References

Bauer, W., & Vinograd, J. (1970) *J. Mol. Biol. 47,* 419-435.
Ben-Hur, E. (1975) *Isr. J. Chem. 13,* 221-228.
Bresloff, J. L., & Crothers. D. M. (1975) *J. Mol. Biol. 95,* 103-123.
Brodie, S., Giron, J., & Latt, S. A. (1975) *Nature (London) 253,* 470-471.
Burton, D., Hyde, J. E., & Walker, I. O. (1975) *FEBS Lett. 55,* 77-80.
Clewell, D. B., & Helinski, D. R. (1972) *J. Bacteriol. 110,* 1135-1146.
Cole, R. S. (1970) *Biochim. Biophys. Acta 217,* 30-39.
Cole, R. S., Levitan, D., & Sinden, R. R. (1976) *J. Mol. Biol. 103,* 39-59.
Compton, J. L., Bellard, M., & Chambon, P. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73,* 4382-4386.
Crothers, D. M. (1968) *Biopolymers 6,* 575-584.
Dall-Acqua, F., Marcinani, S., Ciavatta, L., & Rodighiero, G. (1971) *Z. Naturforsch. B 26,* 561-569.
Dove, W., & Davidson, N. (1962) *J. Mol. Biol. 5,* 467-478.
Gellert, M., Smith, C. E., Neville, D., & Felsenfeld, G. (1965) *J. Mol. Biol. 11,* 445-457.
Hanson, C. V., Shen, C.-K. J., & Hearst, J. E. (1976) *Science 193,* 62-64.
Henson, P., & Walker, I. O. (1970) *Eur. J. Biochem. 14,* 345-350.
Honda, B. M., Baillie, D. L., & Candido, E. P. M. (1974) *FEBS Lett. 48,* 156-159.
Isaacs, S. T., Shen, C.-K. J., Hearst, J. E., & Rapoport, H. (1977) *Biochemistry 16,* 1058-1064.
Lawrence, J.-J., & Daune, M. (1976) *Biochemistry 15,* 3301-3307.
Lepecq, J.-B., & Paoletti, C. (1967) *J. Mol. Biol. 27,* 87-106.
Li, H. J., & Crothers, D. M. (1969) *J. Mol. Biol. 39,* 461-477.
Modrich, P., & Zabel, D. (1976) *J. Biol. Chem. 251,* 5866-5874.
Musajo, L., & Rodighiero, G. (1970) *Photochem. Photobiol. 11,* 27-35.
Noll, M., Thomas, J. O., & Kornberg, R. D. (1975) *Science 187,* 1203-1206.
Ohlenbusch, H. H., Olivera, D. M., Tuan, D. Y. H., & Davidson, N. (1967) *J. Mol. Biol. 25,* 299-315.
Parrish, J. A., Fitzpatrick, T. B., Tanenbaum, L., & Pathak, M. A. (1974) *N. Engl. J. Med. 291,* 206-210.
Peacocke, A. R., & Skerrett, J. N. H. (1956) *Trans. Faraday Soc. 52,* 261-279.
Scatchard, G. (1949) *Ann. N.Y. Acad. Sci. 51,* 660-672.
Scott, B. R., Pathak, M. A., & Mohn, G. R. (1976) *Mutat. Res. 39,* 29-74.
Shen, C.-K. J., & Hearst, J. E. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73,* 2649-2653.
Shen, C.-K. J., & Hearst, J. E. (1977) *J. Mol. Biol. 112,* 495-507.
Waring, M. J. (1965) *J. Mol. Biol. 13,* 269-282.
Wiesehahn, G. P., Hyde, J. E., & Hearst, J. E. (1977) *Biochemistry 16,* 925-932.

# High-Pressure Liquid Chromatography in Polynucleotide Synthesis†

Hans-Joachim Fritz,‡ Ramamoorthy Belagaje, Eugene L. Brown,§ Rebecca H. Fritz, Roger A. Jones,¶ Robert G. Lees,‖ and H. G. Khorana*

ABSTRACT: Reverse phase high-pressure liquid chromatography (HPLC) using columns containing microparticulate materials with bonded octadecyl groups has been developed as a rapid and efficient method for the separation of nucleosides, nucleotides, and, in particular, of protected oligonucleotides which are standard intermediates in the stepwise synthesis of deoxyribopolynucleotides. Reported are extensive studies of the influence of the different purine and pyrimidine bases, of protecting groups, of the phosphate groups, and of the chain lengths of oligonucleotides on their retention on such columns. Further, the application of HPLC in the stepwise synthesis of an oligonucleotide, d(G-G-A-A-G-C-T-T-A-A-C), has been described. The methods, which are herein described, lend themselves to separations on a preparative scale and effect a marked reduction (up to 50%) in the time required for the synthesis of oligonucleotides.

The current methodology for the total synthesis of biologically specific DNA involves (a) the chemical synthesis of short deoxyribopolynucleotides corresponding to the entire two strands and (b) the end-to-end joining of deoxyribopolynucleotides following enzymatic phosphorylation of the 5′-hydroxyl end groups by the use of the polynucleotide ligase (Khorana et al., 1976). Chemical synthesis has so far remained

† From the Departments of Biology and Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139. *Received September 12, 1977.* This investigation was supported by the National Cancer Institute of the National Institutes of Health, U.S. Public Health Service (Grant No. CA11981), and the National Science Foundation (Grant No. PCM73-06757). This is paper 146 in the series Studies on Polynucleotides. The preceding paper is by Sekiya et al., 1976.

‡ Recipient of Postdoctoral Fellowships from NATO (1974-1975) and Deutsche Forschungsgemeinschaft (1976). Present address: Institute für Genetik der Universität zu Köhn, 5 Köhn 41, Weyertal 121, Federal Republic of Germany.

§ Recipient of a Postdoctoral Fellowship (1974-1976) of the National Institutes of Health (Fellowship No. CA01599).

¶ Recipient of a Postdoctoral Fellowship (1975-1976) of the National Institutes of Health (Fellowship No. GM05388). Present address: Department of Chemistry, Douglass College, Rutgers University, New Brunswick, N.J. 08903.

‖ Present address: Department of Chemistry, Fordham University, New York, N.Y. 10458.

0006-2960/78/0417-1257$01.00/0 

the slowest phase in the total synthesis, and developments which reduce the time and effort involved at this stage would greatly facilitate the synthesis of sequence-specific double-stranded DNA. In the bulk of the published work, the time-consuming anion exchange chromatography using DEAE-cellulose, both at the level of protected and unprotected oligonucleotides, has been extensively used to purify and characterize the products at every stage. On the whole, separation and characterization of the intermediates and products, rather than the synthetic procedures, have accounted for the greater part of the total time expended in chemical synthesis. In recent years, increasing attention has, therefore, been given to developing more rapid methods for the separation of reaction mixtures, and one general principle, which has been applied with increasing success (Agarwal et al., 1976a), aims at increasing the hydrophobicity of the protecting groups in the synthetic intermediates. This, in turn, enables the use of more convenient and rapid solvent extraction procedures (see also the accompanying paper; Jones et al., 1978) in the fractionation of reaction mixtures.

Recently, the technique of high-pressure liquid chromatography (HPLC)[1] has been successfully applied in rapid separation of synthetic compounds in a variety of organic fields: e.g., vitamin $B_{12}$ (Woodward, 1973; Eschenmoser & Wintner, 1977); benzo[*a*]pyrene derivatives (Yagi et al., 1977; Nakanishi et al., 1977), 24,25-dihydroxyvitamin $D_3$ (Madhok et al., 1977); allenic retinals (Nakanishi et al., 1976); and tetrahydrocannabinols (Rasdan et al., 1973). Nucleic acids and their compounds have been separated by HPLC using anion exchange resins (Horvath, 1973; Gabriel & Michaelewsky, 1973; Hartwick & Brown, 1975; Van Boom & DeRooy, 1977; Gait & Sheppard, 1977) or anion exchange resins with superimposed lipophilic characteristics, e.g., RPC-5 (Pearson et al., 1971; Burd et al., 1975). In the present paper, we have investigated the use of HPLC in the separation of protected and unprotected oligo- and polynucleotides. The work has resulted in the development of a highly efficient method for the analytical separation of products formed at different stages of polynucleotide synthesis. The method is based on using columns in which the stationary phase consists of fully porous silica particles in the 10 $\mu$m size range, with octadecyl groups covalently linked to their surface. The mobile phases consist of neutral buffer systems containing varying amounts of acetonitrile. The influence of the different purine and pyrimidine bases, of the protecting groups, of the phosphate groups, and of the chain length of the retention of oligonucleotides on such columns has been studied, and, finally, the method has been applied systematically to the analysis and separation of the intermediates in the synthesis of the undecanucleotide, d(G-G-A-A-G-C-T-T-A-A-C). The latter was required as a part of the study of structure–function relationships in the $tRNA^{Tyr}_1$ gene promoter (Brown et al., 1977).

The above studies have provided a rational correlation between the structures of the oligonucleotides and their retention times on the above columns and, furthermore, the scope and usefulness of the HPLC technique in polynucleotide synthesis have been greatly enhanced by the work described in the accompanying paper (Jones et al., 1978). Analysis of unprotected oligonucleotides has also been carried out on the same columns and this method, which is also herein described, is comparable in sensitivity to the two-dimensional analysis (electrophoresis and homochromatography of 5′-$^{32}$P-labeled oligonucleotides (Sanger et al., 1974)). A brief account of the major findings of the present work has already been given (Fritz et al., 1978) and further applications in the preparative separations of oligonucleotides are being described separately (R. Belagaje, E. L. Brown, H.-J. Fritz, R. G. Lees, and H. G. Khorana, manuscript in preparation).

## Materials and Methods

High-pressure liquid chromatography was performed on a system consisting of the following components available from Waters Associates, Milford, Mass: two M6000A solvent delivery systems, a 660 solvent programer, a U6K injector, a 440 UV detector operating at wavelengths of 254 and 280 nm, a $\mu$Bondapak $C_{18}$ column (0.4 × 30 cm), and a Houston Instruments Omniscribe TM chart recorder. Pumps and solvent programer were operated in such a way that one pump delivered the aqueous buffer and the other acetonitrile. For the synthetic oligonucleotides described herein, the $\mu$Bondapak $C_{18}$ column has approximately $1.5 \times 10^4$ theoretical plates. This number ($N$) is derived from the distance ($a$) of the apex of the peak from the injection point and the width ($b$) of the peak at the baseline by the equation $N = 16(a/b)^2$. For preparative separations the column and detector were replaced with a Waters Associates Bondapak $C_{18}$/Porasil B column (0.7 × 183 cm) and an Altex Model 151 UV detector equipped with a preparative flow cell operating at a wavelength of 280 nm. Other materials and methods used for preparative HPLC will be described elsewhere (R. Belagaje, E. L. Brown, H.-J. Fritz, R. G. Lees, and H. G. Khorana, manuscript in preparation).

The aqueous buffers used were 0.1 M ammonium acetate and 0.1 M triethylammonium acetate (pH 7.0). The latter buffer was prepared by dissolving the calculated amount of triethylamine and slightly less than the equivalent amount of acetic acid in deionized water that had been passed through a 0.45-$\mu$m Millipore filter and titrating with acid to a pH of 7.0. Ammonium acetate solutions were passed through a similar Millipore filter. The triethylamine used to prepare the TEAA buffer was distilled from chlorosulfonic acid and then from potassium hydroxide. Reagent grade acetonitrile was distilled over a Vigreux column (50 cm) and then passed through a 0.5-$\mu$m Millipore filter for organic solvents.

Samples were dissolved in aqueous buffer containing 50% ethanol at a concentration of approximately 30 $A_{260}$ units/mL. These solutions were passed through a 0.45-$\mu$m Millipore filter using a filter disk 13 mm in diameter, screwcap filter holders, and disposable plastic syringes. Samples thus prepared could be stored at −20 °C for several months without significant decomposition. In most of the work, 0.1 M ammonium acetate containing varying percentages of acetonitrile was used as the mobile phase. A typical injection consisted of 5 $\mu$L of the

[1] The abbreviations used in this and the accompanying paper are as follows. The one-letter symbols for nucleosides and the symbols for the protecting groups on the bases, sugars, or phosphates in the polynucleotides are according to the IUPAC–IUB Commission on Biochemical Nomenclature Recommendations (1970), *J. Biol. Chem. 245*, 5171. A hyphen represents internal phosphate; p represents the terminal phosphate. When placed to the left of the nucleoside initial, p indicates phosphate group at the 5′ terminus. The symbol for the protecting group on this phosphate is placed at the left of the letter p: (Ac), acetyl; (*n*-$BuPh_2Si$), *n*-butyldiphenylsilyl; (*s*-$BuPh_2Si$), *sec*-butyldiphenylsilyl; (*t*-$BuPh_2Si$), *tert*-butyldiphenylsilyl. These symbols are placed at the right of the nucleoside initial to indicate their location on the 3′-hydroxyl group: (MeOTr), monomethoxytrityl, is placed at the left of the nucleoside initial to indicate its location on the 5′-hydroxyl group; an, *p*-anisoyl; bz, benzoyl; ib, isobutyryl. These immediately precede the one-letter symbol for the nucleoside and represent the protecting groups on the heterocyclic rings. Other abbreviations include: HPLC, high-pressure liquid chromatography; TPSE, 2-(*p*-tritylphenyl)sulfonylethyl; DIEA, diisopropylethylamine; TEAA, triethylammonium acetate; TEAB, triethylammonium bicarbonate; TPS, 2,4,6-triisopropylbenzenesulfonyl chloride; $CH_3CN$, acetonitrile.

sample solutions. The flow rate for the μBondapak $C_{18}$ column was 2 mL/min.

*Mono- and Oligonucleotides.* The protected deoxynucleosides, d(MeOTr)ibG, and the N-protected mononucleotides, d(pbzA), d(panC), and d(pibG), were prepared as described earlier (Khorana et al., 1961; Ralph & Khorana, 1961; Büchi & Khorana, 1972). Di-, tri-, and tetranucleotides were synthesized by the standard method given below and isolated by the solvent extraction procedure (Agarwal et al., 1976a; Jones et al., 1978).

*Synthesis of the Protected Oligonucleotides.* The condensation procedures for the synthesis of the di-, tri-, and oligonucletides were similar to those described earlier (Weber & Khorana, 1972; Büchi & Khorana, 1972; Agarwal et al., 1976a). Thus, both the mono- or oligonucleotide components were dissolved in dry pyridine and the mixture was rendered anhydrous by repeated evaporation, in vacuo, of added dry pyridine (four or five times). After each evaporation the flask was opened to the atmosphere of a drybox, under a positive pressure of dry $N_2$, and containing phosphorus pentoxide. During the last evaporation of the pyridine, the minimal amount of the solvent necessary for complete solubilization of the reaction components was allowed to remain and TPS was then added inside the drybox. The reaction mixture was kept at room temperature for 4 to 6 h with exclusion of moisture. The reaction was terminated by cooling followed by addition of 1 M DIEA in pyridine and an equal amount of water. After the solution was kept at room temperature for 16 h it was concentrated to a gum and quickly taken up in 0.2 M TEAB. After nonnucleotidic components were partially removed by solvent extraction, the desired product was isolated either by additional extractions or precipitation of the aqueous phase followed by preparative HPLC. The 3′-*O*-acetate group was removed by treatment of the oligonucleotide dissolved in pyridine/ethanol/water (4:3:3, v/v) with an equal volume of 2 M sodium hydroxide for 5 min at 0 °C.

The preparation of 3′-*O*-butyldiphenylsilyl ethers of mono- and oligonucleotides has been described in the accompanying paper (Jones et al., 1978).

*Semipreparative Purification of Unprotected Oligonucleotides.* The product of chemical synthesis, in the fully protected form (10–100-nmol range), was deprotected by the standard method of successive treatments with ammonia and pyridine–acetate buffer (Weber & Khorana, 1972). After evaporation of the solvent in vacuo, the residue was taken up in 0.2 mL of 0.2 M TEAB and separated from the by-products of deprotection by extraction with ethyl acetate and adsorption onto a small column (0.5–1.0 mL) of DEAE-cellulose (bicarbonate form). A 0.05 M TEAB wash eluted nonnucleotidic material while 1.0 M TEAB immediately brought off the oligonucleotide. The latter eluate was diluted with water, filtered through a 0.45-μm Millipore filter, and lyophilized. The residue was dissolved in 0.1 M TEAA to a final concentration of approximately 100 nmol/mL and then applied to the standard μBondapak $C_{18}$ column, 10 nmol being used per injection. Samples that had been previously purified by DEAE-cellulose chromatography in the presence of 7 M urea and contained 0.01 M Tris-HCl (pH 7.6) were either directly injected into the column or first desalted by adsorption onto a DEAE-cellulose column as described above. The effluent containing the purified oligonucleotide was collected, diluted with 0.02 M TEAB in order to reduce the acetonitrile concentration to 5% or less, and then lyophilized several times. The oligonucleotide was dissolved in 0.01 M Tris-HCl (pH 7.6) and stored at −85 °C.

*Two-Dimensional Fingerprinting.* Fingerprinting of the partial digest with snake venom phosphodiesterase was performed by previously described procedures (Sanger et al., 1974).

TABLE I: Retention Times of Protected and Unprotected 2′-Deoxyribonucleosides and Nucleotides on the μBondapak $C_{18}$ Column.

| Compd | Acetonitrile (% in 0.1 M aq $NH_4OAc$) | Retention time (min) |
|---|---|---|
| dC | 5 | 2.5 |
| dG | 5 | 5.0 |
| dT | 5 | 5.6 |
| dA | (a) 5 | 11.1 |
| | (b) 7 | 6.1 |
| pdC | 1 | 1.9 |
| pdT | 1 | 3.6 |
| pdG | 1 | 4.8 |
| pdA | (a) 1 | 9.2 |
| | (b) 4 | 2.2 |
| d(pibG) | 12 | 2.8 |
| d(pbzA) | 12 | 3.8 |
| d(panC) | (a) 18 | 2.5 |
| | (b) 20 | 2.4 |
| d[pT(Ac)] | 8 | 2.6 |
| d[pibG(Ac)] | 12 | 7.4 |
| d[panC(Ac)] | 20 | 3.9 |
| d(MeOTr) | 55 | 5.0 |
| d(MeOTr)bzA | 55 | 6.7 |
| d(MeOTr)anC | 55 | 10.1 |
| d[pT(*t*-$BuPh_2Si$)] | 42 | 2.7 |
| d[pibG(*s*-$BuPh_2Si$)] | 42 | 5.1 |
| d[panC(*s*-$BuPh_2Si$)] | (a) 42 | 5.5 |
| | (b) 55 | 3.5 |
| d[pbzA(*n*-$BuPh_2Si$)] | (a) 45 | 4.7 |
| | (b) 50 | 2.9 |

## Results and Discussion

*Nucleosides, Nucleotides, and Their Protected Derivatives.* The chromatographic system operated by virtue of the differences in the lipophilic interactions between the individual components of the reaction mixture and the nonpolar coating on the silica particles of the column. Elution with a polar mobile phase, therefore, brings off the components in the order of increasing lipophilicity, the least lipophilic component emerging first. The first set of data, which support this conclusion, is given in Table I. This table lists the retention times of unprotected and protected 2′-deoxyribonucleosides and -nucleotides on the μBondapak $C_{18}$ column. The 2′-deoxynucleosides require only a low (5%) acetonitrile concentration to give these four compounds a high mobility. Changing the acetonitrile concentration from 5 to 7%, which increases the lipophilicity of the eluent, decreases the retention time of 2′-deoxyadenosine from 11.1 to 6.1 min. The corresponding phosphomonoesters, being much more hydrophilic than the nucleosides, require only 1% acetonitrile in the mobile phase for their elution. It is interesting to note that the order of elution of pdT and pdG is the reverse of that of the corresponding nucleosides. In the above two unprotected series of compounds, cytosine derivatives have the highest mobility while the adenine derivatives have the lowest mobility.

The three mononucleotides containing the standard N-protecting groups were studied next (Table I). The protected cytosine nucleotide, d(panC), showed the lowest mobility of the amino protected mononucleotides. It can also be seen for

TABLE II: Retention Times of Protected Di-, Tri-, and Tetranucleotides with Free 5′-Phosphate End Groups on the μBondapak $C_{18}$ Column.

| Oligonucleotide | Acetonitrile (% in 0.1 M aq $NH_4OAc$) | Retention time (min) |
|---|---|---|
| d(pT-T) | 5 | 3.3 |
| d[pT-T(Ac)] | 8 | 3.8 |
| d(pbzA-T) | 15 | 2.3 |
| d(pbzA-bzA) | (a) 15 | 16.5 |
| | (b) 20 | 3.0 |
| d[pbzA-bzA(Ac)] | 20 | 5.2 |
| d(pibG-bzA) | 20 | 2.7 |
| d[pibG-bzA(Ac)] | 20 | 4.0 |
| d(pibG-anC) | (a) 17 | 6.2 |
| | (b) 20 | 3.2 |
| d[pibG-anC(Ac)] | 20 | 6.5 |
| d(pbzA-anC) | 20 | 3.2 |
| d[pbzA-anC(Ac)] | 20 | 5.0 |
| d(pbzA-bzA-ibG) | (a) 20 | 6.5 |
| | (b) 22 | 4.2 |
| d[pbzA-bzA-ibG(Ac)] | 22 | 8.0 |
| d(pbzA-bzA-anC) | 22 | 4.3 |
| d[pbzA-bzA-anC(Ac)] | 22 | 8.0 |
| d(pbzA-anC-bzA-anC) | 20 | 5.8 |
| d[pbzA-anC-bzA-anC(Ac)] | 20 | 9.1 |

this group of compounds that d(pbzA) elutes slightly slower than d(pibG). The retarding effect of the 3′-*O*-acetyl group can be seen by comparing the retention times for the third and fourth sets of compounds in Table I. Finally, the particularly strong retarding effect of the 5′-*O*-monomethoxytrityl and 3′-*O*-butyldiphenylsilyl groups is evident upon inspection of the retention times of the last two sets of compounds in Table I.

*Protected Di-, Tri-, and Tetranucleotides.* Next, it was appropriate to study the behavior of protected oligonucleotides carrying 5′-phosphate groups, which serve as the incoming blocks for elongation of the oligonucleotide chains at the 3′-hydroxyl end. The results obtained with a variety of di-, tri-, and tetranucleotides are assembled in Table II. These general comments may be made. Firstly, as expected, an oligonucleotide carrying a 3′-*O*-acetyl group is retained on the column markedly longer than the corresponding oligonucleotide carrying only a 3′-hydroxyl group. This is true for all of the oligonucleotides studied and listed in Table II. The second general point, which would also have been expected, is that the behavior of an oligonucleotide on the column can, to an approximation, be derived from its base composition. In other words, their behavior would be a summation of the mobilities of protected mononucleotides listed in Table I. For example, of the four dinucleotides, d(pT-T), d(pbzA-T), d(pbzA-bzA), and d(pbzA-anC), d(pT-T) elutes faster than any of the other dinucleotides. In the other three dinucleotides, in which the nucleotide at the 3′ end is systematically varied, the order of elution is determined by the latter nucleotide. Thus, d(pbzA-T) elutes before d(pbzA-bzA), which in turn elutes before d(pbzA-anC). Thirdly, it is interesting that the relative influence of the polar phosphate groups in reducing the affinity of the oligonucleotide for the column seems to diminish with an increase in the chain length. Furthermore, there is a trend to higher retention times with increasing chain length, at least up to a tetranucleotide. This is most clearly evident from a comparison of oligonucleotides having the same base composition but different chain length. For example, with a mobile phase of 20% acetonitrile in 0.1 M aqueous ammonium acetate, d(pbzA-anC) eluted in 3.2 min while its dimer, d(pbzA-anC-bzA-anC), eluted with a retention time of 5.8 min.

*The Hydrophobic Protecting Groups.* Classically, the hydrophobic methoxytrityl group has been used for protecting the 5′-hydroxyl group of the nucleoside with which the synthesis starts. Examples of the marked retarding effect of this group on the HPLC column have been given above (Table I) and a more detailed discussion is given in the following section. More recently, additional hydrophobic protecting groups, e.g., the TPSE group, have been introduced with advantage into polynucleotide synthesis methodology. The retention times on the HPLC column of a number of protected mono-, di-, and trinucleotides carrying the TPSE moiety on the 5′-phosphate group as well as of a few oligonucleotides carrying, in addition, another strongly hydrophobic group at the 3′-hydroxyl end were compared and the results are shown in Table III. Firstly, as seen for all the four protected mononucleotides (group I, Table III), the presence of the TPSE group causes strong retention on the column. Thus, whereas pdT (Table I) elutes in 3.6 min using only 1% acetonitrile in ammonium acetate, d[(TPSE)pT] shows a retention time of 4.0 min at as high as 45% acetonitrile concentration. Furthermore, although the TPSE derivatives of the mononucleotides showed a sequence of elution which was the same as that observed in the absence of the TPSE group, nevertheless, the latter group clearly dominates the elution behavior since mono-, di-, and trinucleotides carrying the TPSE group all showed very similar retention times.

Finally, in group III, Table III, is shown the influence of the 3′-*O*-butyldiphenylsilyl group on elution from the column. Thus, relative to d[(TPSE)pbzA-bzA], which eluted in 3.8 min at 45% acetonitrile content, the corresponding 3′-*O*-*t*-$BuPh_2Si$ derivative showed 8.9-min retention even at 55% acetonitrile concentration. In a second set, the TPSE derivatives of the trinucleotide d(pbzA-bzA-anC) were compared. As is seen (group III, Table III), the 3′-*O*-acetyl derivative showed a retention time of 15.0 min at 38% acetonitrile while the *s*-$BuPh_2Si$ derivatives required 48% acetonitrile at a retention time of 8.5 min.

The use of the hydrophobic silyl groups in the selective retardation, and hence unambiguous identification, is described in the accompanying paper (Jones et al., 1978).

*The Use of HPLC in the Stepwise Synthesis of Deoxyribopolynucleotides.* The standard methodology starts with 5′-methoxytrityl and N-protected deoxynucleosides and stepwise elongation is accomplished by condensation with protected mono-, di-, and higher oligonucleotide blocks. The initial studies with components listed in Tables I–III were essential for a rational use of HPLC in the actual separations of the condensation products at various stages of polynucleotide synthesis. Clearly, the aim at every step would be to be able to achieve separation of the desired condensation product and to be able to rapidly verify its identity. A number of factors would influence its emergence before or after the starting methoxytrityl-containing oligonucleotide. The first and most important influence would be the composition of the incoming oligonucleotide block originally carrying the 5′-phosphate group. A second factor would be the acetonitrile content of the eluting buffer. (Elution of the methoxytrityl oligonucleotide is very sensitive to slight changes in $CH_3CN$ concentration, while the oligonucleotides containing 5′-phosphate groups are relatively insensitive.) Thirdly, the initial products of condensations

TABLE III: Retention Times of TPSE-Protected Mono-, Di-, and Trinucleotides on the μBondapak $C_{18}$ Column.

| Oligonucleotide | Acetonitrile (% in 0.1 M aq $NH_4OAc$) | Retention time (min) |
|---|---|---|
| Group I | | |
| d[(TPSE)pT] | 45 | 4.0 |
| d[(TPSE)pibG] | (a) 42 | 7.8 |
| | (b) 45 | 5.8 |
| d[(TPSE)pbzA] | 45 | 6.2 |
| d[(TPSE)panC] | 45 | 6.8 |
| Group II | | |
| d[(TPSE)pbzA-T(Ac)] | 45 | 3.3 |
| d[(TPSE)pibG-ibG)] | (a) 42 | 4.3 |
| | (b) 45 | 3.4 |
| d[(TPSE)pbzA-bzA] | (a) 38 | 14.0 |
| | (b) 45 | 3.8 |
| d[(TPSE)panC-ibG(Ac)] | 45 | 5.4 |
| Group III | | |
| d[(TPSE)pbzA-bzA(*t*-$BuPh_2Si$)] | 55 | 8.9 |
| d[(TPSE)pbzA-bzA-anC(Ac)] | 38 | 15.0 |
| d[(TPSE)pbzA-bzA-anC(*s*-$BuPh_2Si$)] | 48 | 8.5 |

TABLE IV: Retention Times of Oligonucleotides Carrying 5′-Monomethoxytrityl End Groups on the μBondapak $C_{18}$ Column.

| Oligonucleotide | Acetonitrile (% in 0.1 M aq $NH_4OAc$) | Retention time (min) |
|---|---|---|
| Group I | | |
| d[(MeOTr)ibG-ibG(Ac)] | 40 | 4.9 |
| d[(MeOTr)ibG-ibG] | (a) 40 | 2.8 |
| | (b) 38 | 7.5 |
| | (c) 35 | 13.6 |
| d[(MeOTr)ibG-ibG-bzA-bzA] | 35 | 5.5 |
| d[(MeOTr)ibG-ibG-bzA-bzA-ibG-anC(Ac)] | 35 | 7.1 |
| d[(MeOTr)ibG-ibG-bzA-bzA-ibG-anC] | 35 | 5.8 |
| d[(MeOTr)ibG-ibG-bzA-bzA-ibG-anC-T-T(Ac)] | (a) 32 | 6.3 |
| | (b) 36[a] | 6.7 |
| d[(MeOTr)ibG-ibG-bzA-bzA-ibG-anC-T-T] | (a) 32 | 5.9 |
| | (b) 36[a] | 5.9 |
| d[(MeOTr)ibG-ibG-bzA-bzA-ibG-anC-T-T-bzA-bzA-anC(Ac)] | (a) 32 | 6.3 |
| | (b) 36[a] | 7.8 |
| d[(MeOTr)ibG-ibG-bzA-bzA-ibG-anC-T-T-bzA-bzA-anC] | (a) 32 | 5.9 |
| | (b) 36[a] | 6.9 |
| Group II | | |
| d[(MeOTr)bzA-anC-ibG-T-T] | (a) 37 | 4.3 |
| | (b) 35 | 9.2 |
| d[(MeOTr)bzA-anC-ibG-T-T-ibG-bzA(Ac)] | (a) 37 | 4.3 |
| | (b) 33 | 16.2 |
| d[(MeOTr)bzA-anC-ibG-T-T-ibG-bzA] | 35 | 7.9 |
| d[(MeOTr)bzA-anC-ibG-T-T-ibG-bzA-ibG-bzA] | 35 | 7.0 |
| d[(MeOTr)bzA-anC-ibG-T-T-ibG-bzA-ibG-bzA-bzA-bzA-ibG] | 35 | 4.6 |
| Group III | | |
| d[(MeOTr)ibG-anC-T-anC] | 35 | 6.7 |
| d[(MeOTr)ibG-anC-T-anC-ibG-ibG(Ac)] | 35 | 8.5 |
| d[(MeOTr)ibG-anC-T-anC-ibG-ibG] | 35 | 6.8 |
| d[(MeOTr)ibG-anC-T-anC-ibG-ibG-ibG-ibG-anC(Ac)] | 35 | 8.8 |
| d[(MeOTr)ibG-anC-T-anC-ibG-ibG-ibG-ibG-anC] | 35 | 6.6 |

[a] These elutions were performed using the acetonitrile concentrations indicated as dissolved in 0.1 M aqueous TEAA.

performed until now have carried an acetyl group at the 3′ end. This group is selectively deblocked in order to perform the next condensation. The product carrying the 3′-hydroxyl group would differ significantly from the initial product carrying the 3′-*O*-acetyl group in its retention on the HPLC column since, as described in Tables I and II as well as in Table IV, the acetyl group has a significant retarding effect. Indeed, this can be exploited for the characterization (see below) and for enhancing the separation of chain elongation reaction mixtures (R. Belagaje, E. L. Brown, H.-J. Fritz, R. G. Lees, & H. G. Khorana, manuscript in preparation) by either leaving or removing the acetyl group from the desired products. In the present work, three sets of polynucleotide syntheses were carried out in which extensive use was made of HPLC. In order to illustrate the different factors in the separations achieved by HPLC, the three groups of oligonucleotides involved in the

d(MeOTr)ibG

↓ (1) d[pibG(Ac)] + TPS
↓ (2) $OH^-$

d[(MeOTr)ibG-ibG]

↓ (1) d[pbzA-bzA(Ac)] + TPS
↓ (2) $OH^-$

d[(MeOTr)ibG-ibG-bzA-bzA]

↓ (1) d[pibG-anC(Ac)] + TPS
↓ (2) $OH^-$

d[(MeOTr)ibG-ibG-bzA-bzA-ibG-anC]

↓ (1) d[pT-T(Ac)] + TPS
↓ (2) $OH^-$

d[(MeOTr)ibG-ibG-bzA-bzA-ibG-anC-T-T]

↓ (1) d[pbzA-bzA-anC(Ac)] + TPS

d[(MeOTr)ibG-ibG-bzA-bzA-ibG-anC-T-T-bzA-bzA-anC(Ac)]

↓ (1) $NH_3$
↓ (2) $H^+$

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

FIGURE 1: Steps in the synthesis of the undecanucleotide, d(G-G-A-A-G-C-T-T-A-A-C).

three syntheses and their retention times are compiled in Table IV. Furthermore, of the three syntheses, that of the undecanucleotide d[(MeOTr)ibG-ibG-bzA-bzA-ibG-anC-T-T-bzA-bzA-anC] is described in detail. The synthetic steps used in this synthesis are shown in Figure 1, the behavior of all the synthetic intermediates on the HPLC column being compiled under group I in Table IV.

*The Undecanucleotide, d[(MeOTr)ibG-ibG-bzA-bzA-ibG-anC-T-T-bzA-bzA-anC] (Figure 1).* The starting material dG, its *N*-isobutyryl derivative dibG, and the methoxytrityl derivative of the latter, d(MeOTr)ibG, were all analyzed by HPLC. The results (Figure 2) were consistent with those given in Table I. Thus, to elute dG, dibG and d(MeOTr)ibG at comparable retention times, acetonitrile concentrations of 7, 20, and 55%, respectively, were required.

Condensation of d(MeOTr)ibG with d[pibG(Ac)] afforded d[(MeOTr)ibG-ibG(Ac)], which, in turn, was deacetylated to give d[(MeOTr)ibG-ibG]. The chromatogram from the coinjection of the acetylated and deacetylated product is shown in Figure 3A. Comparison of the percentages of acetonitrile needed for the elution of the dinucleotide monophosphate (40%) relative to that of d(MeOTr)ibG (55%) and of d[pibG(Ac)] (20%) shows that the condensation product has an affinity for the column which is intermediate between those of its components. Furthermore, as repeatedly observed, the 3′-*O*-acetyl end group has a marked retarding effect on the mobility of the lipophilic products. The characteristic retarding effect of the 3′-*O*-acetyl group is useful for confirming unambiguous assignment of the peaks to the starting material and to the product in chromatograms of condensation reaction mixtures. Thus, samples of a reaction mixture are analyzed before and after treatment with alkali. The peak corresponding to the starting material exhibits the same retention time in both chromatograms whereas the product, following loss of the acetyl group, will have a shorter retention time. Further examination of Figures 3A and 3B shows that a minor variation in acetonitrile concentration in the mobile phase has a very

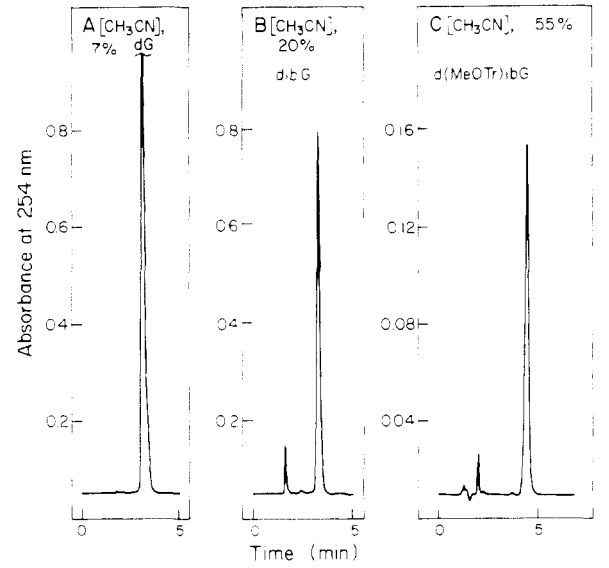


FIGURE 2: HPLC elution profiles of 2′-deoxyriboguanosine (A), 2-*N*-isobutyryl-2′-deoxyriboguanosine (B), and 2-*N*-isobutyryl-5′-*O*-monomethoxytrityl-2′-deoxyriboguanosine (C). The mobile phase was 0.1 M aqueous ammonium acetate containing 7% (A), 20% (B), and 55% (C) acetonitrile, respectively.

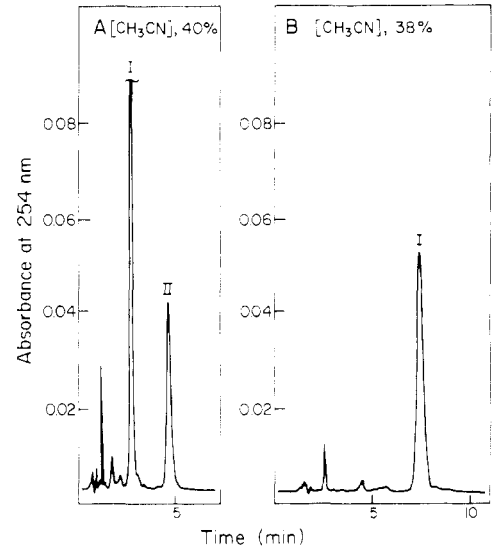


FIGURE 3: HPLC elution profile of d[(MeOTr)ibG-ibG] (I) and d[(MeOTr)ibG-ibG(Ac)] (II). The mobile phase was 0.1 M aqueous ammonium acetate containing 40% (A) and 38% (B) acetonitrile, respectively. To prepare the dinucleoside monophosphate, d[(MeOTr)-ibG-ibG(Ac)], an anhydrous pyridine solution (35 mL) of d(MeOTr)ibG (1.58 g, 2.59 mmol), d[pibG(Ac)] (3.14 g, 5.06 mmol), and TPS (2.98 g, 9.61 mmol) was allowed to react for 5.5 h. The protected dinucleoside monophosphate was isolated by solvent extraction and was obtained in 79% yield.

significant effect on the retention time of d[(MeOTr)ibG-ibG] (2.8 min at 40% to 7.5 min at 38%). This behavior, which is typical for fully protected oligonucleotides with 5′-*O*-monomethoxytrityl end groups, is further discussed below.

Condensation of d[(MeOTr)ibG-ibG] and d[pbzA-bzA(Ac)] followed by deacetylation of the reaction product yielded d[(MeOTr)ibG-ibG-bzA-bzA] (Figure 4A), which in turn was condensed with the dinucleotide, d[pibG-anC(Ac)], to give the hexanucleotide, d[(MeOTr)ibG-ibG-bzA-bzA-ibG-anC(Ac)] (Figure 4B). In the tetranucleotide preparation there was a large difference in retention time be-

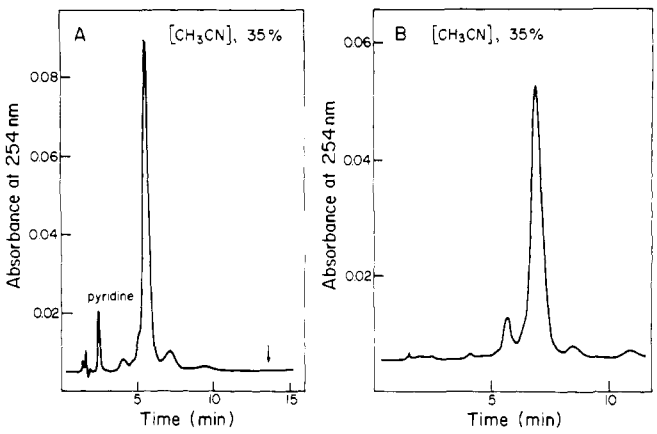


FIGURE 4: Elution profile of d[(MeOTr)ibG-ibG-bzA-bzA] (A) and d[(MeOTr)ibG-ibG-bzA-bzA-ibG-anC(Ac)] (B). The mobile phase for each chromatogram was 0.1 M aqueous ammonium acetate containing 35% acetonitrile. In panel A the expected elution point of the dinucleoside monophosphate is indicated by the arrow. To prepare the tetranucleotide, an anhydrous pyridine solution (40 mL) of d[(MeOTr)ibG-ibG] (1.06 g, 1.0 mmol), d[pbzA-bzA(Ac)] (3.0 g, 3.0 mmol), and TPS (3.0 g, 9.5 mmol) was allowed to react for 5.5 h. The desired d[(MeOTr)ibG-ibG-bzA-bzA(Ac)] was isolated by solvent extraction and then deacetylated and further purified by preparative HPLC (13% yield). The hexanucleotide was synthesized by treating an anhydrous pyridine solution (10 mL) of d[(MeOTr)ibG-ibG-bzA-bzA] (140 mg, 67 μmol) and d[pibG-anC(Ac)] (580 mg, 518 μmol) with TPS (535 mg, 1.7 mmol) for 5.5 h. The desired d[(MeOTr)ibG-ibG-bzA-bzA-ibG-anC(Ac)] was isolated by preparative HPLC and obtained in 78% yield.

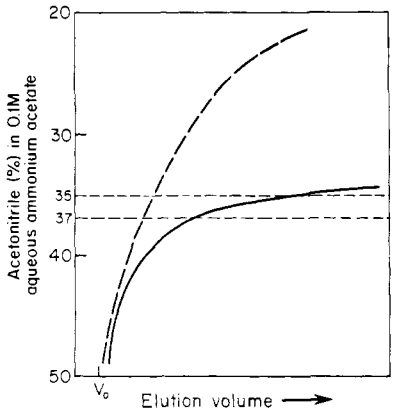


FIGURE 5: Comparative profiles of the influence of solvent composition on retention of oligonucleotides carrying 5′-*O*-monomethoxytrityl groups and oligonucleotides with 5′-phosphate or 5′-hydroxyl end groups. The solid line corresponds to the change in elution volume with the change in percent acetonitrile in 0.1 M aqueous ammonium acetate for any 5′-monomethoxytritylated oligonucleotide while the dashed line gives the corresponding value for any 5′-phosphate or 5′-hydroxyl derivative. For example, the relative elution volume for each class of compound at 35% acetonitrile is given by the distance from the void volume point, $V_0$, to the point where the solid and dashed lines cross the 35% acetonitrile line.

tween the starting material (13.6 min) and the deacetylated product (5.5 min). On the other hand, the acetylated hexanucleotide and its tetranucleotide precursor showed a much smaller difference in retention time at 35% acetonitrile: 7.1 min for the hexanucleotide vs. 5.5 min for the tetranucleotide. Because of this small difference in retention time the preparative HPLC did not completely resolve the two oligonucleotides and, therefore, only the hexanucleotide fractions free from the tetranucleotide were pooled.

As mentioned above, 5′-*O*-monomethoxytrityl oligonucleotides are very sensitive in their retentivity to small changes

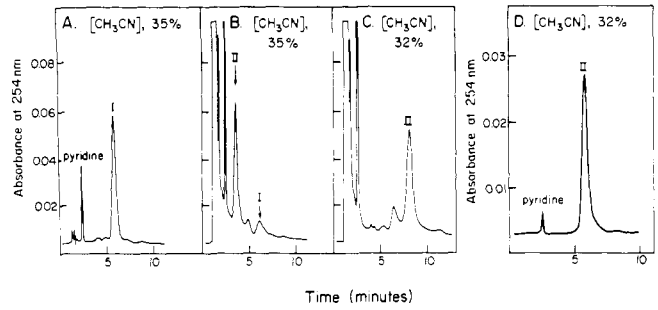


FIGURE 6: Elution profile of d[(MeOTr)ibG-ibG-bzA-bzA-ibG-anC] (A). The total reaction mixture from preparation of d[(MeOTr)ibG-ibG-bzA-bzA-ibG-anC-T-T(Ac)] (B and C) and the isolated octanucleotide after deacetylation (D). Peaks I and II corresponded to the hexa- and octanucleotide, respectively. The mobile phase for each chromatogram was 0.1 M aqueous ammonium acetate containing the indicated amount of acetonitrile. For preparation of the octanucleotide, an anhydrous pyridine solution (6 mL) of d[(MeOTr)ibG-ibG-bzA-bzA-ibG-anC] (80 mg, 30 μmol), d[pT-T(Ac)] (403 mg, 500 μmol), and TPS (500 mg, 1.65 mmol) was allowed to react for 4 h. After the usual workup and deacetylation, the desired d[(MeOTr)ibG-ibG-bzA-bzA-ibG-anC-T-T] was isolated by preparative HPLC in 30% yield.

in the concentration of acetonitrile in the mobile phase, whereas oligonucleotides with 5′-phosphate end groups do not show any marked changes in retention time with corresponding changes in acetonitrile concentration. This striking difference, which is illustrated in Figure 5, can be highly advantageous in analysis of condensation mixtures, as is illustrated in the preparation of the octanucleotide, d[(MeOTr)ibG-ibG-bzA-bzA-ibG-anC-T-T]. The starting hexanucleotide, whose HPLC profile is shown in Figure 6A, was condensed with d[pT-T(Ac)]. High-pressure liquid chromatographic analysis of the resulting mixture with 35% acetonitrile in the mobile phase gave the profile shown in Figure 6B. By a coinjection of the crude reaction mixture and the starting material, the peak with a retention time of 6.3 min was identified as the hexanucleotide. The identification of the peak with a retention time of 3.7 min as the octanucleotide followed from the significant increase in the retention time (8.2 min) of the peak when the acetonitrile content in the mobile phase was decreased from 35 to 32% (Figure 6C). In contrast, the fast eluting compounds present in Figure 6B did not exhibit any detectable change in retention time on lowering the acetonitrile concentration and these must, therefore, be the dinucleotide and its pyrophosphate. The purity of the octanucleotide, d[(MeOTr)ibG-ibG-bzA-bzA-ibG-anC-T-T], after its isolation by preparative HPLC, is shown in Figure 6D.

Finally, the last step involved the condensation of the protected octanucleotide d[(MeOTr)ibG-ibG-bzA-bzA-ibG-anC-T-T] with the trinucleotide d[pbzA-bzA-anC(Ac)]. The total condensation mixture was analyzed by HPLC using 35% acetonitrile in aqueous TEAA[2] (Figure 7A), the latter buffer being more effective for separation than ammonium acetate, which had been used until now. Of the three peaks obtained, peak I was identified as the unreacted octanucleotide (confirmation by coinjection with the starting material). Peak II was identified as the 3′-*O*-acetyl derivative of the octanucleotide since, after a brief alkaline treatment, the mobility of the product was identical with that of the starting material as

[2] The same buffer has been used in all subsequent work. The higher resolving power of this buffer, relative to ammonium acetate, could be due to the strong influence of the counterion on the partition of the anionic oligonucleotide between the stationary and the mobile phase.

TABLE V: Retention Times of Unprotected Oligonucleotides with 5′-Hydroxyl End Groups on the μBondapak $C_{18}$ Column.

| Entry no. | Oligonucleotide | Acetonitrile (% in 0.1 M aq TEAA) | Retention time (min) |
|---|---|---|---|
| 1 | d(G-C) | 8 | 3.5 |
| 2 | d(G-C-G) | (a) 8 | 5.0 |
| | | (b) 9 | 3.3 |
| 3 | d(G-C-G-T-C) | (a) 9 | 5.7 |
| | | (b) 10 | 3.4 |
| 4 | d(G-C-G-T-C-A-T) | 11 | 4.2 |
| 5 | d(G-C-G-T-C-A-T-T-T) | 12 | 4.3 |
| 6 | d(G-C-G-T-C-A-T-T-T-G-A-T-A) | 12 | 7.5 |
| 7 | d(A-A) | 8 | 9.0 |
| 8 | d(A-A-T) | (a) 10 | 6.3 |
| | | (b) 11 | 4.6 |
| 9 | d(A-A-T-T) | (a) 10 | 10.5 |
| | | (b) 11 | 6.3 |
| 10 | d(A-A-T-T-C) | (a) 10 | 11.8 |
| | | (b) 11 | 7.0 |
| 11 | d(A-A-T-T-C-T-T) | 12 | 7.5 |
| 12 | d(A-A-T-T-C-T-T-T-C) | (a) 12 | 9.2 |
| | | (b) 13 | 4.8 |
| 13 | d(T-G-A-T-G-C-G-C-C-C-C) | 12 | 2.5 |
| 14 | d(T-G-A-C-G-C-G-C-C-G-C) | 12 | 2.8 |
| 15 | d(G-G-A-A-G-C-G-G-G-G-C) | (a) 10 | 8.8 |
| | | (b) 11 | 3.2 |
| 16 | d(T-G-T-A-A-A-G-T-G-T-T) | 12 | 5.7 |
| 17 | d(A-C-G-T-T-G-A-G-A-A-A-G) | 12 | 4.0 |
| 18 | d(G-C-T-C-C-C-T-T-A-G-C-A) | 12 | 5.0 |
| 19 | d(A-T-C-A-G-C-A-G-G-A-C-G) | 12 | 5.1 |
| 20 | d(T-C-A-A-C-G-T-A-A-C-A-C) | 12 | 5.8 |

well.[3] Peak III contained the desired undecanucleotide whose separate elution profile is also shown in Figure 7B. The above assignments of peaks I–III (Figure 7A) were further confirmed by the two-dimensional fingerprinting method.

*The Dodecanucleotide, d[(MeOTr)bzA-anC-ibG-T-T-ibG-bzA-ibG-bzA-bzA-bzA-ibG], and the Nonanucleotide, d[(MeOTr)ibG-anC-T-anC-ibG-ibG-ibG-ibG-anC].* Table IV includes two additional sets of synthetic intermediates: one involved in the synthesis of the dodecanucleotide, d[(MeOTr)bzA-anC-ibG-T-T-ibG-bzA-ibG-bzA-bzA-bzA-ibG] (group II), and the second involved in the synthesis of the nonanucleotide, d[(MeOTr)ibG-anC-T-anC-ibG-ibG-ibG-ibG-anC] (group III). While detailed documentation of these syntheses will be given elsewhere (R. Belagaje, E. L. Brown, H.-J. Fritz, R. G. Lees, & H. G. Khorana, manuscript in preparation; E. L. Brown, R. Belagaje, H.-J. Fritz, R. A. Jones, & H. G. Khorana, manuscript in preparation), the data relevant to HPLC are included in order to be able to assess the separation method more generally. In group II, the hepta-, nona-, and dodecanucleotides could not be completely freed of their respective precursors by the use of HPLC alone. Likewise, the nonanucleotide listed in group III was freed of the unreacted hexanucleotide only after a semipreparative, μBondapak $C_{18}$ column (0.7 × 30 cm) became available.

*Separation of Unprotected Oligonucleotides by HPLC.* Up to this point, the present study has focused on the use of HPLC for expediting organochemical synthesis by acceleration of the separation of synthetic intermediates. However, it is clear that HPLC could potentially provide an invaluable method of high resolution for the purification and characterization of fully unprotected oligonucleotides. This has been demonstrated now to be the case by the following study of a large number of examples.

Table V shows the retention times of many synthetic fully unprotected oligonucleotides of varying size. It should be noted that all the compounds listed carry a 5′-hydroxyl group and not a 5′-phosphate group. The first general point, which is evident, and which would have been expected, is that the ace-

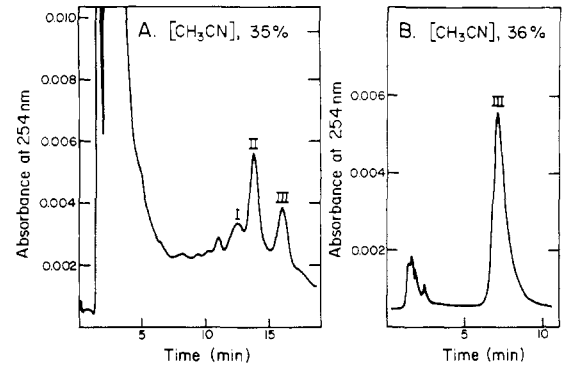


FIGURE 7: (A) HPLC elution profile of the total reaction mixture from preparation of d[(MeOTr)ibG-ibG-bzA-bzA-ibG-anC-T-T-bzA-bzA-anC(Ac)]. Peaks I, II, and III correspond to the octanucleotide, 3′-acetylated octanucleotide, and undecanucleotide, respectively. (B) Chromatogram of the isolated undecanucleotide. The mobile phase was 0.1 M aqueous TEAA containing 35 and 36% acetonitrile, respectively. To prepare this undecanucleotide, an anhydrous solution (1.4 mL) of d[(MeOTr)ibG-ibG-bzA-bzA-ibG-anC-T-T] (19 mg, 15 μmol), d[pbzA-bzA-anC(Ac)] (180 mg, 100 μmol), and TPS (130 mg, 435 μmol) was allowed to react for 16 h. Preparative HPLC of this reaction mixture gave the undecanucleotide in 19% yield.

[3] Some acetate ions must have been present in the trinucleotide block and this, on activation with TPS, caused 3′-O-acetylation. Occurrence of this side reaction has been encountered from time to time and fresh precipitation of this oligonucleotide block just before use is essential (see, e.g., Jacob & Khorana, 1965).

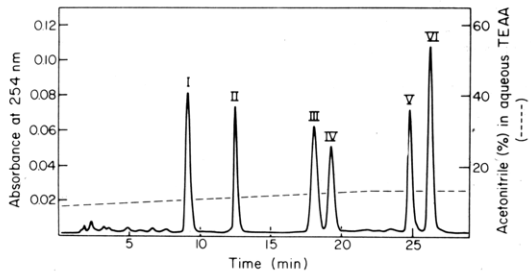


FIGURE 8: The HPLC elution profile of an artificial mixture of six unprotected oligonucleotides. The mobile phase was 0.1 M aqueous TEAA with varying percentages of acetonitrile, as shown. The composition of each peak was as follows: peak I contained d(A-A), peak II contained d(A-A-T), peak III corresponded to d(A-A-T-T), peak IV corresponded to d(A-A-T-T-C), peak V contained d(A-A-T-T-C-T-T), and peak VI contained d(A-A-T-T-C-T-T-T-C).

tonitrile concentrations in the aqueous buffer required for elution are much lower than those found for the protected oligonucleotides. Secondly, the retention is very sensitive to the base composition of the oligonucleotides, as can be seen from the relatively large difference in retention time for the dinucleoside monophosphates d(G-C) and d(A-A) and the trinucleoside diphosphates, d(G-C-G) and d(A-A-T). Similarly, to choose examples from higher oligonucleotides, examination of retention times for the undecanucleotides, **13–16**, and the dodecanucleotides, **17–20**, again showed that the method is very sensitive to base compositions. In particular, the undecanucleotides, d(T-G-A-T-G-C-G-C-C-C-C) (**13**) and d(T-G-A-C-G-C-G-C-C-G-C) (**14**), have nearly identical sequences (only the fourth and tenth bases from the 5′ end are different) yet they are readily resolved on the μBondapak column.

Another general observation can be made by comparing the retention times of oligonucleotides **1–6** and **7–12**. Compounds **1–6** form a set of successively longer oligonucleotides; the first five compounds (in protected form) had served as intermediates in the synthesis of the tridecanucleotide **6**. Similarly, oligonucleotides **7–12** form a second set. In both these sets of oligonucleotides the trend is to longer retention times with increasing chain length. The resolving power of the method was further demonstrated by a coinjection of all the six oligonucleotides of the second set. As seen in the resulting elution pattern (Figure 8), the six components were readily separated in less than 30 min. To separate this mixture on a DEAE-cellulose column in the presence of 7 M urea, the standard method, would have required several days. A similar separation was obtained for the first set of compounds (**1–6**) listed in Table V.

Although there may not be very large differences in retention between the different unprotected oligonucleotides, the pertinent question in utility of the method is its capacity to achieve high resolution and separation of the required product from the contaminating impurities. The superiority of the HPLC method to the DEAE-cellulose–urea method, a method which has invariably been used as the final step in purification, was shown in several examples (R. Belagaje, E. L. Brown, M. Gait, H. G. Khorana, & K. E. Norris, manuscript in preparation; E. L. Brown, R. Belagaje, H.-J. Fritz, R. A. Jones, & H. G. Khorana, manuscript in preparation), two of which are given below.

In one example, the crude product (~3.1 nmol) obtained upon deprotection of d[(MeOTr)ibG-ibG-bzA-bzA-ibG-anC-T-T-bzA-bzA-anC(Ac)] was purified by HPLC on this reverse phase column (four injections) to afford d(G-G-A-A-G-C-T-T-A-A-C) (1.4 nmol). A chromatogram from the analysis of the undecanucleotide is shown in Figure 9A. Purity of the resulting preparation was confirmed by the two-dimensional fingerprint of a partial snake venom phosphodiesterase digest of the isolated undecanucleotide (Figures 9B and 9C). Conversely, it may be concluded that the high-pressure liquid chromatographic method is of comparable sensitivity to the two-dimensional method for determining sample homogeneity.

In another example, a preparation of the tridecanucleotide, d(G-C-T-T-C-C-C-G-A-T-A-A-G), required for the total synthesis of the tyrosine tRNA gene (Agarwal et al., 1976b) was evidently contaminated with at least one additional tridecanucleotide, although the sample had been carefully purified by the DEAE-cellulose–urea method (T. Sekiya, T. Takeya, E. L. Brown, R. Belagaje, R. Contreras, H.-J. Fritz, M. Gait, R. G. Lees, K. Norris, M. Ryan, & H. G. Khorana, manuscript in preparation). Analysis of this tridecanucleotide by the HPLC method (Figure 10A) and the two-dimensional fingerprinting procedure (Figure 10B) showed that this preparation was contaminated by a mixture of at least four oligonucleotidic impurities. However, as indicated by the chromatogram shown in Figure 10A, the major component could be easily isolated by chromatography on a μBondapak $C_{18}$ column. Indeed, chromatography of this mixture (2.8 $A_{260}$

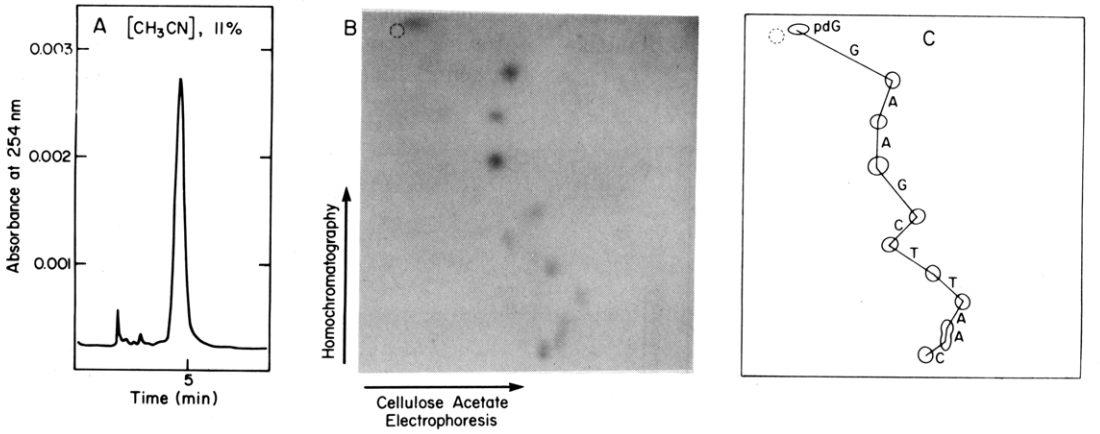


FIGURE 9: (A) Chromatogram of HPLC-purified d(G-G-A-A-G-C-T-T-A-A-C). The mobile phase was 0.1 M aqueous TEAA containing 11% acetonitrile. (B) A two-dimensional fingerprint of a partial snake venom phosphodiesterase digest of the undecanucleotide after it was phosphorylated using [$\gamma$-$^{32}$P]ATP and polynucleotide kinase. (C) An artist's conception of the two-dimensional fingerprint shown in panel B. The dashed circle in panel B indicates the position of the dye marker, xylene cyanol.

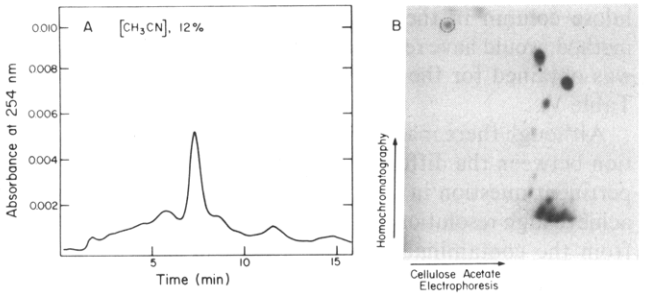


FIGURE 10: (A) HPLC elution pattern of impure tridecanucleotide, d(G-C-T-T-C-C-C-G-A-T-A-A-G). The mobile phase was 0.1 M aqueous TEAA containing 12% acetonitrile. (B) A two-dimensional fingerprint of a partial snake venom phosphodiesterase digest of this tridecanucleotide.

units) afforded purified tridecanucleotide (0.48 $A_{260}$ unit) which upon reanalysis was found to be free of most contaminants (Figure 11).

*General Comments.* The present work has demonstrated the great value of the HPLC method in the very rapid separation of synthetic intermediates in polynucleotide synthesis as well as in the purification of completely deprotected oligonucleotides. The influence of the different protecting groups currently used in synthetic methodology on the retention of the mono- and oligonucleotides on the HPLC columns was systematically investigated in order to provide the necessary background information. While the use of HPLC has already been made in the synthesis of a number of polynucleotide segments (R. Belagaje, E. L. Brown, H.-J. Fritz, R. G. Lees, & H. G. Khorana, manuscript in preparation) the stepwise synthesis of three segments has been discussed in the present paper. It is clear from the results that higher predictability and certainty would be advantageous in the separation of the desired condensation product following condensation with incoming di- and trinucleotide protected blocks. In actual fact, the separation depends, as discussed above, very much on the nature of the methoxytrityl-containing component and of the incoming block. Thus, the hexanucleotide, d[(MeOTr)ibG-ibG-bzA-bzA-ibG-anC(Ac)] (Table IV, group I), showed a longer retention time than its tetranucleotide precursor, evidently due to the relatively high lipophilicity of the added dinucleotide, d[pibG-anC(Ac)]. On the other hand, at the next step of the synthesis, addition of d[pT-T(Ac)] resulted in a relatively large decrease in the retention time for the octanucleotidic product, d[(MeOTr)ibG-ibG-bzA-bzA-ibG-anC-T-T(Ac)]. This was clearly caused by the small lipophilic contribution of the thymine residues relative to the standard hydrophilic contribution of the two phosphate dissociations of the added dinucleotide.

Similarly, other cases, which have already been pointed out above, were observed (groups II and III of Table IV) in which the condensation products could not be separated from the starting methoxytrityl oligonucleotides by HPLC alone. The value of HPLC in the synthetic work would be greatly enhanced if a new concept could be introduced into the total strategy which would ensure separation and unmistakable identification of the required product following each synthetic step. Thus, if the incoming di- or trinucleotide block were to carry on its 3′-hydroxyl group a very strongly hydrophobic group, the lipophilic contribution made by it to the elongated oligonucleotide chain would overwhelm other effects. Therefore, the desired condensation product would be set apart from the rest of the reaction components by virtue of its uniquely high retention in the HPLC column. Development of such a concept has indeed been realized and its successful application in polynucleotide synthesis is described in the accompanying paper (Jones et al., 1978).

The HPLC method undoubtedly offers many possibilities for separation in the nucleic acid field. Thus, it should be applicable to separations of nucleosides, nucleotides, and polynucleotides of the ribo series, especially for isolation and characterization of tRNA components. Many of the naturally occurring modified nucleosides have a more or less high degree of alkylation as their major unusual structural feature, which can be expected to give these derivatives enhanced affinity to the stationary phase. More systematic work is needed for evaluation of optimal chromatographic conditions for the different classes of compounds encountered in the ribo series; however, very encouraging preliminary results have already been obtained with modified ribonucleosides and mononucleotides as well as with the entire tRNAs (Fritz et al., 1978; J. E. Heckman, M. Silberklang, & H.-J. Fritz, unpublished results). It is tempting to suggest that this chromatography system may also be applicable for fractionation of DNA restriction fragments as has been observed with HPLC on RPC-5 columns (Landy et al., 1976; Hardies & Wells, 1976).

Finally, technical advances are needed in the design and preparation of reverse phase chromatography sorbents. Two sorbent-related problems have been encountered with the use of the μBondapak $C_{18}$ column: (1) 3′-*O*-butyldiphenylsilyl-nucleoside 5′-monophosphates exhibited broad and distorted peaks and (2) less than quantitative recoveries were found for the semipreparative purifications. Both problems were apparently due to incomplete functionalization of the exposed silanol groups on the silica support which lead to mixed-mode interactions, e.g., absorption, between the sample and the stationary phase. Technical improvements in the derivatization of silica particles are needed to provide new reverse-phase sorbents with more homogeneous surfaces that will eliminate

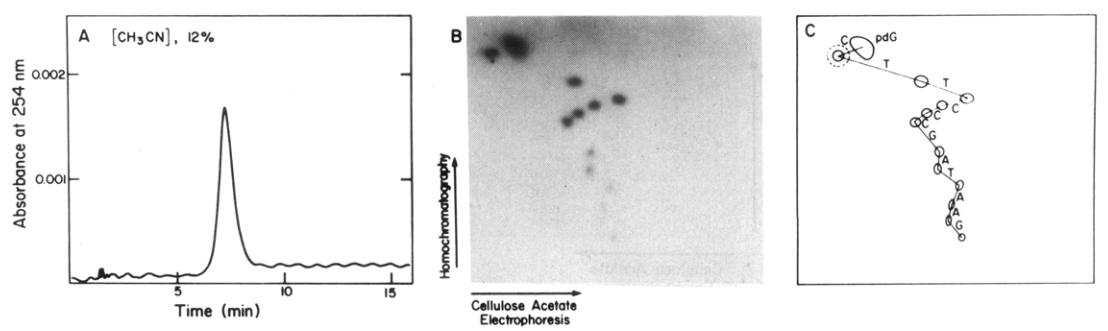


FIGURE 11: (A) Elution profile of HPLC-purified tridecanucleotide, d(G-C-T-T-C-C-C-G-A-T-A-A-G). The mobile phase was 0.1 M aqueous TEAA containing 12% acetonitrile. (B) A two-dimensional fingerprint of a partial snake venom phosphodiesterase digest of the purified tridecanucleotide. (C) An artist's conception of the two-dimensional fingerprint shown in panel B.

these mixed-mode associated phenomena.

## References


Agarwal, K. L., Berlin, Y., Fritz, H.-J., Gait, M., Kleid, D. G., Lees, R. G., Norris, K. E., Ramamoorthy, B., Smirnov, V., & Khorana, H. G. (1976a) *J. Am. Chem. Soc. 98,* 1065.

Agarwal, K. L., Caruthers, M. H., Büchi, H., van de Sande, J. H., & Khorana, H. G. (1976b) *J. Biol. Chem. 251,* 604.

Brown, E. L., Belagaje, R., Fritz, H.-J., Fritz, R. H., & Norris, K. (1977) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 36,* 732.

Büchi, H., & Khorana, H. G. (1972) *J. Mol. Biol. 72,* 251.

Burd, J. F., Larson, J. E., & Wells, R. D. (1975) *J. Biol. Chem. 250,* 6002.

Eschenmoser, A., & Wintner, C. E., (1977) *Science 196,* 1410.

Fritz, H.-J., Belagaje, R., Brown, E. L., Fritz, R. H., Jones, R. A., Lees, R. G., & Khorana, H. G. (1978) *Adv. Chem. Ser.* (in press).

Gabriel, T. F., & Michaeleswky, J. E. (1973) *J. Chromatogr. 80,* 263.

Gait, M. J., & Sheppard, R. C. (1977) *Nucleic Acids Res. 4,* 1135.

Hardies, S. C., & Wells, R. D. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73,* 3117.

Hartwick, R. A., & Brown, P. R. (1975) *J. Chromatogr. 112,* 651.

Horvath, C. (1973) *Method Biochem. Anal. 21,* 79.

Jacob, T. M., & Khorana, H. G. (1965) *J. Am. Chem. Soc. 87,* 368.

Jones, R. A., Fritz, H.-J., & Khorana, H. G. (1978) *Biochemistry 17* (following paper in this issue).

Khorana, H. G., Turner, A. F., & Vizsolyi, J. P. (1961) *J. Am. Chem. Soc. 83,* 686.

Khorana, H. G., Agarwal, K. L., Besmer, P., Büchi, H., Caruthers, M. H., Cashion, P. J., Fridkin, M., Jay, E., Kleppe, K., Kleppe, R., Kumar, A., Loewen, P. C., Miller, R. C., Minamoto, K., Panet, A., RajBhandary, U. L., Ramamoorthy, B., Sekiya, T., Takeya, T., & van de Sande, H. J. (1976) *J. Biol. Chem. 251,* 565, and accompanying papers.

Landy, A., Foeller, C., Reszelbach, R., & Dudock, B., (1976), *Nucleic Acids Res. 3,* 2575.

Madhok, T. C., Schnoes, H. K., & DeLuca, H. F. (1977) *Biochemistry 16,* 2142.

Nakanishi, K., Yudd, A. P., Crouch, R. K., Olson, G. L., Cheung, H. C., Govindjee, R., Ebrey, T. G., & Patel, D. J. (1976) *J. Am. Chem. Soc. 98,* 236.

Nakanishi, K., Kasai, H., Cho, H., Harvey, R. G., Jeffrey, A. M., Jennette, K. W., & Weinstein, I. B. (1977) *J. Am. Chem. Soc. 99,* 258.

Pearson, R. L., Weiss, J. F., & Kelmers, A. K. (1971) *Biochim. Biophys. Acta 228,* 770.

Ralph, R. K., & Khorana, H. G. (1961) *J. Am. Chem. Soc. 83,* 2926.

Rasdan, R. K., Uliss, D. B., & Dalzell, H. C. (1973) *J. Am. Chem. Soc. 95,* 2361.

Sanger, F., Donelson, J. E., Coulson, A. R., Kössel, H., & Fischer, D. (1974) *J. Mol. Biol. 90,* 315.

Sekiya, T., Contreras, R., Küpper, H., Landy, A., & Khorana, H. G. (1976) *J. Biol. Chem. 251,* 5124.

Van Boom, J. H., & DeRooy, J. F. M. (1977) *J. Chromatogr. 131,* 169.

Weber, H., & Khorana, H. G. (1972) *J. Mol. Biol. 72,* 219.

Woodward, R. B. (1973) *Pure Appl. Chem. 33,* 145.

Yagi, H., Akagi, H., Thakher, D. R., Mah, H. D., Koreeda, M., & Jerina, D. M. (1977) *J. Am. Chem. Soc. 99,* 2358.