of sufficient quantities of these in the high purity required for X-ray and NMR studies is still a formidable challenge. The choice of synthetic strategy for scaling-up depends on the required length and the choice of purification protocols becomes crucial when one needs large amounts of oligonucleotides in high purity. We have previously analysed the suitability of various supports (Ganesh 1985; Minganti *et al* 1985) for solid-phase synthesis of longer (~ 23-mer) oligodeoxynucleotides in milligram amounts. In this paper, we report an optimised procedure for solution-phase phosphotriester synthesis of d(CACGTG)** in milligram amounts. The method is of general application for synthesis of short oligodeoxynucleotides (6–8-mers).

---

*To whom all correspondence should be addressed
**Structure of DNA fragments is represented only by the base sequence, e.g., in d(CACGTG), d stands for deoxy.

2. Chemical synthesis

2.1 *Choice of synthetic method*

In the past, a lot of effort has been directed towards synthesis of oligodeoxynucleotides by solid-phase synthesis and by different chemical approaches. These methods result in smaller quantities of less pure products than does synthesis in solution but meet most of the present needs of molecular biologists. The solid-phase method works best for smaller amounts (from micrograms up to a milligram); the scaling-up to prepare larger amounts (milligrams) introduces considerable problems. (Gait 1984; Minganti *et al* 1985). However, little effort has been concentrated till now on improving the solution-phase synthesis. Solution-phase synthesis appears to be the best method for obtaining large amounts of short oligomers required for spectroscopic, structural and physiochemical studies. The ultimate purity of the product is high because of successive purifications after each step of condensation. The phosphotriester chemistry is the only method available at present for solution-phase synthesis and our present approach is therefore dependent on the use of this methodology.

2.2 *The synthetic cycle*

Our present phosphotriester method is an optimised extension of that reported by Chaudhari *et al* (1984). Here (scheme 1) the DNA chain is extended in the 3' → 5' direction as in normal solid-phase synthesis. Each synthetic cycle consists of three consecutive steps: (1) condensation, (2) deprotection and (3) purification. During the first step (scheme 1) a 3'-terminal block such as 1 is condensed with a 3'-(2-chlorophenyl) phosphate ester of an N-protected 2'-deoxyribonucleotide 2. The condensing agent used is either 1-(mesitylene-2-sulphonyl)-3-nitro-1,2,4-triazole (MSNT) or triisopropyl benzenesulphonyl chloride (TPSCl) in combination with a nucleophilic catalyst such as 1-methylimidazole (Efimov *et al* 1982). The reaction solvent used is acetonitrile for the first 3 cycles, and pyridine for further cycles. The progress of the reaction is monitored by thin layer chromatography over silica gel.

The 5',3'-O-protected dinucleotide (3) product of the first step is then deprotected at the 5'-position in the second step. This is achieved by treatment of the dinucleotide 3 with a solution of phenyl dihydrogen phosphate in chloroform: ethanol. A deep orange-red colour is produced instantaneously due to the liberated dimethoxytrityl cation. After work-up, the product 5 is taken to the third step of the cycle for rapid purification over silica gel.

The column was eluted first with chloroform (1-2 bed volumes) to remove non-nucleotidic impurities such as dimethoxytritanol liberated during the deprotection. The desired product 5 is then eluted with anhydrous tetrahydrofuran-pyridine (3:1) and recovered from the eluant. After drying, this product is used as the 5'-hydroxyl component for initiating the next cycle (scheme 1).
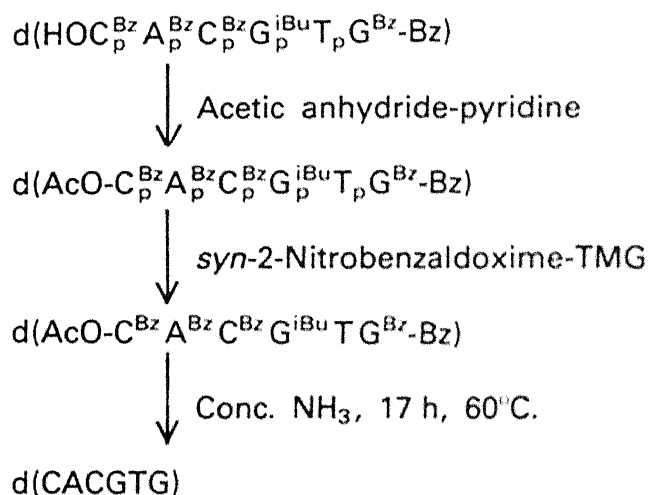
2.3 *Deprotection and purification*

At the end of the complete synthesis, the product oligomer obtained carries a terminal 5'-hydroxyl group (scheme 2). This is acetylated with acetic anhydride-

DMT = 4,4'-Dimethoxytriphenylmethyl
Bz = Benzoyl
iBu = Isobutyryl
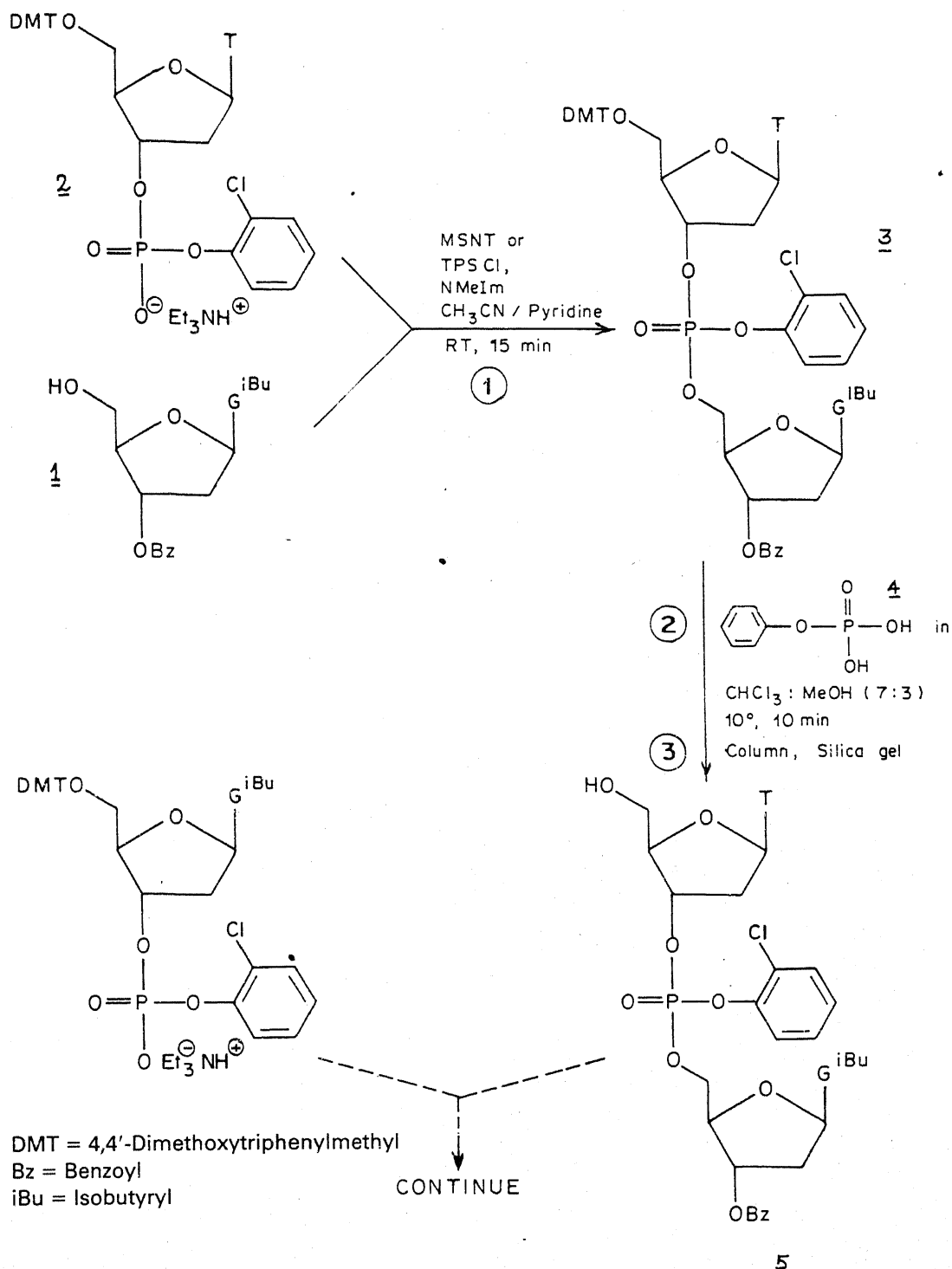
Scheme 1

pyridine and then subjected to a two-stage deprotection procedure. First, it is treated with *syn*-2-nitro benzaldoximate reagent to remove the 2-chlorophenyl phosphate protection and this is followed by treatment with concentrated ammonia to effect all N- and O-deprotections. The totally unblocked product is passed through a Sephadex G-15 gel filtration column to remove small molecule impurities. The product which elutes out in the void volume is then purified by fast

$$d(HOC_p^{Bz}A_p^{Bz}C_p^{Bz}G_p^{iBu}T_pG^{Bz}\text{-Bz})$$

↓ Acetic anhydride-pyridine

$$d(AcO\text{-}C_p^{Bz}A_p^{Bz}C_p^{Bz}G_p^{iBu}T_pG^{Bz}\text{-Bz})$$

↓ *syn*-2-Nitrobenzaldoxime-TMG

$$d(AcO\text{-}C^{Bz}A^{Bz}C^{Bz}G^{iBu}TG^{Bz}\text{-Bz})$$

↓ Conc. NH_3, 17 h, 60°C.

$$d(CACGTG)$$

d = Deoxy
Ac = Acetyl
A^{Bz} = 6-N-Benzoyl-2'-deoxyadenosine
C^{Bz} = 4-N-Benzoyl-2'-deoxycytidine
G^{iBu} = 2-N-Isobutyryl-2'-deoxyguanosine
p = Phosphate groups protected by 2-chlorophenyl

Scheme 2

protein liquid chromatography (FPLC) on an anion-exchange Mono Q column using a linear gradient from 0·1 M aqueous sodium hydroxide to 0·1 M aqueous sodium hydroxide containing 0·8 M sodium chloride. Figure 1 illustrates an analytical FPLC chromatogram of the crude product. The major peak from a preparative run was collected, neutralised to pH 7, and then desalted over a Sephadex G-15 column.

Using this procedure, we have synthesised oligomers (6–8-mers) in high purity and in milligram amounts. The isolated yields varied in the range of 70–85% per synthetic cycle leading to a 25–30% overall yield before FPLC purification.

3. Discussion

The modifications used in our procedure for solution-phase synthesis are based on previous experience (Effimov *et al* 1982; Rajendrakumar *et al* 1985) with solid-phase synthesis. The 3'-terminal nucleotide 1 used here carries an acetyl or benzoyl group at the 3' position whereas in the solid-phase synthesis it is attached to the polymer matrix. The phosphate component 2 added during every cycle is identical to the phosphotriester monomers of solid-phase synthesis. The solvent used for the condensation during the first three steps was acetonitrile since reactions are faster in this solvent compared to pyridine (Effimov *et al* 1982). However, after the trinucleotide stage, because of the low solubility of the 5'-hydroxyl component in acetonitrile, the solvent used was pyridine. The condensing agent used was either MSNT or TPSCl in combination with a nucleophilic catalyst such as 1-methylimidazole.
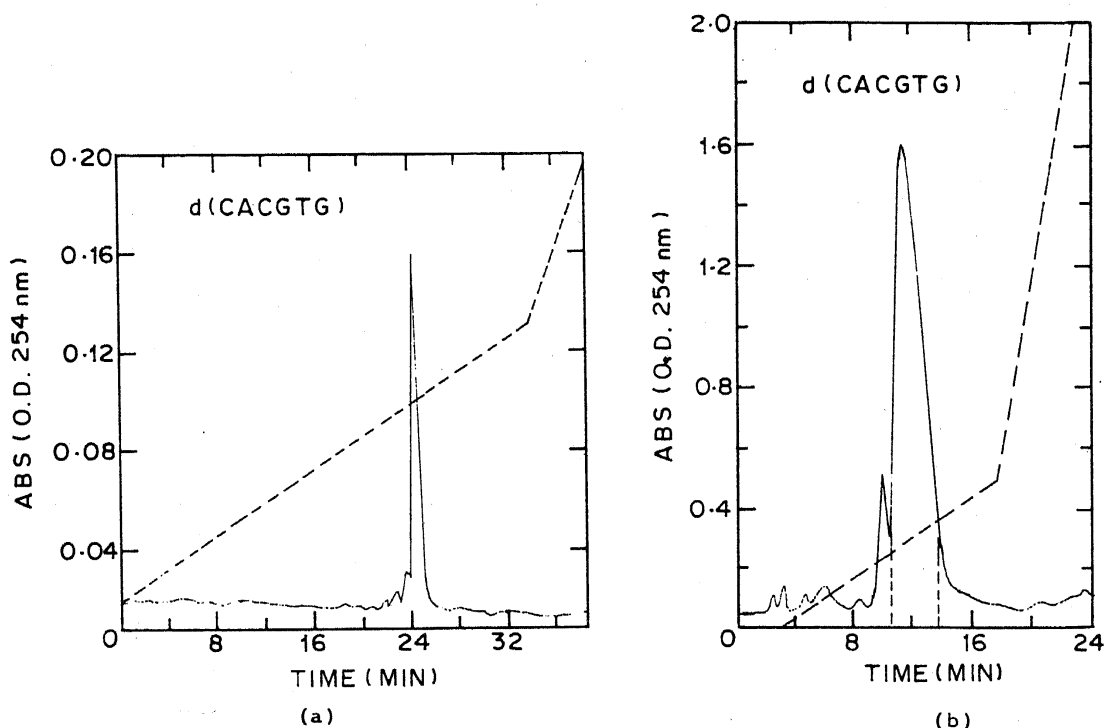
Figure 1. Analytical (a) and preparative (b) FPLC chromatograms. FPLC (Pharmacia) system consisted of two P 500 pumps, a GP 250 gradient programmer, a 5 MPa mixer, a V-7 sample injector and a DP-2 dual wavelength detector fixed at 254 nm. Column: Mono Q (Pharmacia) 5/5 (analytical) and 10/10 (preparative). Solvents: 0·1 M aq NaOH pH 12 (low) and 0·1 M aq. NaOH with 0·8 M NaCl *pH* 12 (high). Flow rate: 0·5 ml/min (analytical) and 1·5 ml/min (preparative).

The reactions can be conveniently monitored by thin layer chromatography over silica gel. The phosphate monomers $\underline{2}$ added at every cycle have low r_f value ($\sim 0·04 - 0·05$) and show up as orange-red spots (trityl-positive) on t.l.c. after acid-spray. The 5'-hydroxy component $\underline{1}$ cannot be visualised in a similar way; however, when heated after the acid-spray it shows up as a dark spot. The phosphate component $\underline{2}$ is always taken in slight excess (1·2 equivalents) over the hydroxy component $\underline{1}$. The completion of the condensation reaction is signalled by the total disappearance of $\underline{1}$ and appearance of a higher r_f trityl-positive spot due to $\underline{3}$. Beyond the trinucleotide stage, the phosphate components are taken in 1·4-equivalent excess over the hydroxy component to drive the reaction to completion. It should be pointed out that for the eventual success of this method, the hydroxy component must disappear totally at every step; otherwise it is very difficult to separate it from the product since they have very close r_f values.

The 5'-deprotection of the product $\underline{3}$ is done with phenyl dihydrogen phosphate $\underline{4}$. The reaction with this reagent in addition to being faster seems to cause less depurination as compared to either benzenesulphonic acid or dichloroacetic acid. The reaction can be assayed by t.l.c. as the trityl-positive product $\underline{3}$ is converted into a slightly slower moving trityl-negative product $\underline{5}$. The liberated tritanol moves almost with the solvent front. The end product of the cycle ($\underline{5}$) is purified by a rapid column chromatography over silica gel (Chaudhari *et al* 1982).

It can be seen from table 1 that the total time required for the first synthetic cycle is about 150 min. The four other cycles required for synthesis of a hexamer

Table 1. Time table for each cycle of coupling

		(Min)
Step 1	Reaction	15
	Work-up	15
Step 2	Reaction	10
	Work-up	10
Step 3	Column	60–80
Total		~150 min cycle

were carried out in the same way. After the five cycles had been completed, the material obtained after the final column chromatography was acetylated with acetic anhydride-pyridine before subjecting it to deprotection steps as in solid phase synthesis. After the ammonia treatment, the product was passed over a gel filtration column (Sephadex G-15) to separate the oligonucleotide product from the deprotecting agent and other small molecule impurities. The product (6–8 mers) elutes out in the void volume. The crude product thus obtained is almost always 90–95% pure as shown by analytical anion-exchange FPLC (figure 1). However, for spectroscopic and crystallographic work, this material is further purified by preparative FPLC over anion-exchange FPLC column. The detailed purification strategies for preparative purification, particularly optimised for self complementary sequences, are being reported elsewhere.

To substantiate the purity and authenticity of the product, we have shown here (figure 2) the 270 MHz ^{1}H NMR spectrum of the synthesised d(CACGTG). The aromatic proton signals from the six base residues, appearing in the region $\delta 7.0-8.5$ clearly support the base composition. The presence of a single thymidine residue is indicated by a lone methyl signal at $\delta 1.45$ and a singlet due to H6 at $\delta 7.05$. The presence of two cytidines is indicated by two doublets for H6 protons, one from each cytidine, at $\delta 7.2$ and $\delta 7.6$. The doublets for the corresponding H5 protons of the two cytidines occur at $\delta 5.92$ and $\delta 5.32$, the former overlapping with the glycosidic H1' sugar protons.

The hexamer d(CACGTG), which is self-complementary, exists as a duplex under the spectrum-recording conditions, as evidenced by a good dispersion and general non-equivalence of signals seen in the ^{1}H NMR spectrum. Single-stranded structures, being random coils, give averaged spectrum which is just an addition spectrum of four bases showing equivalence among bases of the same kind. The total assignments of all signals (sugar and base protons) in d(CACGTG) have been completed by two-dimensional NMR techniques, and from COSY and NOESY experiments. These assignments conform to the symmetry of a duplex structure. These and other spectroscopic analyses (CD, ^{31}P NMR) which are reported elsewhere have shown that this sequence has a distorted B-DNA conformation. The sequence d(CACGTG) is homologous to part of the bacteriophage P22 DNA operator region which is involved in interaction with mnt repressor protein. The structural results reported here are therefore of considerable importance to studies of mechanisms of DNA-protein interactions.
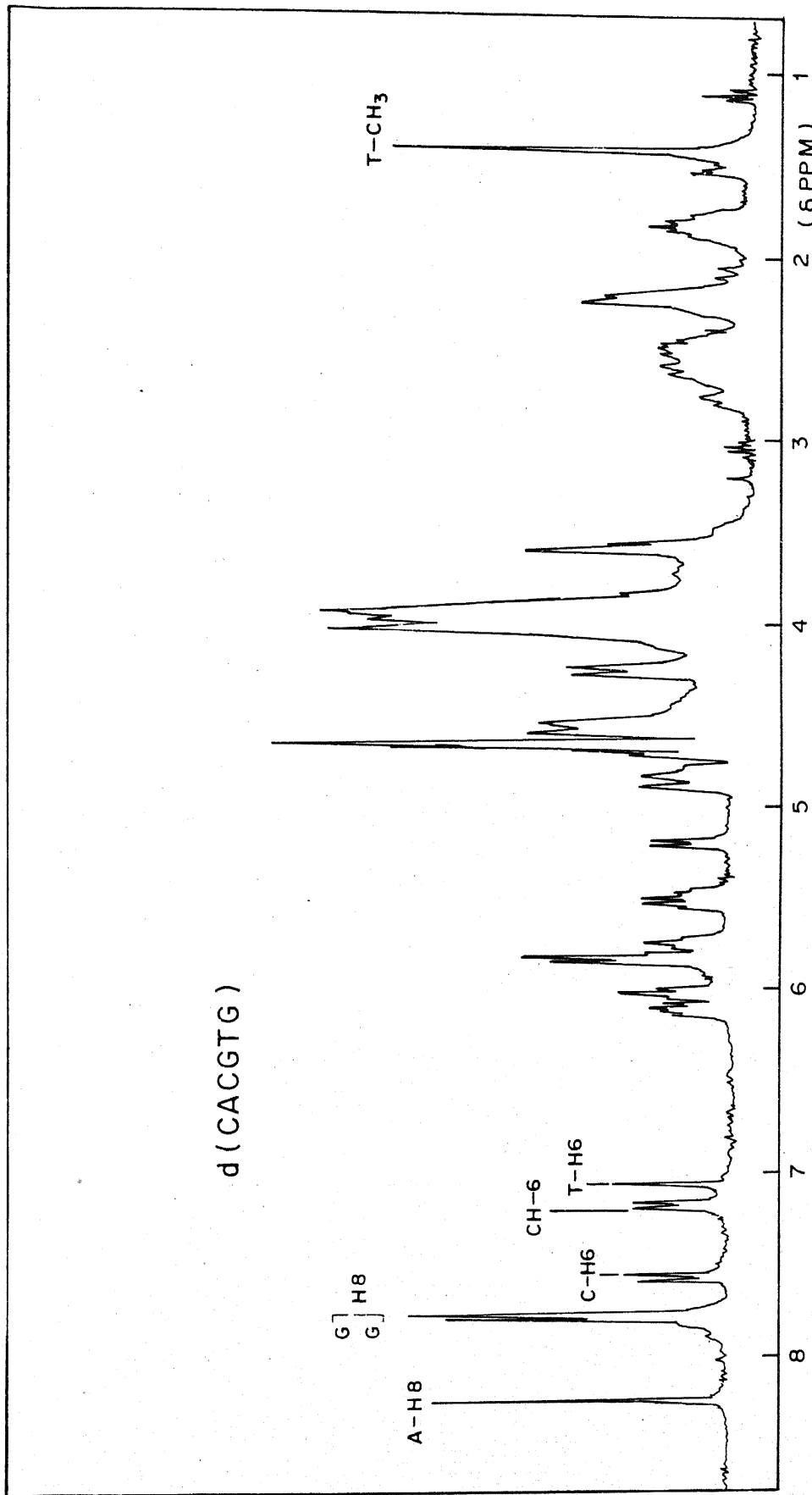
Figure 2. 270 MHz proton NMR spectrum of d(CACGTG) at 25°C. Sample concentration: 2 mg in 0·5 ml (~ 4 mm) of phosphate buffer (0·02 M), pH 7·5. Solution was lyophilized (3 times) and redissolved in 0·5 ml D_2O.

4. Experimental

The N,O-protected nucleotides and nucleosides were synthesised according to procedures reported previously (Rajendrakumar *et al* 1985). Methylimidazole and phenyl dichlorophosphate were procured from Fluka, Switzerland. Pyridine GR (E Merck, India) was refluxed and distilled over ninhydrin followed by distillation over calcium oxide. Acetonitrile and dichloromethane (E Merck, India) were distilled over phosphorus pentoxide before use. 2,4,6-Triisopropylbenzenesulphonyl chloride (Sigma, USA) was recrystallised from hexane. 1-(Mesitylene-2-sulphonyl)-3-nitro-1,2,4-triazole (MSNT) was prepared according to Gait *et al* (1984). All column chromatographic purifications were done over Merck Kieselgel 60 (Art. 9385) and monitored by t.l.c. over pre-coated fluorescent silica gel t.l.c. plates (Merck Art. 5554). The spots were visualised by UV lamp and by spraying with 60% perchloric acid-ethanol (3:2) for trityl derivatives. Compounds without the trityl group gave dark spots on spraying with the above reagent followed by heating. The 3'-terminal block **1** was prepared according to Denny *et al* (1982) and phenyl dihydrogen phosphate was synthesised by hydrolysis of phenyl dichlorophosphate (Owen *et al* 1974). The general protocol of the chemical synthesis is illustrated by the following procedure for the chemical synthesis of d(CACGTG).

Step (1): The phosphodiester block DMT-T-p (**2**, 140 mg, 0·18 mmole) and the 3'-terminal nucleoside (**1**, HO-dG-Bz, 72 mg, 0·15 mmole) were dried by coevaporation (2 times) and dissolved in with dry acetonitrile (1 ml/0·1 mmole of **1**, 1·5 ml) under anhydrous conditions and treated with the condensing agent (0·45 mmole, TPSCl 140 mg or MSNT 133 mg) and 1-methylimidazole (0·9 mmole, 72 μl). The reaction mixture was stirred at room temperature. The reaction, followed by t.l.c., was essentially complete within 10 min. Excess reagents were destroyed by treatment with aq. $NaHCO_3$ and the product was extracted into chloroform (3 × 20 ml). The chloroform layer was washed with water (10 ml) and the dried (Na_2SO_4) organic layer was concentrated under reduced pressure to yield a colourless foamy material of the protected dinucleotide **3**. The total time to complete step 1 was 30 min. The product was taken to step 2 without any characterization.

Step (2): The material from step 1 was dissolved in chloroform-methanol (95:5, v/v, 6 ml), cooled at 10°C and treated with solid phenyl dihydrogen phosphate **4** (260 mg, 1·5 mmole). An instantaneous orange-red colour was produced. After 5–10 min (check t.l.c.) the reaction mixture was diluted with chloroform (20 ml) and washed with aq. $NaHCO_3$. The dried organic layer on evaporation gave a foamy gum which was purified by step 3 to give **5**.

Step (3): The material from step 2 was dissolved in chloroform (1·5 ml) and loaded uniformly over a short column (2 cm i.d.) of Merck silica gel (8 g). The column was washed with chloroform (~ 20 ml), when a pale yellow band separated and eluted out. This was then followed by elution with anhydrous tetrahydrofuran-pyridine (3:1, v/v, 40 ml) when the desired product, **5** was obtained in the eluant. This was recovered after evaporation under reduced pressure and repeatedly coevaporated with dichloromethane. The product **5** (T_pG^{Bz}-Bz, 120 mg, 90% yield) was dried over P_2O_5-KOH in a vacuum desiccator and used for initiating the next cycle.

Table 2 Reaction conditions* for synthetic cycles.

3'-component (mg, mmole)	5'-component (mg, mmole)	Condensing agent (mg, mmole)	N-Methyl-imidazole (μl, mmole)	Product (mg, yield)
d(T$_p$G^{Bz}-Bz) (115, 0·13)	DMT-dG-$^{Bz}_p$ (140, 0·15)	TPSCl (117, 0·39)	62; 0·78	d(G$^{Bz}_p$T$_p$G^{Bz}-Bz) (144, 80%)
d(G$^{Bz}_p$T$_p$G^{Bz}-Bz) (140, 0·1)	DMT-dC$^{Bz}_p$ (130, 0·14)	TPSCl (90, 0·3)	48; 0·6	d(C$^{Bz}_p$G$^{Bz}_p$T$_p$G^{Bz}-Bz) (140, 70%)
d(C$^{Bz}_p$G$^{Bz}_p$T$_p$G^{Bz}-Bz) (140, 0·08)	DMT-dA$^{Bz}_p$ (106, 0·11)	MSNT (71, 0·24)	38; 0·48	d(A$^{Bz}_p$C$^{Bz}_p$G$_p$T$^{Bz}_p$G^{Bz}-Bz) (126, 70%)
d(A$^{Bz}_p$C$^{Bz}_p$G$^{Bz}_p$T$_p$G^{Bz}-Bz) (120, 0·05)	DMT-dC$^{Bz}_p$ (72, 0·08)	MSNT (45, 0·15)	25; 0·3	d(C$^{Bz}_p$A$^{Bz}_p$C$^{Bz}_p$G$^{Bz}_p$T$_p$G^{Bz}-Bz) (144, 70%)

The 3'-component, 5'-component, condensing agent and N-methylimidazole in the given amounts were dissolved in 1·5 ml of acetonitrile or pyridine and stirred at room temperature for 15 minutes before work-up.

Four more cycles were carried out similarly according to conditions shown in table 2. After the final cycle, the product obtained was subjected to the following sequence to yield the completely deblocked d(CACGTG).

d(AcO-C$^{Bz}_p$A$^{Bz}_p$C$^{Bz}_p$G$^{Bz}_p$T$_p$G^{Bz}-Bz):
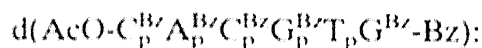
d(C$^{Bz}_p$A$^{Bz}_p$C$^{Bz}_p$G$^{Bz}_p$T$_p$G^{Bz}-Bz) (140 mg) was treated with acetic anhydride (0·6 ml) and pyridine (1·5 ml) at room temperature for 2 h. The reaction mixture was poured into ice-water and stirred for 10 min. It was then extracted into chloroform (25 ml), washed with aq. NaHCO$_3$ and dried over Na$_2$SO$_4$. The organic layer on concentration gave a foamy acetate product (130 mg, 92%).

d(CACGTG): The acetate obtained above was dissolved in dioxane-water (1 : 1, v/v, 10 ml) and treated with *syn*-2-nitrobenzaldoxime (400 mg) followed by tetramethylguanidine (0·27 ml). The reaction mixture was kept at room temperature for 14 h and then heated at 60°C for 3 h. It was lyophilized and then treated with concentrated ammonia (30 ml) in a sealed flask for 20 h at 60°C. The ammonia was evaporated and the product was passed through a Sephadex G-15 column (bed volume 120 ml) and eluted with 20% methanol-water. The eluted fractions were monitored by UV detector and the major peak eluting in the void volume was lyophilized. The residue was then purified over FPLC (for conditions see legend to figure 1) to obtain d(CACGTG) (70 mg, 30% overall yield).

5. Conclusions

In this paper, we have demonstrated an optimized protocol for a rapid, large-scale chemical synthesis of short oligodeoxynucleotides by solution-phase phosphotriester chemistry. In experienced hands, the protocol consumes only 150 min per cycle and can be conveniently scaled up to yield 6–8-mers in hundred-milligram

quantities. As indicated by ^{1}H NMR, the purity of the products after a preparative FPLC is extremely good and is acceptable for spectroscopic and crystallographic work.

Acknowledgements

References

Chaudhari B, Reese C B and Weclawek K 1984 *Tetrahedron Lett.* **25** 4037
Denny W A, Leupin W and Kearn D R 1982 *Helv. Chim. Acta* **65** 2372
Effimov V A, Reverdatto S V and Chakmakhcheva O G 1982 *Nucleic Acids Res.* **10** 6675
Gait M J (ed.) 1984 *Oligonucleotide synthesis, a practical approach* (Oxford: IRL Press)
Ganesh K N 1985 *Curr. Sci.* **54** 103
Minganti C, Ganesh K N, Sproat B S and Gait M J 1985 *Anal. Biochem.* **147** 63
Owen G R, Ranson C J, Reese C B, Van Boom J H and Herschied J D H 1974 *Synthesis* 704
Rajendrakumar G V, Sivakamisundari N and Ganesh K N 1985 *Proc. Indian Acad. Sci. (Chem. Sci.)* **95** 357
Sonveaux E 1986 *Bioorg. Chem.* 274, and references cited therein