Volume 13 Number 2 1985 Nucleic Acids Research

Methylation of thymine residues during oligonucleotide synthesis

X.Gao, B.L.Gaffney, M.Senior, R.R.Riddle and R.A.Jones

Department of Chemistry, Rutgers, The State University of New Jersey, New Brunswick, NJ 08903,USA

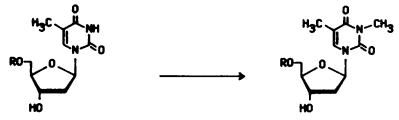
Received 3 December 1984; Accepted 3 January 1985

ABSTRACT

Thymine residues in an oligodeoxyribonucleotide are subject to methylation at N^3 by the internucleotide methyl phosphotriester linkages. This alkylation occurs most rapidly in the presence of a strong base such as DBU, but also takes place, at a much slower rate, during oligonucleotide synthesis.

INTRODUCTION

The phosphoramidite approach to oligonucleotide synthesis is a rapid, efficient method with few apparent side reactions.[1] There is, however, no doubt that some side reactions do take place. Hplc analysis of the crude material obtained at the completion of a synthesis always shows a handful of small peaks centered around the product peak. These are not failure sequences. The failure sequences are removed by hplc before detritylation. They must represent, therefore, sequences which are incompletely deprotected, partially depurinated, or otherwise modified. Since even extensive treatment with aqueous ammonia will not eliminate these contaminants, incomplete deprotection cannot be the explanation. Depurination is a well known problem, and is a likely contributor. Base modification may be another source. Modified bases, even if present in only quite small amounts, could have serious consequences for certain types of physical and biochemical experiments. It has been shown that in phosphate triester synthesis reaction of the appropriate O^6 or O^4 amide oxygen with condensing agents, followed by ammonolysis, is capable of converting deoxyguanosine to 2-amino deoxyadenosine, thymidine to 5-methyl deoxycytidine, and uridine to cytidine.[2-4] Moreover, since base modification ordinarily gives rise to sequences which do not differ in either length or charge from the desired product, the modified molecules are not likely to be separated from the product by the commonly used gel or ion-exchange techniques.[5] The ability of even reversed-phase hplc to resolve such sequences is limited. We have, therefore, begun to



1 2

Figure 1. a, R = H; b, R = 4,4'-dimethoxytrityl (DMT).

re-examine synthetic methods and procedures to attempt to identify the kinds of modifications, if any, which may be produced, and their source, so that they may be avoided.

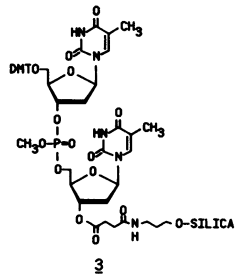
RESULTS

Reaction of Trimethyl Phosphate with Thymidine

The high yield alkylation of thymidine by trimethyl phosphate (TMP) in basic aqueous solution has been reported.[6] As a model for the reaction of the internucleotide methyl triester we have examined the reaction in organic solvents. We have found that, in the presence of either 1,8-diazabicyclo [5.4.0]undec-7-ene (DBU) or triethylamine (TEA), thymidine is efficiently alkylated exclusively at N^3. Both 3-methyl thymidine (2a) and its 5'-dimethoxytrityl (DMT) derivative (2b) have been obtained in excellent yield (Figure 1). While both DBU and triethylamine were effective, DBU gave significantly faster alkylation (Table 1). For example, a solution of 1b in pyridine containing 10% trimethylphosphate showed complete conversion to 2b in a few hours in the presence of 10% DBU, while 10% TEA gave only about 13% formation of 2b after two days. This is consistent with the pH dependence reported for the reaction in water.[6] Higher concentrations of trimethylphosphate, which may be a better model for the intrastrand reactions which can take place with

Table 1. The Dependence of the Rate of Alkylation of dT (1a) and d[DMT-T] (1b) on Base Strength and TMP Concentration.

Reactant	Reaction System	TMP (%)	Time	Alkylation (%)
1b	10% DBU/pyr	10	3 hr	95
1b	10% TEA/pyr	10	2 days	13
1b	10% TEA/TMP	90	2 days	90
1a	10% TEA/TMP	90	2¼ hr	7
1a	10% TEA/TMP	90	33 hr	46



3

Figure 2. d[DMT-TpT]-silica.

an oligonucleotide, dramatically increase the reaction rate. Thus 1b was 90% methylated when trimethylphosphate was used as the solvent along with 10% TEA.

Treatment of the Dinucleoside Methyl Phosphate d[DMT-TpT]-silica (3) with DBU

The dinucleoside methyl phosphate d[DMT-TpT]-silica (3, Figure 2) was treated with 10% DBU in pyridine for 6.5 hours followed by the normal triethylamine/thiophenol deprotection, cleaved from the support with aqueous ammonia, and examined by hplc. Three peaks were observed (Figure 3). These compounds were separated by hplc, detritylated, and then degraded using venom phosphodiesterase and alkaline phosphatase. Under these conditions compound

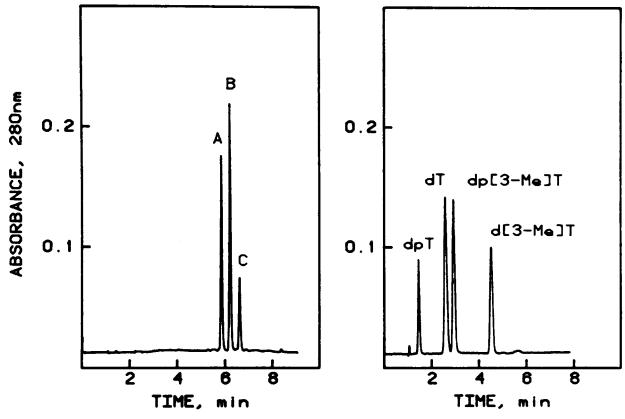


Figure 3. Analytical hplc of the mixture obtained by treatment of d[DMT-TpT]-silica (3) with DBU for 6.5 hours (left), using a gradient of 20 to 50% CH_3CN:0.1 M TEAA in 5 min at 4 mL/min, and (right) the mixture of deoxynucleosides and deoxynucleotides obtained upon detritylation and degradation with venom phosphodiesterase of the components present in peak B of the left panel, using a gradient of 6 to 20% CH_3CN:0.1 M TEAA in 5 min at 4 mL/min.

A gave only dT, while compound B gave equal amounts of dT and d(3-methyl)T, and compound C gave only d(3-methyl)T. Thus A is unmodified, B is monomethylated and C is dimethylated. Since there is only one methyl phosphate per dimer, the presence of a dimethylated compound indicates that interstrand alkylation occurred. A sample of B was also degraded using venom phosphodiesterase alone. Under these conditions the 3' nucleoside is obtained as its 5' phosphate, making it possible to distinguish between the 5' and 3' ends of the dimer. B now gave a mixture of four compounds (Figure 3); dpT, dT, dp(3-Me)T, and d(3-Me)T. From the relative quantities of these four compounds peak B is seen to be a 60:40 mixture of, respectively, the 3' and 5' monomethylated dimers. This mixture is not resolvable even after detritylation. The presence of 40% of the 5' methylated compound is somewhat surprising in that, from examination of models, alkylation of the 5' nucleoside only appears to be possible via an interstrand reaction. Thus not only did the interstrand reaction occur, as shown by the presence of dimethylated dimer, but it also was reasonably competitive with the intrastrand reaction by which the 3' nucleoside may have reacted. This result suggests that at a loading of 77 μmol/g the surface of the Fractosil 500 is densely populated.

Treatment of d[DMT-TpTpTpTpTpT]-silica with DBU

The dimer 3 has the lowest possible ratio of methylphosphate to thymidine (1:2). By elongating 3 to the corresponding hexamer, d[DMT-TpTpTpTpTpT]-silica (4), a molecule with a nearly equal ratio (5:6) of methylphosphate to thymidine is obtained. Portions of the hexamer were then treated with a 10% solution of DBU in pyridine, for times from 0.5 to 46 hours. The samples were washed free of DBU, treated with thiophenol/triethylamine and cleaved from the silica with aqueous ammonia. Analysis of each sample by hplc showed heterogeneity, but was not able to resolve the mixture into separate peaks. The broad product peak was isolated in each case without attempted fractionation through the peak itself. Detritylation with 80% acetic acid then gave mixtures like that shown in Figure 4 from the 0.5 hour sample. Each of the components of the mixture was then isolated and characterized by enzymatic degradation. A total of seven components were obtained, corresponding to 0, 1, 2, 3, 4, 5, and 6 modified thymidine nucleosides per strand. In general, these could only be resolved based on the number, not the location, of the 3-methyl thymidines. The peaks are somewhat broader than normal because of this underlying heterogeneity. The dT and d(3-Me)T obtained by enzymatic degradation of peak 2, corresponding to two (3-Me)T residues, are also shown in Figure 4. The retention times of T_6 and the 1-6 methylated compounds are

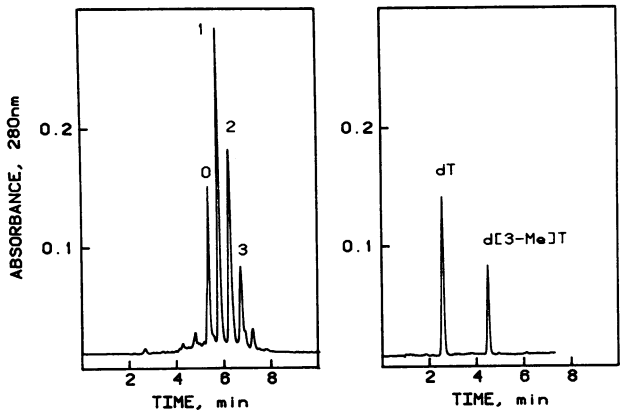


Figure 4. Analytical hplc of the mixture obtained by treatment of d[DMT-TpTpTpTpTpT]-silica (4) with DBU for one-half hour, using the 6 to 20% gradient described in Figure 3, where 0, 1, 2, and 3 indicate the number of 3-methyl thymine residues present in those components (left), and (right) the mixture of deoxynucleosides obtained upon degradation of the component(s) present in peak 2 of the left panel, using the 6 to 20% gradient described in Figure 3.

listed in Table 2, along with the experimentally determined percentages of d(3-Me)T. The relative quantities of these methylated compounds produced at 0.5, 6.5 and 46 hours are shown in Figure 5, and the total percentage of methylation at each time is shown in Figure 6.

Analysis for 3-methyl Thymidine Formed During Oligonucleotide Synthesis and Deprotection

The experiments outlined above indicate that thymidine alkylation is at

Table 2. Retention of $d(T)_6$ and Methylated $d(T)_6$.

Number of 3-methyl thymine residues	Retention Time (min)	% d(3-Me)T (obs)[1]
0	5.4	—
1	5.8	17.3
2	6.3	33.5
3	6.7	50.1
4	7.2	66.1
5	7.7	79.1
6	9.0	94.5

[1]Determined after enzymatic degradation.

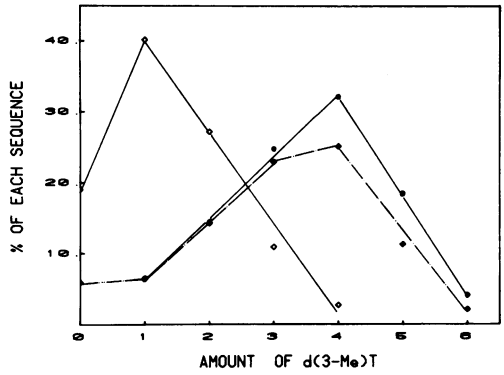


Figure 5. The relative quantities of the 0 to 6 methylated $d(T)_6$ compounds formed by exposure of 4 to DBU for 0.5 (◇), 6.5 (◆), and 46 (●) hours, determined by hplc after deprotection.

least a potential side reaction of oligonucleotide synthesis. We next looked for d(3-Me)T in several oligonucleotides which were being purified while this work was in progress. We routinely purify crude oligonucleotides by hplc on a 7.8 x 30 cm Bondapak column both before and after detritylation.[4,7] The first separation serves to remove failure sequences and other shorter retention debris, so that these compounds do not interfere with the second, and final, purification. From the results presented above, methylated sequences, if present, should be among the longer retention impurities which are removed during the final hplc purification, assuming that they are resolved at all.

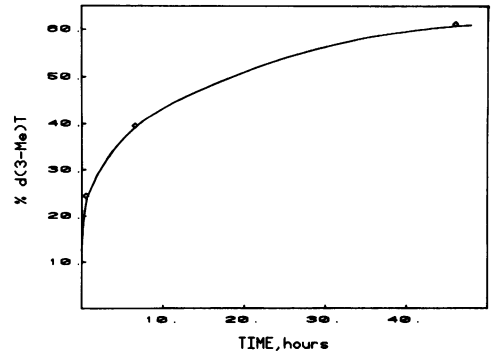


Figure 6. The overall percentage of methylation of 4 as a function of time of exposure to DBU, determined by hplc after deprotection and enzymatic degradation.

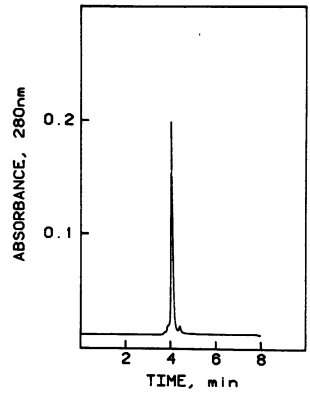


Figure 7. Analytical hplc of crude d[GCTTTCG] prior to final purification, using the 6 to 20% gradient described in Figure 3.

Consider the analytical hplc of d[GCTTTCG] as it was before the second hplc purification (Figure 7). There is a well resolved impurity evident just after the main peak. Both compounds were isolated and enzymatically degraded to give the mixtures of deoxynucleosides shown in Figure 8. One d(3-Me)T is clearly present in the longer retention impurity.

The analytical hplc of d[GCGCGCTT], again as it was prior to final purification, is shown in Figure 9. The main product and the longer retention

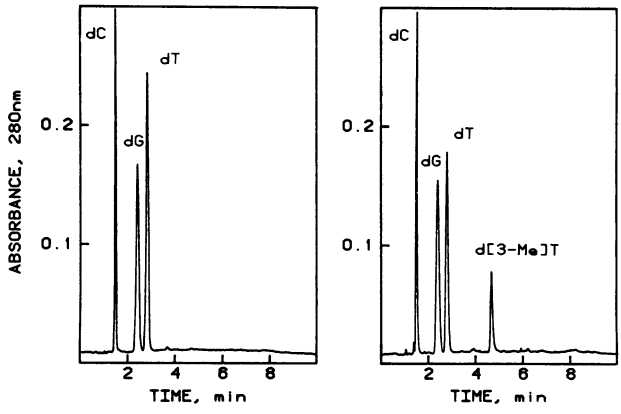


Figure 8. The mixture of deoxynucleosides obtained upon degradation of pure d[GCTTTCG] (left), and (right) from similar degradation of the longer retention impurity (Figure 7), after isolation of each by hplc, using the 6 to 20% gradient described in Figure 3.

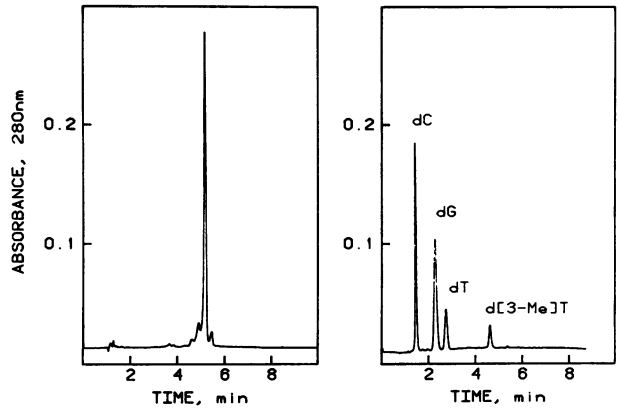


Figure 9. Analytical hplc of crude d[GCGCGCTT] prior to final purification (left), and (right) the mixture of deoxynucleosides obtained upon degradation of the longer retention impurity (Figure 10, peak B) after isolation by hplc. The gradient used in each case is the 6 to 20% gradient described in Figure 3.

impurity were both isolated and degraded. The mixture obtained upon degradation of the isolated impurity, also shown in Figure 9, once again contains d(3-Me)T. The actual preparative purification of a portion of the crude product is shown in Figure 10. A total of 116 OD_{260} of material was isolated;

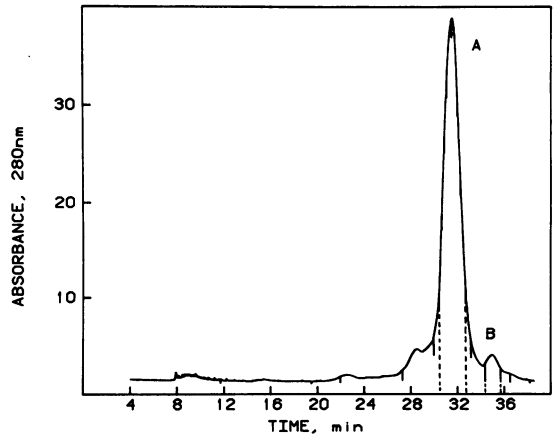


Figure 10. Preparative hplc of a portion of crude d[GCGCGCTT] using a gradient of 5 to 20% CH_3CN:0.1 M TEAA in 30 min at 2 mL/min. The pure product (A) and the longer retention impurity (B) were obtained from the combined fractions between the dashed lines as indicated.

102 OD_{260} of pure d[GCGCGCTT] (A) and 3 OD_{260} of the impurity (B).

There was one sequence in which we could not find evidence of methylation: d[TTCGCGCG]. There was no longer retention peak in the analytical hplc and no d(3-Me)T was found by enzymatic degradation of fractions taken throughout the main peak. Presumably since both thymidines in this sequence are at the 5' end of the molecule they did not have sufficient exposure to the methyl triesters for detectable alkylation to occur.

DISCUSSION

The experiments described above demonstrate that thymine residues in an oligonucleotide are subject to methylation by the internucleotide methyl triesters. The extent of alkylation can be minimized by avoiding strong bases, but some alkylation is probably unavoidable so long as the methyl group is used for triester protection. Furthermore, although we have been able to isolate pure compounds by using reversed-phase hplc, the use of ion-exchange hplc or gel electrophoresis is unlikely to resolve these methylated sequences since they do not differ in either length or charge from the unmodified compound.[5] Even the reversed-phase hplc purification must be limited to sequences of about 10 to 15 residues depending on the exact sequence. Moreover, it is these longer sequences that are likely to have the greatest extent of alkylation, because of their longer exposure to the methyl triesters.

It should be noted that the sequences described above were synthesized manually on scales of 500 to 1000 mg of functionalized silica. On these scales the condensation cycle time we use is longer than that used by small scale automated systems. Thus the extent of methylation we find is likely to be correspondingly larger. For most purposes, the small amounts of 3-methyl thymidine produced may have no detectable effect. However, this side reaction should be kept in mind where trace contamination could be a problem or if the synthesis of an unusually long sequence is contemplated. In such cases an alternative phosphate protecting group should be used.

EXPERIMENTAL

^{1}H and ^{13}C nmr spectra were obtained on a Varian CFT-20 nmr spectrometer. UV spectra were taken from a Varian 118C uv spectrometer. Melting points were determined on a Thomas-Hoover apparatus and are uncorrected. Thin layer chromatography (TLC) was carried out on Eastman Kodak chromatogram sheet 13181. Reagents and procedures not specified below were as reported previously.[4,7]

Preparation of 3-methyl Thymidine

To a solution of DMT-T (2 mmol, 1.1 g) in pyridine (27 mL), DBU (20.1 mmol, 3 mL) and TMP (26 mmol, 3 mL) were added. The reaction mixture was stirred at room temperature for 3 h. The solution was then partitioned between H_2O and CH_2Cl_2 and the aqueous layer was extracted with CH_2Cl_2. The combined organic layers were concentrated under vacuum and the residue was purified on a silica gel column using CH_3OH:CH_2Cl_2 (1.5:98.5) as eluent. The combined product fractions were concentrated to give 2b as an amorphous solid (0.982 g, 1.76 mmol, 88%). UV_{max} (CH_3OH) 267 nm; UV_{sh} 280 nm; UV_{min} 257 nm.

2b (0.5 g, 0.9 mmol) was dissolved in 22 mL of a solution of trichloroacetic acid (2%, 2.7 mmol) in CH_2Cl_2. After 5 min, excess pyridine was added and the solution was concentrated under vacuum. The mixture was fractionated on a Dowex AG 1-X2 column using H_2O as eluent. The combined product fractions were concentrated to give 2a as a white solid (0.2 g, 0.8 mmol, 90%). A sample recrystallized from acetone and ethyl ether had mp 128-130°C (mp of a commercial sample 127-129°C). UV_{max} (CH_3OH) 266 nm ($\varepsilon 7.86 \times 10^3$), UV_{min} 235 nm ($\varepsilon 2.04 \times 10^3$). ^{13}C NMR (d_7-DMF, D_2O) δ 163.4 (C_4), 150.2 (C_2), 134.0 (C_6), 108.4 (C_5), 85.9 ($C_{4'}$), 84.6 ($C_{1'}$), 69.3 ($C_{3'}$), 60.1 ($C_{5'}$), 38.2 ($C_{2'}$), 26.5 (3-Me), 11.2 (5-Me). ^1H NMR consistent with that reported in the literature.[6] Anal. Cald. for $C_{11}H_{16}N_2O_5$: C, 51.55; H, 6.29; N, 10.93. Found: C, 51.43; H, 6.45; N, 10.79.

Reaction of Trimethyl Phosphate with 1a or 1b (Table 1)

To 1a (~10 mg) or 1b (~15 mg) was added a 1 mL portion of the solution indicated in Table 1. Reactions were allowed to proceed at room temperature and were monitored by hplc on an analytical column (Radial-Pak μ Bondapak C_{18} Cartridge). A gradient of 6-20% acetonitrile:0.1 M triethylammonium acetate (TEAA) in 5 min at 4 mL/min was used for reactions involving 1a. A gradient of 55-75% was used for reactions of 1b. Under these conditions the retention times for 1a and 2a were 3.0 and 4.9 min, respectively, and for 1b and 2b, 3.5 min and 4.4 min. The relative peak heights detected at 280 nm were used to determine the extent of the reaction.

Treatment of d[DMT-TpT]-silica (3) or d[DMT-TpTpTpTpTpT]-silica (4) with DBU

A 100-230 mg sample of the silica-bound oligonucleotide was treated with a 10% solution of DBU in pyridine for the times indicated in the text. The silica was washed with CH_3CN and deprotection carried out as indicated below. A sample of 3-methyl thymidine obtained from enzymatic degradation was found to give an identical uv spectrum to that of 3-methyl thymidine obtained by

chemical synthesis.

General Procedure for Oligonucleotide Synthesis

Detritylation of the silica-bound material was effected by three to four 30 second treatments with chilled (-20°C) 2% trichloroacetic acid in CH_2Cl_2, alternating with washes of chilled CH_2Cl_2, and followed by quenching with pyridine. The silica was further washed with three portions of dry CH_3CN and dried by passing N_2 through it.

To the dried silica-bound oligonucleotide was added a solution of tetrazole in CH_3CN (3 equiv relative to the amount of phosphoramidite to be added), followed by a solution of the appropriate phosphoramidite (10-20 equiv) in CH_3CN. All operations were performed by syringe, under a positive pressure of N_2, in vessels sealed with septa. The reaction was allowed to proceed for twenty minutes. The excess reagents were then removed by filtration and the support was washed with three portions of CH_3CN and one portion of lutidine:THF:H_2O (2:2:1).

Oxidation was carried out by treatment for one minute with excess fresh 0.2 M I_2 in the same lutidine:THF:water mixture. After washing with CH_3CN, unreacted sites were capped by treatment with 10% acetic anhydride in dry pyridine for five minutes. The support was then washed with three portions of CH_3CN and three portions of CH_2Cl_2.

Coupling efficiencies, as determined by spectrophotometric analysis of the filtrates obtained upon detritylation, were generally about 95%.

General Procedures for Deprotection of Oligonucleotides

The silica-bound oligonucleotide was treated with thiophenol:triethylamine:dioxane (1:1:2) for 1 h. The silica was then washed with CH_3CN and CH_3OH successively and treated with conc. aqueous ammonia for 48 h. The mixture was filtered and the silica was washed with H_2O and CH_3OH. The combined aqueous solution was concentrated under vacuum and the residue was applied to a Bio Gel P-2 column using NH_3:EtOH:H_2O (0.1:20:80) as eluent. The product fractions were combined, concentrated and purified on a semi-preparative C-18 Bondapak column using a gradient of 20-50% CH_3CN:0.1 M TEAA in 25 min at a flow rate of 2 mL/min. Product fractions were concentrated and detritylated by treatment with 80% acetic acid for 20 min. The solution was concentrated, water was added, and the mixture was extracted with ether three times. The aqueous layer was then concentrated and the residue was purified on the semi-preparative column as described above using a gradient of 6-20% CH_3CN:TEAA in 30 min at 2 mL/min (see Figure 8).

General Procedure for Enzymatic Degradation

Using Venom Phosphodiesterase and Alkaline Phosphatase. To a solution (1-10 OD_{260}) of the oligonucleotide in 100μL of 0.1 M TEAA, pH 10, was added 10μL of a solution of snake venom phosphodiesterase (ca. 0.01 unit) and 10μL of a solution of alkaline phosphatase (ca. 1 unit). The sample was stored at 45°C. Degradation is generally complete (hplc) within one to four hours.

Using Venom Phosphodiesterase Alone. To a solution (1-10 OD_{260}) of the oligonucleotide in 100μL of 0.1 M KH_2PO_4, pH 10, was added 10μL of a solution of venom phosphodiesterase (ca. 0.01 unit). The sample was stored at 45°C. Degradation was usually complete (by hplc) within one to four hours.

ACKNOWLEDGEMENTS

This work was supported by grants from the Biomedical Research Support Grant, the American Cancer Society (CH248) and the National Institutes of Health (GM31483). The authors thank Mr. James Snyder of this department for use of his data plotting program, Snyder-Writer®.

REFERENCES

1. McBride, L.J. and Caruthers, M.H. (1983) Tetrahedron Letters 24, 2953-2956.
2. Reese, C.B. and Ubasawa, A. (1980) Tetrahedron Letters 21, 2265-2268.
3. Sung, W.L. (1981) J. Chem. Soc. Chem. Commun., 1089.
4. Gaffney, B.L., Marky, L.A., and Jones, R.A. (1984) Tetrahedron 40, 3-13.
5. Haupt, W. and Pingovel, A. (1983) J. Chromatography 260, 419-427.
6. Tanabe, T., Yamauchi, K., and Kinoshita, M. (1979) Bull. Chem. Soc. Japan 52, 204-207.
7. Gaffney, B.L., Marky, L.A. and Jones, R.A. (1984) Biochemistry 23, 0000-0000.