

## CIV. Total Synthesis of the Structural Gene for an Alanine Transfer Ribonucleic Acid from Yeast. Chemical Synthesis of an Icosadeoxynucleotide Corresponding to the Nucleotide Sequence 21 to 40

H. WEBER AND H. G. KHORANA

The icosanucleotide having the sequence d-G-A-A-C-C-G-G-A-G-A-C-T-C-T-C-C-A-T-G† has been chemically synthesized. This and the icosanucleotide described in the succeeding paper are the longest polydeoxynucleotides with specific non-repeating nucleotide sequences which have been synthesized by chemical means so far. The icosanucleotide described in this paper corresponds to the nucleotide sequence 21 to 40 of the DNA strand complementary to the major yeast transfer RNA and, accordingly, has polarity opposite to that of the RNA. The synthesis involved the following steps using fully protected intermediates. Synthesis of d-MMTr-G<sup>Bz</sup>pA<sup>Bz</sup>pA<sup>Bz</sup> and condensation with d-pC<sup>An</sup>pC<sup>An</sup>.OAc to give the pentanucleotide, d-MMTr-G<sup>Bz</sup>pA<sup>Bz</sup>pA<sup>Bz</sup>pC<sup>An</sup>pC<sup>An</sup>; condensation of the latter with the trinucleotide d-pG<sup>iB</sup>pG<sup>iB</sup>pA<sup>Bz</sup>.OAc to yield the octanucleotide, d-MMTr-G<sup>Bz</sup>pA<sup>Bz</sup>pA<sup>Bz</sup>pC<sup>An</sup>pC<sup>An</sup>pG<sup>iB</sup>pG<sup>iB</sup>pA<sup>Bz</sup>. Then followed three successive condensations with the tetranucleotide blocks, d-pG<sup>iB</sup>pA<sup>Bz</sup>pC<sup>An</sup>pT.OAc, d-pC<sup>An</sup>pTpC<sup>An</sup>pC<sup>An</sup>.OAc and finally with d-dC<sup>An</sup>pA<sup>Bz</sup>pTpG<sup>iB</sup>.OiB. The deoxyguanosine residue in the last-mentioned block was labeled with <sup>3</sup>H to facilitate the purification and characterization, with regard to the 3'-OH terminus, of the icosanucleotide. The final product and the intermediate polydeoxynucleotides were characterized thoroughly in a variety of ways, both in the protected and the unprotected form.

### 1. Introduction

This and the succeeding paper (CV) describe syntheses of two icosanucleotides with specific sequences. These polydeoxynucleotides are the longest that have been chemically synthesized to date. The icosanucleotides are complementary to each other through one-half of their lengths (Fig. 1) and together span the nucleotide sequence 21 through 50 of the double-stranded DNA representing the gene for the major yeast alanine transfer RNA (CIII). The availability of the two icosanucleotides shown in Figure 1 made possible initial studies of the DNA-joining enzymes (paper CIII; Gupta *et al.*, 1968; Khorana *et al.*, 1968) and of the DNA polymerases (Gupta & Khorana, 1968; Kleppe, Ohtsuka, Kleppe, Molineux & Khorana, 1971). The studies with the DNA-joining enzymes, in turn, provided the theoretical basis for planning the

† The system of abbreviations for the protected oligo- and polynucleotides has been defined previously (Schaller & Khorana, 1963). The abbreviation iB stands for the isobutyryl group; MMT is monomethoxybutyl. For the unprotected oligodeoxynucleotides, the system used is that in current general use. In addition, a simplification has been introduced in that the letter p, occurring repetitively between the nucleotide initials, is deleted and in its place a hyphen is used. Thus, the oligonucleotide d-GpApApCpCpGpGpA is simply written as d-G-A-A-C-C-G-G-A. When so written, the nucleoside (G) at the left end in the above oligonucleotide carries a free 5'-OH group and the nucleoside at the right end (A) carries a free 3'-OH group. Further abbreviations, used mostly in the Experimental Methods section, are defined later. This system is used in all the accompanying papers.

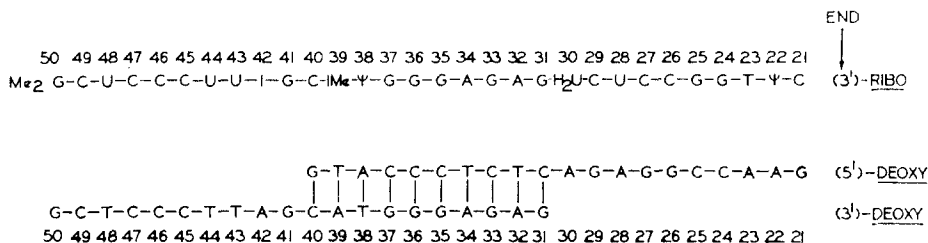


Fig. 1. Structures of the two partly complementary icosanucleotides representing the nucleotide sequence 21 to 50 of the double-stranded DNA corresponding to the gene for the yeast alanine transfer RNA. Note the opposite polarities of the two polydeoxynucleotide chains and their relation to the polarity of the transfer RNA sequence. As explained in the preceding paper (CIII), in the derivation of the DNA sequence from that of the alanine transfer RNA sequence, the assumption has been made that the minor bases are formed by modification of the four common ribonucleotide units; only the latter are believed to be used in transcription by the RNA polymerase.

total synthesis of the gene. The resulting plan has been presented and discussed in the preceding paper (CIII). The icosanucleotide whose synthesis is the subject of the present paper is segment 6 in the synthetic plan, the relevant part of which is shown in Figure 2.

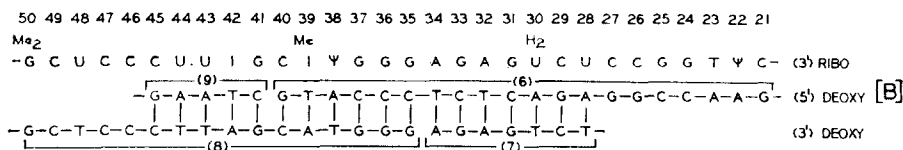


Fig. 2. Chemically synthesized polydeoxynucleotides corresponding to sequences 21 to 50 of the gene for the yeast alanine transfer RNA. The chemically synthesized segments are as shown within brackets. The icosanucleotide, whose synthesis is described in this paper, is segment 6 in the plan for the middle [B] part of the gene.

As already stated in the preceding paper, syntheses of the icosanucleotides described in this and the succeeding paper (CV) were undertaken at a time when the DNA-joining enzymes had not been discovered. While it was clear at the start of the present work that the basic concept for the future synthesis of bihelical DNA must utilize the inherent ability of complementary polynucleotide chains to form bihelical complexes, definitive knowledge of the minimal lengths of the polynucleotide segments to be synthesized chemically was not available. The decision to synthesize polydeoxynucleotides as large as but no longer than the icosanucleotides was determined primarily by two lines of reasoning. First, this length was just about the practical limit† of synthetic methodology with reasonable assurance of separation and unambiguous characterization of the products. Second, previous experience with the stabilities of short duplexes and the template size requirements of the DNA polymerases made it reasonably certain that structures sustained by ten base pairs (Fig. 1) would be adequate for testing many of the ideas then conceived for building long bihelical DNA (Khorana, 1967).

† This limit is, of course, to a great extent arbitrary but it is set primarily with regard to the total amount of effort required (1) in the preparation of adequate amounts of the protected oligonucleotide blocks and (2) in separation and characterization of the products. No doubt, by scaling up the intermediate preparations, substantially longer chain lengths of nucleotide units can be assembled.

## 2. Discussion of Experimental Methods

### (a) General methodology

The principles used in the present work as well as in the synthetic work reported in the accompanying papers (CV to CXI) have been developed previously in this laboratory (Khorana, 1967, 1968). Thus, internucleotide bonds are formed by condensation of the 5'-phosphomonoester end-group of a protected mono- or oligonucleotide with the 3'-hydroxyl end-group of a suitably protected deoxynucleoside, deoxymono- or oligonucleotide†. For the protection of the 5'-hydroxyl function, the mono-*p*-methoxytrityl group is customarily used, while an acetyl group is used for the protection of the 3'-hydroxyl group. For synthesis of the oligonucleotide blocks carrying a 5'-phosphate end-group, use of the  $\beta$ -cyano-ethyl group is common for the protection of the phosphomonoester group. While the use of the *N*-benzoyl group and the *N*-anisoyl group is standard for the protection of the adenine and cytosine rings, respectively, the protection of the amino group in the guanine ring cannot be claimed to have been completely satisfactory in the past. The *N*-acetyl group was introduced some years ago (Ralph, Connors, Schaller & Khorana, 1963) and it has been used in all the synthetic work reported so far, but its disappearance to a significant extent during the course of synthetic and isolation procedures has repeatedly been observed (e.g. Kössel, Buchi & Khorana, 1967)‡. In view of the extended nature of the present task, the use of an alternative group was highly desirable. The very stable *N*-benzoyl group (Ralph *et al.*, 1963; Smith, Drummond & Khorana, 1961) was used at the start of the present work (see below) as well as in work reported in an accompanying paper (CVIII). However, it was soon discovered to have alternative, undesirable properties. Thus, the presence of a low  $p^K$  (about 8.5) for the amide hydrogen dissociation was manifest and, as a consequence, there was marked retardation of compounds (especially those that were guanine rich) on DEAE-cellulose columns. Further investigation (see the succeeding paper) showed the *N*-isobutyryl group to be considerably more stable than the *N*-acetyl group and this group has been used for the protection of all the deoxyguanosine residues in the present work, except for the first one at the 5'-terminus. The same group has been used in the work reported in the accompanying papers. On the whole, the isobutyryl group is quite satisfactory although great care is still required during its use. Thus, the control of pH during prolonged chromatography is necessary and the duration of the repetitive alkaline treatments at successive steps must be kept to a minimum. Even so, in the present and in the accompanying work, detectable loss of the isobutyryl group was sometimes encountered§ and, therefore, search for a completely satisfactory protecting group for the guanine ring continues.

Except in the preparation of a few dinucleotides (see also accompanying papers) where dicyclohexylcarbodiimide was used as the condensing agent, aromatic sulfonyl

† Disadvantages of the alternative general approach using condensation of a 3'-phosphate end-group with the 5'-hydroxyl group of a second component have been discussed elsewhere (Weimann & Khorana, 1962).

‡ Furthermore, if no counter measures are taken to control the pH (see Experimental Methods section) the pH of the triethylammonium bicarbonate solutions during extended column chromatography rises to a value in the neighborhood of 9 and this leads to a slow but perceptible loss of *N*-acetyl groups.

§ For example, in the preparation of d-pG<sup>18</sup>pG<sup>18</sup>pA<sup>22</sup> (see Experimental Methods section) this could have been due largely to the alkaline treatment (much more than minimal) given for the removal of the  $\beta$ -cyanoethyl group. In later work, much shorter and gentler alkaline treatment was found to be adequate (paper CVIII).

chlorides were used most often for the purpose. Both mesitylenesulfonyl chloride (Jacob & Khorana, 1964) and tri-isopropylbenzenesulfonyl chloride (Lohrmann & Khorana, 1966) showed similar efficiency for the preparation of smaller oligonucleotides; however, the higher reaction rates obtained with MS† made this reagent preferable for reactions involving longer oligonucleotides. Based on the more recent results of a detailed study (Hachmann & Khorana, 1969), the amount of MS used for every phosphoryl dissociation was lower than that used in previous work (Kössel, Moon & Khorana, 1967; Ohtsuka & Khorana, 1967); nevertheless, the MS:P ratio‡ used (around 1.0) was still higher than that indicated to be optimal for work with thymidine oligonucleotides. It was felt that because of the bulky *N*-protected rings, especially purine deoxynucleoside residues, somewhat more forcing conditions were desirable. (The yields in condensation reactions between purine nucleotides are invariably lower than those obtained between pyrimidine nucleotides; see also accompanying papers.) The reaction times for condensation were two to three hours at room temperature: those for condensations using TPS being somewhat longer. A summary of the reaction conditions used and the yields obtained at the different steps is given in Table 1.

#### (b) *The oligonucleotide blocks and the synthetic plan*

Except for the first two steps, in which condensations involved addition of mononucleotides, the synthetic plan used the strategy of stepwise elongation of the oligonucleotide chain by addition of preformed oligonucleotide blocks. This is the approach which has been used with considerable success in more recent work on polydeoxynucleotide synthesis (Khorana, 1967, 1968). The most important aspect of the synthetic plan, therefore, was the choice of the oligonucleotide blocks to be used in the various condensation steps. The major considerations were as follows. Previous work has demonstrated that very large excesses of the oligonucleotide blocks carrying 5'-phosphate group relative to the component bearing 3'-hydroxyl group are necessary to obtain reasonable yields in the condensation steps. Further, the yields tend to drop with increase in the chain length of the polynucleotide bearing a 3'-hydroxyl group and in the size of the oligonucleotide block carrying the 5'-phosphate group. Nevertheless, the unusually large size of the polynucleotide chain to be built in the present work necessitated the use of at least tri- and tetranucleotide blocks, particularly in the latter stages of the synthesis. In this way it was hoped to separate successfully the expected components of the reaction mixtures. The current synthetic method involves separations on DEAE-cellulose anion-exchange columns, the method exploiting mainly the differences between the total negative charges carried by the various components. In addition to the over-riding consideration that the condensation product and the unreacted polynucleotide chains be clearly separated, it was desirable that the excess oligonucleotide block should also be recovered pure in the single-step chromatography. This would permit re-use of the oligonucleotide block in subsequent condensations. The re-use of di- and higher oligonucleotide blocks and the maximum

† Abbreviations used in this and the following papers: MS, mesitylenesulfonyl chloride; TPS, tri-isopropyl benzenesulfonyl chloride; DCC, dicyclohexylcarbodiimide; TEAB, triethylammonium bicarbonate; CE, cyanoethyl-; MMT, monomethoxytrityl-; DIEA, di-isopropylethylamine.

‡ The MS : P ratio indicates the molar equivalents of the condensing agent (MS) used per phosphoryl dissociation. The amount of the condensing agent to be used in a condensation reaction is then calculated by adding the total number of phosphoryl group dissociations present in the two mono- or oligonucleotidic components to be condensed. This method, first used by Hachmann & Khorana (1969), is used in this and all the accompanying papers.

TABLE I  
*Summary of conditions and yields of condensation reactions*

| 3'-OH-containing component (A)                                                                                                                                                                                                                   | Amount<br>(μmoles) | 5'-Phosphate-containing<br>component (B)                    | Amount<br>(μmoles) | B/A  | Reagent | Amount<br>(μmoles) | Equivalents†<br>(MS : P) | Pyridine<br>(ml.) | Time<br>(hr) | Yield‡<br>(%) |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------|-------------------------------------------------------------|--------------------|------|---------|--------------------|--------------------------|-------------------|--------------|---------------|
| d-CE-pC <sup>An</sup> pA <sup>Bz</sup>                                                                                                                                                                                                           | 1600               | d-pT-OAc                                                    | 6300               | 3.9  | TPS     | 11,900             | 0.75                     | 10                | 2½           | 46            |
| d-CE-pC <sup>An</sup> pA <sup>Bz</sup> pT                                                                                                                                                                                                        | 660                | d-pG <sup>IB</sup> -OiB                                     | 1870               | 2.8  | MS      | 5040               | 0.88                     | 10                | 1½           | 34            |
| d-CE-pG <sup>IB</sup>                                                                                                                                                                                                                            | 2500               | d-pG <sup>IB</sup> pA <sup>Bz</sup> -OAc                    | 2900               | 1.16 | TPS     | 13,000             | 1.17                     | 15                | 3½           | 25            |
| d-CE-pG <sup>IB</sup> pA <sup>Bz</sup>                                                                                                                                                                                                           | 1800               | d-pC <sup>An</sup> pT-OAc                                   | 2100               | 1.20 | TPS     | 9700               | 0.97                     | 20                | 3¼           | 37            |
| d-CE-pC <sup>An</sup> pT                                                                                                                                                                                                                         | 1300               | d-pC <sup>An</sup> pC <sup>An</sup> -OAc                    | 1930               | 1.5  | TPS     | 8400               | 1.0                      | 15                | 3¼           | 42            |
| d-MMTr-G <sup>Bz</sup> pA <sup>Bz</sup> pA <sup>Bz</sup>                                                                                                                                                                                         | 620                | d-pC <sup>An</sup> pC <sup>An</sup> -OAc                    | 2200               | 3.6  | TPS     | 6600               | 0.84                     | 15                | 5            | 67            |
| d-MMTr-G <sup>Bz</sup> pA <sup>Bz</sup> pA <sup>Bz</sup> pC <sup>An</sup> pC <sup>An</sup>                                                                                                                                                       | 35.6               | d-pG <sup>IB</sup> -OiB                                     | 605                | 17   | TPS     | 1800               | 1.33                     | 3                 | 2¼           | 67            |
| d-MMTr-G <sup>Bz</sup> pA <sup>Bz</sup> pA <sup>Bz</sup> pC <sup>An</sup> pC <sup>An</sup>                                                                                                                                                       | 178                | d-pG <sup>IB</sup> pG <sup>IB</sup> pA <sup>Bz</sup> -OAc   | 460                | 2.6  | TPS     | 3100               | 1.20                     | 10                | 3½           | 23            |
| d-MMTr-G <sup>Bz</sup> pA <sup>Bz</sup> pA <sup>Bz</sup> pC <sup>An</sup> pC <sup>An</sup> pG <sup>IB</sup>                                                                                                                                      | 18.1               | d-pG <sup>IB</sup> pA <sup>Bz</sup> -OAc                    | 203                | 10.6 | TPS     | 590                | 0.84                     | 1.5               | 2½           | 39            |
| d-MMTr-G <sup>Bz</sup> pA <sup>Bz</sup> pA <sup>Bz</sup> pC <sup>An</sup> pC <sup>An</sup> pG <sup>IB</sup> -<br>pG <sup>IB</sup> pA <sup>Bz</sup>                                                                                               | 6.1                | d-pG <sup>IB</sup> pA <sup>Bz</sup> -OAc                    | 157                | 26   | TPS     | 590                | 1.12                     | 1.0               | 3            | 55            |
| d-MMTr-G <sup>Bz</sup> pA <sup>Bz</sup> pA <sup>Bz</sup> pC <sup>An</sup> pC <sup>An</sup> pG <sup>IB</sup> -<br>pG <sup>IB</sup> pA <sup>Bz</sup>                                                                                               | 45.7               | d-pG <sup>IB</sup> pA <sup>Bz</sup> pC <sup>An</sup> pT-OAc | 400                | 8.75 | MS      | 2320               | 1.0                      | 7                 | 2¾           | 28            |
| d-MMTr-G <sup>Bz</sup> pA <sup>Bz</sup> pA <sup>Bz</sup> pC <sup>An</sup> pC <sup>An</sup> pG <sup>IB</sup> -<br>pG <sup>IB</sup> pA <sup>Bz</sup> pG <sup>IB</sup> pA <sup>Bz</sup> pC <sup>An</sup> pT                                         | 20                 | d-pC <sup>An</sup> pTpC <sup>An</sup> pC <sup>An</sup> -OAc | 460                | 23   | MS      | 2180               | 0.90                     | 6                 | 2½           | 34            |
| d-MMTr-G <sup>Bz</sup> pA <sup>Bz</sup> pA <sup>Bz</sup> pC <sup>An</sup> pC <sup>An</sup> pG <sup>IB</sup> -<br>pG <sup>IB</sup> pA <sup>Bz</sup> pG <sup>IB</sup> pA <sup>Bz</sup> pTpC <sup>An</sup> pTpG <sup>An</sup> -<br>pC <sup>An</sup> | 6.4                | d-pC <sup>An</sup> pA <sup>Bz</sup> pTpG <sup>IB</sup> -OAc | 216                | 34   | MS      | 1000               | 0.90                     | 1.5               | 2¼           | 25            |

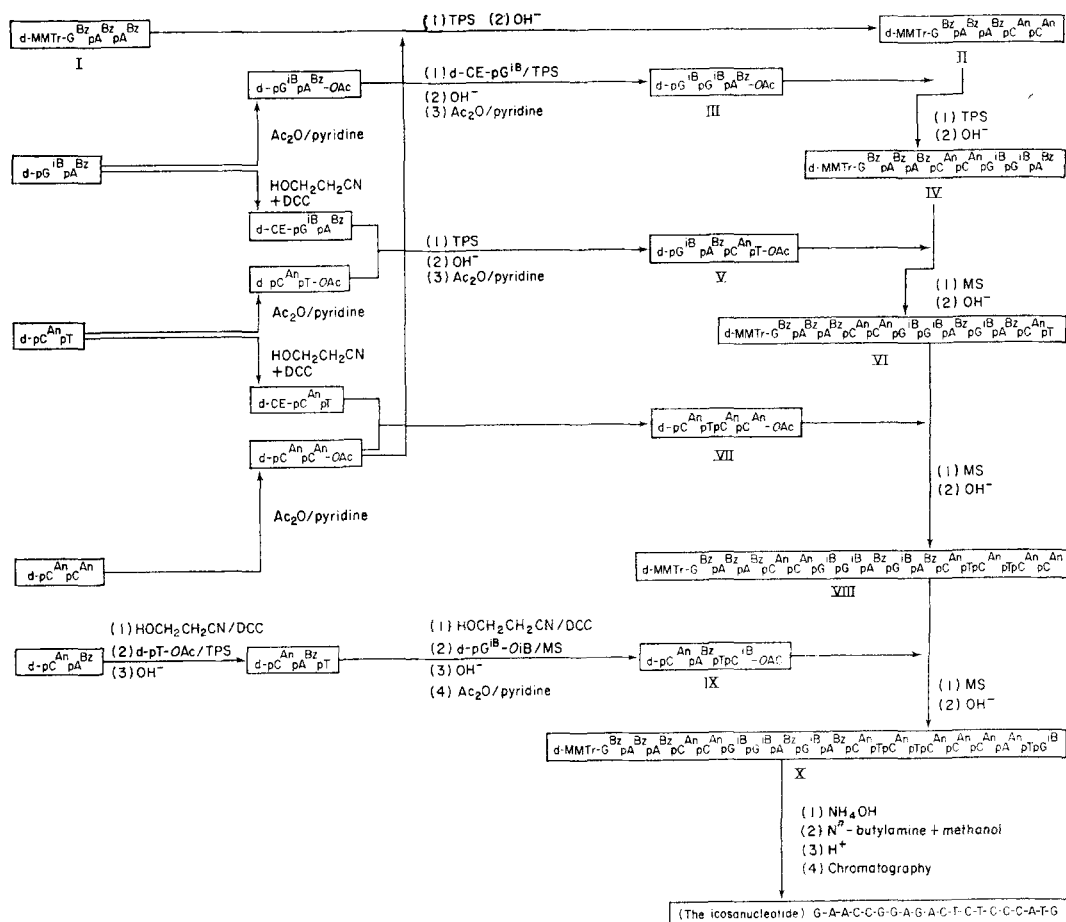
† Molar equivalents of sulfonyl chloride as based on total equivalents of phosphoryl dissociations.

‡ Isolated yield of pure product.

possible use of sets of dinucleotide blocks, e.g.  $pC^{An}pC^{An}$ ,  $d-pC^{An}pT$  and  $d-pG^{iB}pA^{Bz}$ , which were prepared in the present work, were important considerations for economy. Indeed, the present synthesis and that reported in the succeeding paper were far more demanding than any other syntheses carried out in this laboratory, in that they required extremely large quantities of oligonucleotide blocks (several millimoles of each of the dinucleotide blocks and amounts of 0.5 to 1 m-mole of the tri- and tetranucleotide blocks).

CHART 1

*Steps in the synthesis of the icosanucleotides d-G-A-A-C-C-G-G-A-G-A-C-T-C-T-C-C-C-A-T-G*



Based on the above considerations, the plan finally adopted for the synthesis is outlined in Chart 1. Thus, it is seen that from the stage of the trinucleoside diphosphate,  $d-MMTr-G^{Bz}pA^{Bz}pA^{Bz}$  (I), the icosanucleotide was obtained in five condensation steps by successive additions of a dinucleotide, a trinucleotide and three tetranucleotides to the growing chain. Only four protected dinucleotide blocks were used in addition to the protected mononucleotides, and three of the dinucleotides ( $d-pG^{iB}pA^{Bz}$ ,  $d-pC^{An}pT$  and  $d-pC^{An}pC^{An}$ ) were used more than once. The trinucleotide block,  $d-pG^{iB}pG^{iB}pA^{Bz}$  (III), was prepared by the condensation of  $d-pG^{iB}pA^{Bz}$  with the protected mononucleotide  $d-CE-pG^{iB}$ . This method led to a rather poor yield of the

desired product and this result was in fact to be expected, for at least two reasons. First, condensations between end-groups carried by purine nucleotides are in general not very satisfactory and, second, experience has shown that the synthesis of a trinucleotide by the addition of an activated mononucleotide to a protected (cyanoethyl) dinucleotide proceeds usually better than by the alternative route, which involves condensation of an activated dinucleotide to a protected (cyanoethyl) mononucleotide.

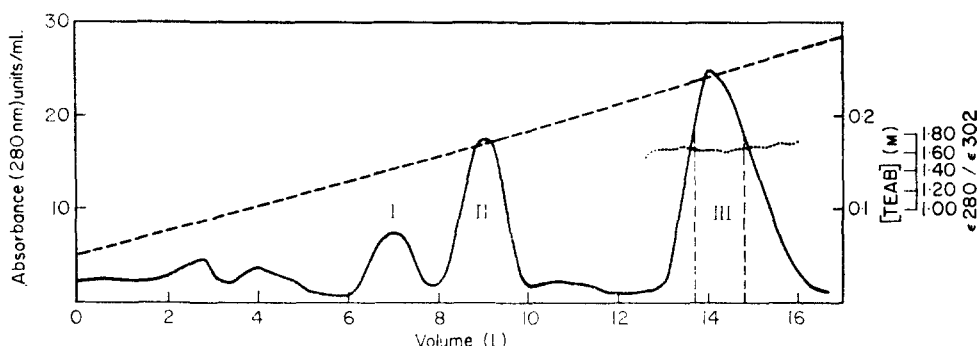


FIG. 3. Condensation of d-CE-pG<sup>1</sup>BpA<sup>2</sup>B<sub>z</sub> with dpC<sup>ααα</sup>pT-OAc. Separation of the reaction products on a DEAE-cellulose (bicarbonate) column (4 cm × 110 cm) pre-equilibrated at 4°C with 0.05 M-TEAB in 40% ethanol. Gradient elution was as shown by the dashed line, 40% ethanol being used as solvent. Peak I is unreacted dpG<sup>1</sup>BpA<sup>2</sup>B<sub>z</sub>, peak II is excess dpC<sup>ααα</sup>pT, and peak III is the tetranucleotide, dpG<sup>1</sup>BpA<sup>2</sup>B<sub>z</sub>pC<sup>ααα</sup>pT. (----) Concn TEAB; (....) ε<sub>280</sub>/ε<sub>302</sub>; (—) absorbance.

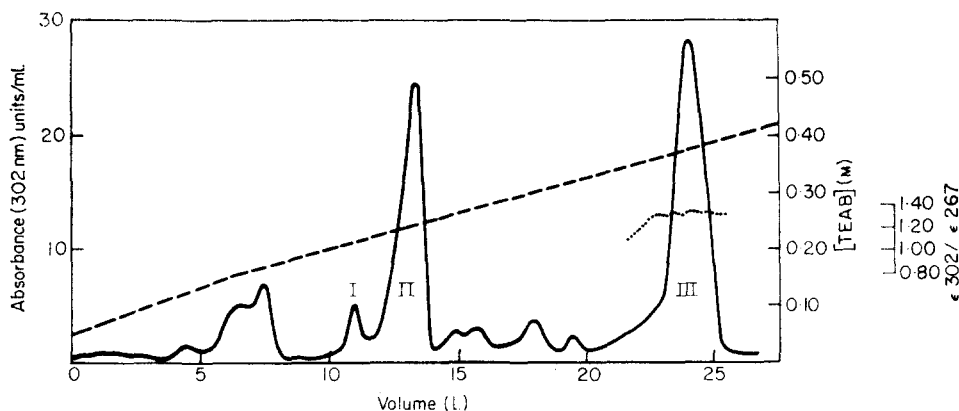


FIG. 4. Condensation of d-CE-pC<sup>ααα</sup>pT with dpC<sup>ααα</sup>pC<sup>ααα</sup>-OAc. Separation of the reaction products on a DEAE-cellulose (bicarbonate) column (4 cm × 110 cm) pre-equilibrated at 4°C with 0.05 M-TEAB in 20% ethanol. Gradient elution was as shown by the dashed line, 20% ethanol being used as solvent. Peak I is unreacted dpC<sup>ααα</sup>pT, peak II is excess dpC<sup>ααα</sup>pC<sup>ααα</sup>, and peak III is the tetranucleotide, dpC<sup>ααα</sup>pTpC<sup>ααα</sup>pC<sup>ααα</sup>. (----) Concn TEAB; (....) ε<sub>302</sub>/ε<sub>267</sub>; (—) absorbance.

The two tetranucleotide blocks, d-pG<sup>1</sup>BpA<sup>2</sup>B<sub>z</sub>pC<sup>ααα</sup>pT (V) and d-pC<sup>ααα</sup>pTpC<sup>ααα</sup>pC<sup>ααα</sup> (VII), were prepared by condensation of two suitably protected dinucleotide blocks as shown in Chart 1. Yields of around 40% (Table 1) were obtained in the coupling reactions, the chromatographic patterns obtained on work-up of the reaction mixtures in these preparations are shown in Figures 3 and 4, respectively. This approach appeared to be superior to the alternative, in which the tetranucleotide block may

be prepared by repetitive addition of a mononucleotide. The third tetranucleotide block, d-pC<sup>An</sup>pA<sup>Bz</sup>pTpG<sup>IB</sup> (IX), on the other hand, was prepared by the stepwise addition of protected mononucleotides to the dinucleotide d-pC<sup>An</sup>pA<sup>Bz</sup> (Chart 1). The chief reason for this was the decision to use radioactively labeled guanine nucleoside in the 3'-terminal position of the icosanucleotide (X). The main reason for the decision was that the synthesis and isolation of an icosanucleotide at the time that this project was undertaken (late 1965) was without parallel in past experience with polynucleotide synthesis. It was hoped that the presence of the label at the 3'-terminus would aid greatly in the isolation, purification and characterization of the icosanucleotide. Therefore, as shown in Chart 1, the dinucleotide, d-pC<sup>An</sup>pA<sup>Bz</sup>, was taken to the trinucleotide, d-pC<sup>An</sup>pA<sup>Bz</sup>pT, and after reprotection of the 5'-phosphomonoester group in this trinucleotide condensation was effected with the protected <sup>3</sup>H-labeled deoxyguanosine 5'-phosphate (d-pG<sup>IB</sup>OiB). After a standard work-up followed by chromatography (Fig. 5) the product was acetylated to give (IX). Actually, for convenience, the labeled tetranucleotide (IX) was prepared on a small scale at a relatively high specific activity and later was diluted with a suitable amount of an unlabeled preparation.

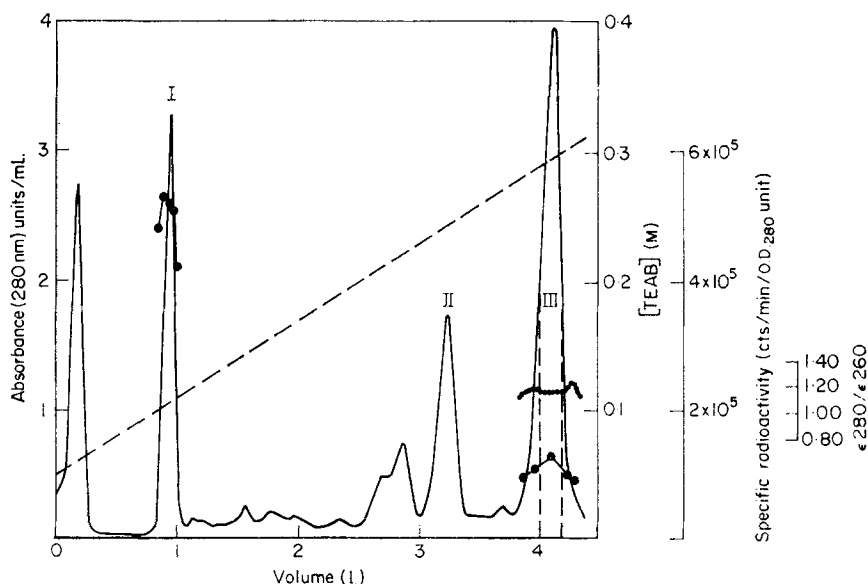


FIG. 5. Condensation of d-CE-pC<sup>An</sup>pA<sup>Bz</sup>pT with d-p[8-<sup>3</sup>H]G<sup>IB</sup>-OiB. Separation of the reaction products on a DEAE-cellulose (bicarbonate) column (1.8 cm × 60 cm) pre-equilibrated with 0.05 M-TEAB in 20% ethanol. Gradient elution was as shown by the dashed line, 20% ethanol being used as solvent. Peak I is the excess mononucleotide, peak II the unreacted trinucleotide, and peak III is the tetranucleotide, dpC<sup>An</sup>pA<sup>Bz</sup>pCpG<sup>IB</sup> (<sup>3</sup>H). (-----) Concn TEAB; (—●—●—, radioactivity; (....)  $\epsilon_{280}/\epsilon_{260}$ ; (—) absorbance.

### (c) The synthesis

The synthesis of the trinucleoside diphosphate (I, Chart 1) was carried out by the stepwise condensation of the protected deoxyguanosine with protected deoxyadenylate. Because of contamination of the commercially available deoxyguanosine by the corresponding ribonucleoside, *N*-benzoyl-5'-*O*-methoxytrityldeoxyguanine was sub-

jected to a rigorous purification step involving chromatography on Silicagel. The work-up procedures for the protected dinucleoside phosphate, d-MMTt-G<sup>Bz</sup>pA<sup>Bz</sup>, and the trinucleoside diphosphate, d-MMTr-G<sup>Bz</sup>pA<sup>Bz</sup>pA<sup>Bz</sup>, were simplified by using a solvent extraction procedure analogous to that developed previously for the thymidine analogue (Jacob & Khorana, 1965). Thus, the use of anion-exchange chromatography was unnecessary for d-MMTr-G<sup>Bz</sup>pA<sup>Bz</sup> and d-MMTr-G<sup>Bz</sup>pA<sup>Bz</sup>pA<sup>Bz</sup> was considerably enriched by solvent extraction before the final chromatographic step.

Condensation of I with d-pC<sup>An</sup>pC<sup>An</sup>-OAc gave the pentanucleotide (II, Chart 1), which was isolated by column chromatography after a standard work-up. The chromatographic pattern obtained is shown in Figure 6; the yield of the isolated

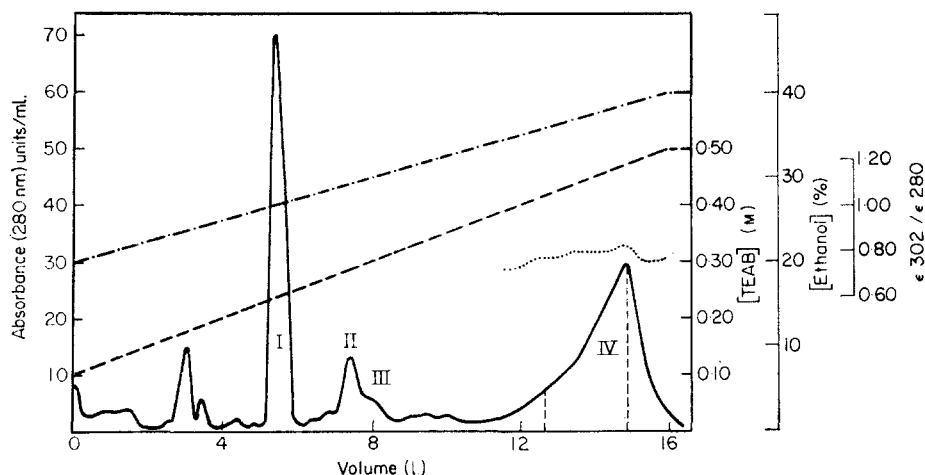


Fig. 6. Condensation of d-MMTrG<sup>Bz</sup>pA<sup>Bz</sup>pA<sup>Bz</sup> with d-pC<sup>An</sup>pC<sup>An</sup>-OAc. Separation of the reaction products on a DEAE-cellulose (bicarbonate) column (4 cm × 120 cm) pre-equilibrated at 4°C with 0.05 M-TEAB in 20% ethanol. Elution was carried out using linear gradients of TEAB and ethanol, as shown by the dashed lines. Peak I contained the excess dinucleotide, peak II mainly the corresponding pyrophosphate. d-MMTrG<sup>Bz</sup>pA<sup>Bz</sup>pA<sup>Bz</sup> was in shoulder III and peak IV contained the product, d-MMTrG<sup>Bz</sup>pA<sup>Bz</sup>pA<sup>Bz</sup>pC<sup>An</sup>pC<sup>An</sup>. (----) Concn TEAB; (---) concn. ethanol; (....)  $\epsilon_{302}/\epsilon_{280}$ ; (—) absorbance.

product was 67%. The preparation of the octanucleotide IV (Chart 1) was carried out by two different routes. In one method, the protected pentanucleotide (II) was condensed with the mononucleotide, d-pG<sup>IB</sup>OiB, to form the hexanucleotide, d-MMTr-G<sup>Bz</sup>pA<sup>Bz</sup>pA<sup>Bz</sup>pC<sup>An</sup>pC<sup>An</sup>pG<sup>IB</sup> (chromatographic pattern in Fig. 7) and the latter was then condensed with the dinucleotide d-pG<sup>IB</sup>pA<sup>Bz</sup>-OAc. The chromatographic pattern obtained after work-up of the latter reaction mixture is shown in Figure 8. In the second method, the protected pentanucleotide (II) was directly condensed with the trinucleotide d-pG<sup>IB</sup>pG<sup>IB</sup>pA<sup>Bz</sup>-OAc. The elution pattern obtained upon chromatography of the reaction mixture is shown in Figure 9. In terms of the yields (Table I), the methods were about equivalent, although by using a larger excess of the protected trinucleotide d-pG<sup>IB</sup>pG<sup>IB</sup>pA<sup>Bz</sup>-OAc, no doubt the yield in the step II → IV could have been improved. (As mentioned above and described in the Experimental Methods section, unusual difficulties were encountered in the preparation of this trinucleotide block.)

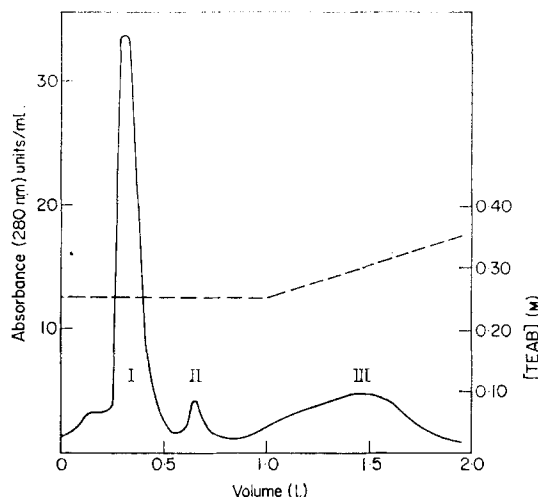


FIG. 7. Condensation of  $d\text{-MMTrG}^{Bz}pA^{Bz}pA^{Bz}pC^{An}pC^{An}$  with  $dpG^{IB} \cdot OiB$ . Separation of the reaction products on a DEAE-cellulose (bicarbonate) column (1.9 cm  $\times$  85 cm) pre-equilibrated at 4°C with 0.05 M-TEAB in 40% ethanol. The elution was carried out using a linear salt concentration gradient as shown by the dashed line, 40% ethanol being used as solvent. Peak I is the excess mononucleotide, peak II is unreacted pentanucleotide, and peak III is the hexanucleotide. (-----) Concn TEAB; (—) absorbance.

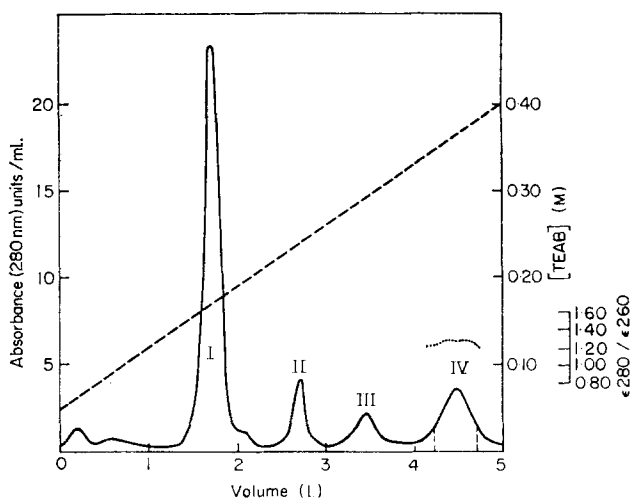


FIG. 8. Condensation of  $d\text{-MMTrG}^{Bz}pA^{Bz}pA^{Bz}pC^{An}pC^{An}pG^{IB}$  with  $dpG^{IB}pA^{Bz} \cdot OAc$ . Separation of the reaction products on a DEAE-cellulose (bicarbonate) column (1.9 cm  $\times$  68 cm) pre-equilibrated at 4°C with 0.05 M-TEAB in 40% ethanol. Gradient elution was as shown by the dashed line, 40% ethanol being used throughout. Peak I was the excess dinucleotide, peak II the pyrophosphate, peak III the hexanucleotide and peak IV the octanucleotide. (-----) Concn TEAB; (....)  $\epsilon_{280}/\epsilon_{260}$ ; (—) absorbance.

Although not included in Chart 1, the preparation of the decanucleotide,  $d\text{-G-A-A-C-C-G-G-A-G-A}$ , was undertaken in a small-scale experiment; this involved the condensation of the protected octanucleotide (IV) with the protected dinucleotide,  $d\text{-pG}^{IB}pA^{B2} \cdot OAc$ . The chromatographic pattern obtained is shown in Figure 10. Full characterization of the decanucleotide is included in the Tables in the Experimental

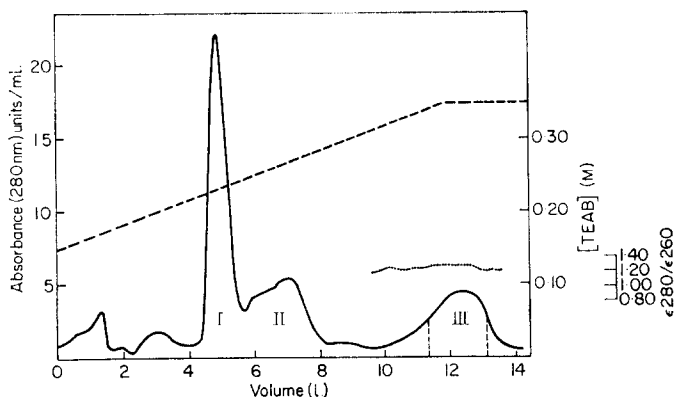


FIG. 9. Condensation of  $d\text{-MMTrG}^{\text{Bz}}\text{pA}^{\text{Bz}}\text{pA}^{\text{Bz}}\text{pC}^{\text{An}}\text{pC}^{\text{An}}$  with  $\text{dpG}^{\text{IB}}\text{pG}^{\text{IB}}\text{pA}^{\text{Bz}}\text{-OAc}$ . Separation of the reaction products on a DEAE-cellulose (bicarbonate) column ( $4\text{ cm} \times 90\text{ cm}$ ) pre-equilibrated at  $4^\circ\text{C}$  with  $0.05\text{ M}$ -TEAB in  $40\%$  ethanol. Gradient elution was as shown by the dashed line,  $40\%$  ethanol being used as solvent. Peak I was the excess trinucleotide, peak II was mainly pentanucleotide, peak III was the octanucleotide. (----) Concn TEAB; (....)  $\epsilon_{280}/\epsilon_{260}$ ; (—) absorbance.

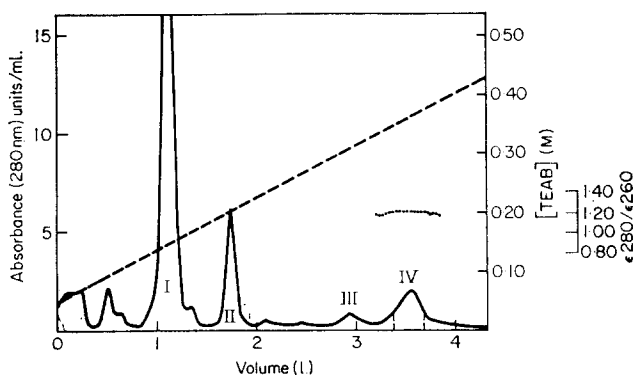


FIG. 10. Condensation of the protected octanucleotide with  $\text{dpG}^{\text{IB}}\text{pA}^{\text{Bz}}$ . Separation of the reaction products on a DEAE-cellulose (bicarbonate) column ( $1.9\text{ cm} \times 68\text{ cm}$ ) pre-equilibrated at  $4^\circ\text{C}$  with  $0.05\text{ M}$ -TEAB in  $40\%$  ethanol. Gradient elution was as shown by the dashed line, the solvent being  $40\%$  ethanol. Peak I was the excess dinucleotide, peak II the pyrophosphate of the dinucleotide, peak III the unreacted octanucleotide and peak IV the decanucleotide. (----) Concn TEAB; (....)  $\epsilon_{280}/\epsilon_{260}$ ; (—) absorbance.

Methods section. There was considerable interest in this decanucleotide for initial studies with DNA ligases and for studies of the stabilities of short DNA duplexes formed from this decanucleotide. The use of the decanucleotide in such studies has already been reported in part and is described in accompanying papers (Gupta *et al.*, 1968; paper CIII; paper CXIII).

Condensation of the octanucleotide (IV) with the tetranucleotide,  $d\text{-pG}^{\text{IB}}\text{pA}^{\text{Bz}}\text{pC}^{\text{An}}\text{-pT-OAc}$  (V), gave the dodecanucleotide (VI), the isolation of which by chromatography is shown in Figure 11. Concurrent studies already referred to with thymidine polynucleotides (Hachmann & Khorana, 1969) had shown that as high an excess of the incoming block as feasible was desirable for reactions to proceed satisfactorily. In the present work, in some cases (e.g. the preparation of the octa- and the dodecanucleotides), the ratio of the block to the long-chain compound was improved by carrying out the condensation in two successive batches, the second batch using the

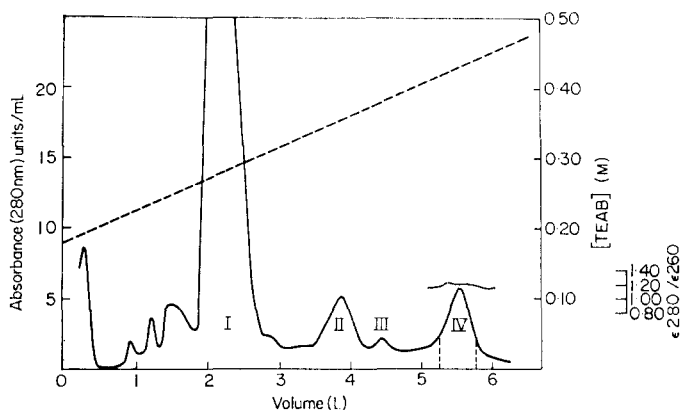


Fig. 11. Condensation of the protected octanucleotide with  $\text{dpG}^{18}\text{pA}^{22}\text{pC}^{22}\text{pT-OAc}$ . Separation of the reaction products on a DEAE-cellulose (bicarbonate) column ( $2.5 \text{ cm} \times 80 \text{ cm}$ ) pre-equilibrated at  $4^\circ\text{C}$  with  $0.05 \text{ M-TEAB}$  in  $40\%$  ethanol. Gradient elution was as shown by the dashed line, the solvent being  $40\%$  ethanol. Peak I was the excess tetranucleotide, peak II was the unreacted octanucleotide, peak III was the pyrophosphate of the tetranucleotide and peak IV was the dodecanucleotide. (----)  $\text{Concn TEAB}$ ; (....)  $\epsilon_{280/260}$ ; (—) absorbance.

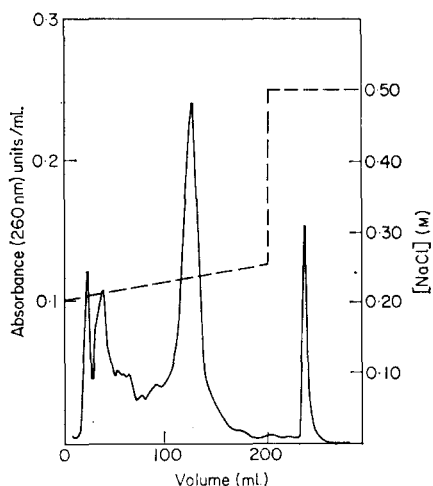


Fig. 12. Chromatography of the dedecanucleotide, after removal of the protecting groups, on a DEAE-cellulose (Whatman DE23, chloride form) column ( $0.5 \text{ cm} \times 90 \text{ cm}$ ). The column was pre-equilibrated at  $4^\circ\text{C}$  with  $0.05 \text{ M-sodium chloride}$  containing  $7 \text{ M-urea}$  and  $0.02 \text{ M-Tris-HCl}$  ( $\text{pH } 7.5$ ). Elution was carried out using a linear concentration gradient of sodium chloride as indicated by the dashed line, in the presence of  $7 \text{ M-urea}$  and  $0.02 \text{ M-Tris-HCl}$ . The smaller peaks at the beginning of the chromatogram consist of non-nucleotide material (no absorption maximum between  $200 \text{ nm}$  and  $320 \text{ nm}$ ). (----)  $\text{Concn NaCl}$ ; (—) absorbance.

block material recovered from the first-batch condensation. Because of the progressively smaller-scale reactions (with respect to the growing chain), this device was not necessary for the two subsequent steps ( $\text{VI} \rightarrow \text{VIII} \rightarrow \text{X}$ ).

For characterization of the dodecanucleotide (VI) a sample was fully deprotected and subjected to anion-exchange chromatography on a DEAE-cellulose column in the presence of  $7 \text{ M-urea}$ . The pattern obtained is shown in Figure 12. The smaller peaks at the beginning of the chromatogram consisted of non-nucleotide material. The main

peak corresponded to the dodecanucleotide, whereas the last peak, which was eluted at higher salt concentration, probably corresponded to the dodecanucleotide which had suffered partial depurination during the acidic removal of the methoxytrityl group. More recent work (see accompanying papers) has focussed attention sharply on the significant extent to which depurination can occur when acetic acid alone is used. More recently (accompanying papers), the method of choice for removal of the methoxytrityl group has used acetic acid buffered with a small amount of pyridine (Ohtsuka & Khorana, 1967)†.

Condensation of the protected dodecanucleotide (VI) with the tetranucleotide (VII), which contained only pyrimidine nucleotides, proceeded reasonably well to give the hexadecanucleotide (VIII). The results of chromatographic elution of the protected product are shown in Figure 13. As in the preceding chromatographic patterns, the

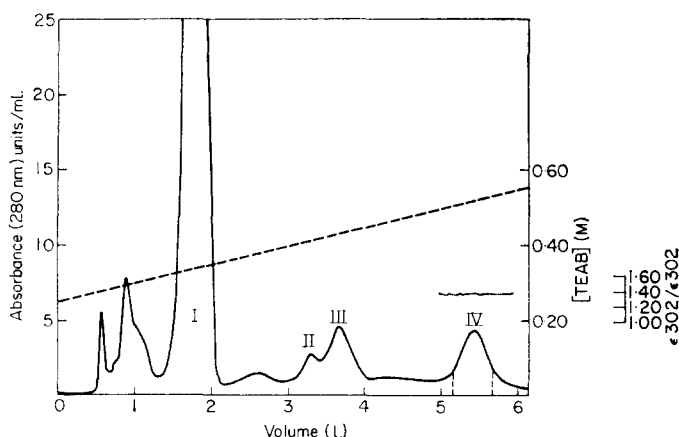


FIG. 13. Condensation of the protected dodecanucleotide with  $\text{dpC}^{An}\text{pTpC}^{An}\text{pC}^{An}\text{-OAc}$ . Separation of the reaction products on a DEAE-cellulose (bicarbonate) column ( $2.5 \text{ cm} \times 80 \text{ cm}$ ) pre-equilibrated at  $4^\circ\text{C}$  with  $0.05 \text{ M}$ -TEAB in  $40\%$  ethanol. Gradient elution was as shown by the dashed line, with  $40\%$  ethanol as solvent. Peak I was the excess tetranucleotide block, peak II was the corresponding pyrophosphate, peak III was the dodecanucleotide and peak IV was the hexadecanucleotide. (—) Concn TEAB; (· · ·)  $\epsilon_{302}/\epsilon_{302}$ ; (—) absorbance.

desired product (VIII) was very well separated from the starting material (VI). For characterization, a sample of the protected hexadecanucleotide was completely deprotected and examined by anion-exchange chromatography in the presence of  $7 \text{ M}$ -urea. The pattern obtained is shown in Figure 14. Once again, the major peak corresponded to the desired material, the small fore-peaks consisting of non-nucleotidic material. For the material eluted last at high salt concentration, the possibility offered above for the dodecanucleotide probably applies here also.

The last step in the synthesis of the protected icosanucleotide (X) involved the condensation of (VIII) with the radioactively labeled (IX) and the chromatographic pattern obtained after work-up of the reaction mixture is shown in Figure 15. As may be seen, the separation of the unreacted hexadecanucleotide (peak III) from the desired product (peak IV) was again satisfactory. The icosanucleotide was the major product

† In contrast, in certain experiments the removal of the methoxytrityl group might not have been complete. This circumstance would also lead to a peak eluted after the desired polynucleotide (paper CXI).

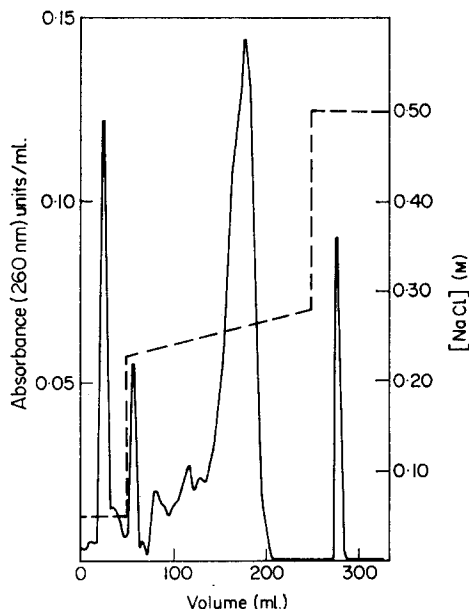


FIG. 14. Chromatography of the hexadecanucleotide, after removal of the protecting groups, on a DEAE-cellulose (Whatman DE23, chloride form) column (0.5 cm  $\times$  90 cm). The column was pre-equilibrated at 4°C with 0.05 M-sodium chloride containing 7 M-urea and 0.02 M-Tris-HCl (pH 7.5). Elution was carried out using a linear concentration gradient of sodium chloride as indicated by the dashed line, in the presence of 7 M-urea and 0.02 M-Tris-HCl. The first two peaks are non-nucleotidic material (no absorption maximum between 200 nm and 230 nm). (----) Concn NaCl; (—) absorbance.

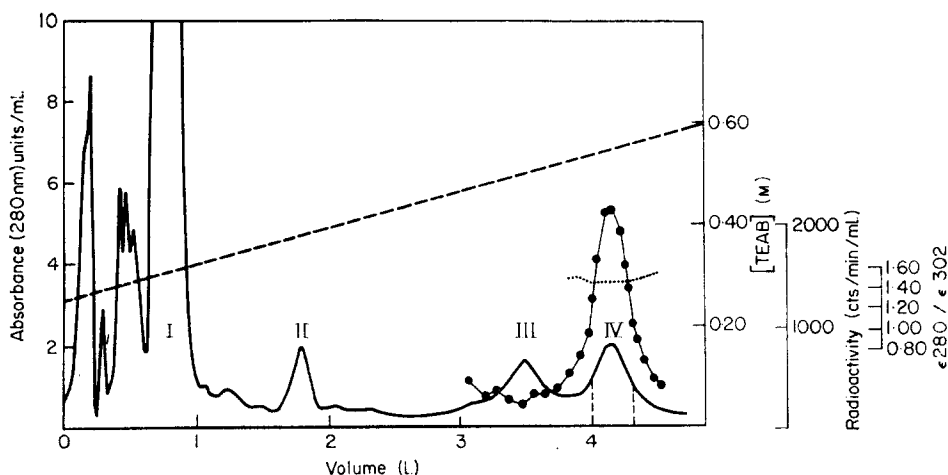


FIG. 15. Condensation of the protected hexadecanucleotide with d-pC<sup>A</sup>n pA<sup>B</sup>z pTpG<sup>18</sup>-OAc. Separation of the reaction products on a DEAE-cellulose (bicarbonate) column (1.8 cm  $\times$  60 cm) pre-equilibrated at 4°C with 0.05 M-TEAB in 40% ethanol. The elution was as shown by the dashed line, the solvent being 40% ethanol. Peak I was the excess tetranucleotide, peak II the corresponding pyrophosphate, peak III the unreacted hexadecanucleotide and peak IV the icosanucleotide. (----) Concn TEAB; (....)  $\epsilon_{280}/\epsilon_{302}$ ; —●—●—, radioactivity; (—) absorbance.

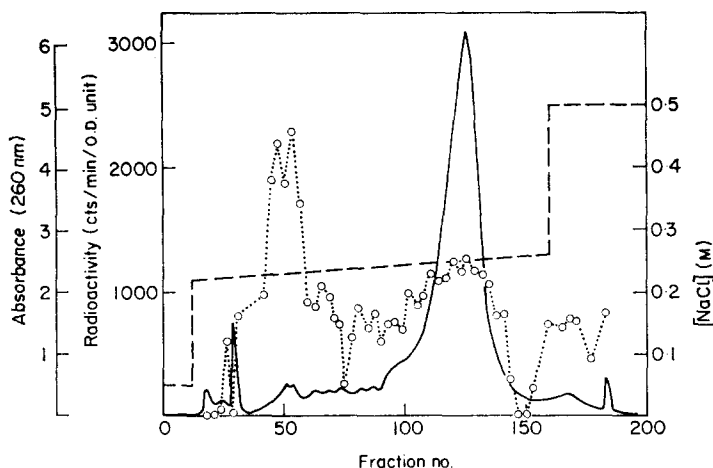


FIG. 16. Chromatography of the  $^3\text{H}$ -labeled icosanucleotide, after removal of the protecting groups, on a DEAE-cellulose (Whatman DE32, chloride form) column (0.5 cm  $\times$  80 cm). The column was pre-equilibrated at room temp. with 0.05 M-sodium chloride containing 7 M-urea and 0.02 M-Tris-HCl (pH 7.5). Elution was carried out using a linear concentration gradient of sodium chloride as indicated by the dashed line, in the presence of 7 M-urea and 0.02 M-Tris-HCl. (----) Concn NaCl; (—) absorbance;  $\circ \circ \circ \circ$ , radioactivity.

in the reaction. It may further be noted that the extinction ratio for 280 nm/302 nm was essentially constant across the icosanucleotide peak (peak IV) and, furthermore, measurement of the radioactivity in the eluate gave immediate assurance of the production of the icosanucleotide. A sample of the protected icosanucleotide (X) was completely deprotected and the product subjected to analytical chromatography in the presence of 7 M-urea. The pattern obtained is shown in Figure 16. The major ultraviolet-absorbing peak corresponded to the icosanucleotide but it is worthy of note that a small amount of radioactive material, presumably the tetranucleotide block, was present as the fore-peak. Purification by the method of chromatography illustrated in Figure 16 has been used on a larger scale in several preparations of the icosanucleotide, which have been used in enzymic studies.

#### (d) *Characterization and removal of the protecting groups*

The spectral and chromatographic properties of the fully protected polynucleotides, the monomethoxytrityl-protected polynucleotides and of the unprotected products are given in Tables 2, 3 and 4, respectively. Compounds up to a chain length of ten were, furthermore, characterized by quantitative determination of the products after degradation with snake venom phosphodiesterase (Table 5). For compounds with chain lengths higher than ten, paper chromatographic purification and characterization proved of doubtful value, due to their low mobility in the solvents that were available at this time for work in this field. Similarly, characterization of such long-chain compounds by nucleotide analysis after enzymic degradation seemed increasingly undesirable, because of the increasing amounts of valuable materials required for reasonably accurate analyses†.

† The more recent commercial availability of a nucleic acid analyzer, the Picker Nuclear apparatus, has enabled analyses on a very small scale, but the general argument made above about increasing uncertainty with increasing chain length is still valid.

TABLE 2  
*Spectral and chromatographic properties of fully protected polynucleotides*

| Compound                                                                                                                                                                                                                                                                                    | $\lambda_{\text{max}}^\dagger$ | $\epsilon_{280}/\epsilon_{260}^\dagger$ | $\epsilon_{280}/\epsilon_{302}^\dagger$ | $R_F$<br>Solvent B | $R_F$<br>Other solvents |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------|-----------------------------------------|-----------------------------------------|--------------------|-------------------------|
| d-pC <sup>An</sup> pC <sup>An</sup>                                                                                                                                                                                                                                                         | 302                            | 1.68                                    | 0.78                                    | 0.38               | 0.28 (A)<br>0.40 (F)    |
| d-pC <sup>An</sup> pC <sup>An</sup> .OAc                                                                                                                                                                                                                                                    | 302                            | 1.68                                    | 0.78                                    | 0.48               |                         |
| d-pC <sup>An</sup> pT                                                                                                                                                                                                                                                                       | 277,302                        | 1.14                                    | 1.04                                    | 0.39               |                         |
| d-pC <sup>An</sup> pT.OAc                                                                                                                                                                                                                                                                   | 277,302                        | 1.14                                    | 1.04                                    | 0.53               |                         |
| d-pG <sup>IB</sup> pA <sup>Bz</sup>                                                                                                                                                                                                                                                         | 280,260                        | 1.05                                    | 2.6                                     | 0.50               | 0.55 (G)                |
| d-pG <sup>IB</sup> pA <sup>Bz</sup> .OAc                                                                                                                                                                                                                                                    | 280,260                        | 1.05                                    | 2.6                                     | 0.66               |                         |
| d-pC <sup>An</sup> pA <sup>Bz</sup>                                                                                                                                                                                                                                                         | 285                            | 1.53                                    | 1.31                                    | 0.43               | 0.48 (G)                |
| d-pG <sup>IB</sup> pG <sup>IB</sup> pA <sup>Bz</sup>                                                                                                                                                                                                                                        | 259,280                        | 0.89                                    | 2.5                                     | 0.42               | 0.49 (G)                |
| d-pG <sup>IB</sup> pG <sup>IB</sup> pA <sup>Bz</sup> .OAc                                                                                                                                                                                                                                   | 259,280                        | 0.89                                    | 2.5                                     | 0.51               |                         |
| d-pC <sup>An</sup> pA <sup>Bz</sup> pT                                                                                                                                                                                                                                                      | 282                            | 1.36                                    | 1.51                                    | 0.23               | 0.41 (G)                |
| d-pG <sup>IB</sup> pA <sup>Bz</sup> pC <sup>An</sup> pT                                                                                                                                                                                                                                     | 279,265 (S)                    | 1.10                                    | 1.66                                    | 0.28               | 0.35 (G)                |
| d-pG <sup>IB</sup> pA <sup>Bz</sup> pC <sup>An</sup> pT.OAc                                                                                                                                                                                                                                 | 279,265 (S)                    | 1.10                                    | 1.66                                    | 0.33               |                         |
| d-pC <sup>An</sup> pTpC <sup>An</sup> pC <sup>An</sup>                                                                                                                                                                                                                                      | 301                            | 1.33                                    | 0.90                                    | 0.21               |                         |
| d-pC <sup>An</sup> pA <sup>Bz</sup> pTpG <sup>IB</sup>                                                                                                                                                                                                                                      | 280,263                        | 1.12                                    | 1.64                                    | 0.33               | 0.18 (G)                |
| d-pC <sup>An</sup> pA <sup>Bz</sup> pTpG <sup>IB</sup> .OAc                                                                                                                                                                                                                                 | 280,263                        | 1.12                                    | 1.64                                    | 0.39               |                         |
| d-MMTrG <sup>Bz</sup> pA <sup>Bz</sup>                                                                                                                                                                                                                                                      | 282,233†                       | 1.26†                                   | 1.80†                                   | 0.83               | 0.70 (A)                |
| d-MMTrG <sup>Bz</sup> pA <sup>Bz</sup> pA <sup>Bz</sup>                                                                                                                                                                                                                                     | 281,234                        | 1.37                                    | 2.1                                     | 0.80               | 0.64 (A)                |
| d-MMTrG <sup>Bz</sup> pA <sup>Bz</sup> pA <sup>Bz</sup> pC <sup>An</sup> pC <sup>An</sup>                                                                                                                                                                                                   | 283                            | 1.39                                    | 1.34                                    | 0.65               | 0.40 (A)                |
| d-MMTrG <sup>Bz</sup> pA <sup>Bz</sup> pA <sup>Bz</sup> pC <sup>An</sup> pC <sup>An</sup> pG <sup>IB</sup>                                                                                                                                                                                  | 282,265 (S)                    | 1.21                                    | 1.43                                    | 0.50               | 0.2 (A)                 |
| d-MMTrG <sup>Bz</sup> pA <sup>Bz</sup> pA <sup>Bz</sup> pC <sup>An</sup> pC <sup>An</sup> pG <sup>IB</sup> pG <sup>IB</sup> pA <sup>Bz</sup>                                                                                                                                                | 282,265 (S)                    | 1.18                                    | 1.50                                    | 0.3                |                         |
| d-MMTrG <sup>Bz</sup> pA <sup>Bz</sup> pA <sup>Bz</sup> pC <sup>An</sup> pC <sup>An</sup> pG <sup>IB</sup> pG <sup>IB</sup> pA <sup>Bz</sup> pG <sup>IB</sup> pA <sup>Bz</sup>                                                                                                              | 282,263                        | 1.16                                    | 1.67                                    |                    |                         |
| d-MMTrG <sup>Bz</sup> pA <sup>Bz</sup> pA <sup>Bz</sup> pC <sup>An</sup> pC <sup>An</sup> pG <sup>IB</sup> pG <sup>IB</sup> pA <sup>Bz</sup> pG <sup>IB</sup> pA <sup>Bz</sup> .pC <sup>An</sup> pT                                                                                         | 280,263                        | 1.15                                    | 1.58                                    |                    |                         |
| d-MMTrG <sup>Bz</sup> pA <sup>Bz</sup> pA <sup>Bz</sup> pC <sup>An</sup> pC <sup>An</sup> pG <sup>IB</sup> pG <sup>IB</sup> pA <sup>Bz</sup> pG <sup>IB</sup> pA <sup>Bz</sup> .pC <sup>An</sup> pTpC <sup>An</sup> pTpC <sup>An</sup> pC <sup>An</sup>                                     | 281,265 (S)                    | 1.15                                    | 1.31                                    |                    |                         |
| d-MMTrG <sup>Bz</sup> pA <sup>Bz</sup> pA <sup>Bz</sup> pC <sup>An</sup> pC <sup>An</sup> pG <sup>IB</sup> pG <sup>IB</sup> pA <sup>Bz</sup> pG <sup>IB</sup> pA <sup>Bz</sup> .pC <sup>An</sup> pTpC <sup>An</sup> pTpC <sup>An</sup> pC <sup>An</sup> pA <sup>Bz</sup> pTpG <sup>IB</sup> | 281,265 (S)                    | 1.12                                    | 1.23                                    |                    |                         |

† In water at neutral pH.

‡ In 50% ethanol.

TABLE 3

*Chromatographic properties of monomethoxytrityl polynucleotides*

| Compound                  | $R_F$     | $R_F$              |
|---------------------------|-----------|--------------------|
|                           | Solvent C | Other solvents     |
| d-MMTrGpA                 |           | 0.63 (A); 0.73 (B) |
| d-MMTrGpApA               | 0.79      | 0.34 (A); 0.46 (B) |
| d-MMTrGpApApCpC           | 0.59      |                    |
| d-MMTrGpApApCpCpG         | 0.44      |                    |
| d-MMTrGpApApCpCpGpGpA     | 0.25      |                    |
| d-MMTr-decanucleotide     | 0.16      |                    |
| d-MMTr-dodecanucleotide   | 0.08      |                    |
| d-MMTr-hexadecanucleotide | 0.04      |                    |

TABLE 4

*Spectral and chromatographic properties of unprotected polynucleotides*

| Compound               | $\lambda_{\max}^\dagger$ | $\epsilon_{280}/\epsilon_{260}^\dagger$ | $R_F$           | $R_F$           | $R_F$              |
|------------------------|--------------------------|-----------------------------------------|-----------------|-----------------|--------------------|
|                        |                          |                                         | Solvent C       | Solvent D       | Other solvents     |
| d-pCpC                 | 270                      | 1.03                                    | 0.32            | 0.42            | 0.14 (B)           |
| d-pCpT                 | 268                      | 0.82                                    | 0.23            | 0.38            |                    |
| d-pGpA                 | 256                      | 0.42                                    | 0.29            | 0.33            |                    |
| d-pCpA                 | 261                      | 0.35                                    | 0.35            | 0.45            |                    |
| d-pGpGpA               | 255                      | 0.46                                    | 0.13            | 0.10            |                    |
| d-pCpApT               | 262                      | 0.55                                    | 0.27            | 0.28            |                    |
| d-pGpApCpT             | 258                      | 0.55                                    | 0.17            | 0.13            |                    |
| d-pCpTpCpC             | 269                      | 0.87                                    | 0.18            | 0.20            |                    |
| d-pCpApTpG             | 258                      | 0.56                                    | 0.17            | 0.08            |                    |
| d-GpA                  | 256                      | 0.42                                    |                 | 0.51            | 0.35 (B); 0.16 (A) |
| d-GpApA                | 256                      | 0.44                                    | 0.35            |                 | 0.05 (A); 0.15 (B) |
| d-GpApApCpC            | 258                      | 0.55                                    | 0.20            | 0.32            |                    |
| d-GpApApCpCpG          | 255                      | 0.55                                    | 0.11            | 0.29            |                    |
| d-GpApApCpCpGpGpA      | 256                      | 0.52                                    | 0.04            | 0.17            |                    |
| d-GpApApCpCpGpGpApGpA  | 255                      | 0.49                                    | 0.02            | 0.04            |                    |
| The dodecanucleotide   | 257                      | 0.58                                    | 0.36 $\ddagger$ | 0.09 $\ddagger$ |                    |
| The hexadecanucleotide | 261                      | 0.64                                    |                 |                 |                    |
| The icosanucleotide    | 259                      | 0.56                                    |                 |                 |                    |

 $\dagger$  In water at neutral pH. $\ddagger$  Relative to the octanucleotide.

The introduction of DEAE-cellulose chromatography in the presence of 7 M-urea by Tomlinson & Tener (1962) has provided an extremely useful method for the separation of polynucleotides, as exemplified by the work of Holley *et al.* (1965) and others (RajBhandary *et al.*, 1967). This method was used in analysis of the unprotected dodeca-, hexadeca- and the icosanucleotides as described above. The same method has been adopted as a standard tool for characterization of the synthetic polydeoxynucleotide segments and has been used with confidence in all the work reported in the accompanying papers.

TABLE 5  
*Venom phosphodiesterase degradation of synthetic polynucleotides*

| Degraded compound      | Degradation products  | Ratio theory  | Ratio found           |
|------------------------|-----------------------|---------------|-----------------------|
| d-pGpGpA†              | dpG : dpA             | 2 : 1         | 2.0 : 1.0             |
| d-pGpApCpT†            | dpG : dpA : dpC : dpT | 1 : 1 : 1 : 1 | 1.1 : 1.0 : 1.0 : 1.0 |
| d-pCpTpCpC†            | dpC : dpT             | 3 : 1         | 3.0 : 0.7             |
| d-pCpApTpG†            | dpC : dpA : dpT : dpG | 1 : 1 : 1 : 1 | 1.0 : 1.0 : 0.9 : 1.0 |
| d-GpApApCpC†           | dG : dpA : dpC        | 1 : 2 : 2     | 0.8 : 2.0 : 2.1       |
| d-GpApApCpCpGpGpA†     | dG : dpA : dpC : dpG  | 1 : 3 : 2 : 2 | 0.9 : 3.0 : 2.0 : 1.9 |
| d-GpApApCpCpGpGpApGpA† | dG : dpA : dpC : dpG  | 1 : 4 : 2 : 3 | 1.0 : 4.0 : 2.0 : 3.1 |

† The degradation products were separated by paper chromatography in solvent D.

‡ The degradation products were first separated by paper chromatography in solvent A. The front half of this paper, containing dG, was then cut off, and the remaining portion of the paper submitted to paper chromatography in solvent D.

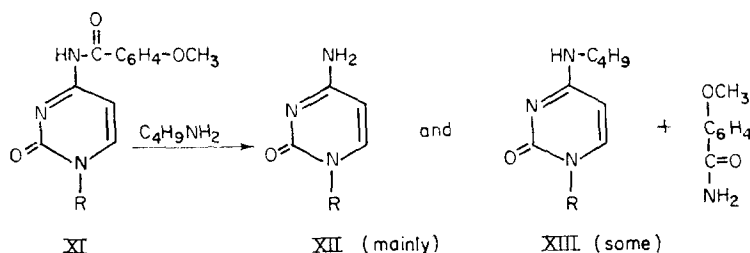
The icosanucleotide was characterized further by end-group analysis. The 5'-terminal nucleotide of the synthetic compound was labeled by phosphorylation using polynucleotide kinase and [ $\gamma$ - $^{32}\text{P}$ ]ATP (Richardson, 1965). After digestion of the polynucleotide with snake venom phosphodiesterase, the  $^{32}\text{P}$  label on paper chromatography travelled as a single spot coincident with the deoxyguanosine 5'-phosphate marker. Furthermore, degradation of the icosanucleotide carrying the  $^3\text{H}$  label at the 3'-terminus, by micrococcal nuclease and spleen phosphodiesterase, resulted in a single radioactive product identical with deoxyguanosine. These results confirmed that both the 5'- and 3'-termini of the synthetic material were deoxyguanosine residues. Furthermore, the fact that the degradation was complete in both experiments showed that all the nucleotidic material contained only the 3'  $\rightarrow$  5' internucleotide bonds. Enzymic joining experiments reported in the accompanying papers have, further, given abundant proof of the correct 3'- and 5'-termini of the icosanucleotide.

Finally, certain observations made during this and the subsequent work on the removal of the protecting groups may be mentioned. Due to the exceptional stability of the benzoyl protecting group on deoxyguanosine residue it proved necessary to give compounds containing this residue a more powerful base treatment than the usual treatment with concentrated ammonia at room temperature; treatment with an anhydrous *n*-butylamine/methanol mixture at room temperature was effective in quantitative removal of the protecting group (paper CVII). Upon such treatment, however, compounds containing deoxycytidine residues were invariably found to contain (by paper chromatography) a faster-moving impurity. The amount of this side-product depended on the number of deoxycytidine residues in the compound (Table 6). Characterization of these side products (see Experimental Methods section) showed that treatment of  $N_{(8)}$ -anisoyl-deoxycytidine compounds with *n*-butylamine probably results in a partial substitution of the  $N_{(8)}$ -amino function by the butylamino group (as illustrated by the partial structures (XI) to (XIII)). In order to avoid this side-product formation, conditions for the removal of the *N*-benzoyl protecting group were modified as described in the Experimental Methods section.

The removal of the monomethoxytrityl group has been commented on above and is further discussed in a subsequent paper(CXI).

TABLE 6  
*Side-products formed by n-butylamine treatment of compounds containing N-anisoyldeoxycytidine residues*

| Parent compound                                       | $R_F$ (solvent D) | Side-product<br>$R_F$ (other) | Relative amount<br>formed (%) |
|-------------------------------------------------------|-------------------|-------------------------------|-------------------------------|
| dpC <sup>An</sup>                                     | 0.76              | 0.35 (A)                      | 5.3                           |
| dpC <sup>An</sup> pC <sup>An</sup>                    | 0.70              | 0.53 (C)                      | 11.0                          |
| dpC <sup>An</sup> pTpC <sup>An</sup> pC <sup>An</sup> | 0.48              | 0.38 (C)                      | 13.0                          |



### 3. Experimental Methods

#### *Materials and general methods*

TPS and MS were obtained commercially (Aldrich Chemical Co., Milwaukee) and recrystallized twice from pentane. Pyridine was purified as described previously (Kossel, Büchi & Khorana, 1967). The protected mononucleotides, d-pA<sup>Bz</sup> and d-pC<sup>An</sup>, were prepared as described earlier (Khorana, Turner & Vizsolyi, 1961; Ralph & Khorana, 1961). d-pG<sup>13</sup>-OiB, d-p[8-<sup>3</sup>H]pG<sup>13</sup>-OiB and d-pG<sup>13</sup> were prepared as described elsewhere (paper CV). The following molar extinction values were used for the protected mononucleotides at neutral pH in aqueous solution: d-pC<sup>An</sup>, 22,500 (302 nm); d-pA<sup>Bz</sup>, 18,300 (280 nm); d-pG<sup>13</sup>, 16,700 (259 nm). Calculations of yields in condensation reactions are based on u.v. absorption spectra theoretically derived by summation of the spectra of the two starting components. Any hypochromicity was ignored in derivation of the spectra of all of the protected oligo- and polynucleotides.

Paper chromatography was carried out by the descending technique using Whatman no. 40 or Whatman no. 1 paper, unless otherwise indicated. The following solvent systems were used: solvent A, isopropanol/conc. ammonia/water (7 : 1 : 2, by vol.); solvent B, ethanol/1 M-ammonium acetate (pH 7.5) (7 : 3, v/v); solvent C, n-propanol/conc. ammonia/water (55 : 10 : 35, by vol.); solvent D, isobutyric acid/conc. ammonia/water (66 : 1 : 33, by vol.); solvent E, n-butanol/acetic acid/water (4 : 1 : 5 by vol.); solvent F, ethanol/1 M-ammonium acetate (pH 3.8) (7 : 3, v/v); solvent G, 1 M-ammonium bicarbonate/ethanol (9 : 1, v/v) (only for chromatography on DEAE-paper). The  $R_F$  values of different compounds are given in Tables 2, 3 and 4.

Paper electrophoresis was performed in a high-voltage apparatus in which the paper (usually Whatman no. 3MM) was immersed in a high boiling petroleum fraction (Varsol). Potassium phosphate buffer (0.03 M, pH 7.1) was used. The compounds were located on

paper chromatograms by examining the chromatograms under u.v. light. The presence of the monomethoxytrityl group in different compounds was detected by spraying the chromatograms with 10% aqueous perchloric acid and drying in warm air. The monomethoxytrityl-containing compounds appeared yellow. Enzymic degradations of the synthetic compounds using purified snake venom phosphodiesterase and spleen phosphodiesterase were performed as described previously (Khorana *et al.*, 1961). The results are given in Table 5.

Tritium was counted in a liquid-scintillation counter. The aqueous samples were added directly to a scintillation medium consisting of: dioxane (900 ml.), methoxyethanol (900 ml.), xylene (300 ml.), naphthalene (168 g), PPO (21 g) and POPOP (1.05 g).

Ultraviolet absorption spectra were taken in water at neutral pH, unless stated otherwise. The abbreviation O.D. unit refers to the extinction of a solution in a 1-cm light path quartz cell. The wavelength at which the absorbance measurements are made are indicated by the number in the suffix, e.g. O.D.<sub>260</sub> refers to extinction being measured at 260 nm.

#### *Acetylation of 3'-hydroxyl groups*

Acetylation of the 3'-hydroxyl groups of mononucleotides and oligonucleotide blocks was carried out by the following general procedure. A mixture of the dry nucleotidic compound (pyridinium or triethylammonium salt) and acetic anhydride (50-fold molar excess) in anhydrous pyridine (four times the volume of acetic anhydride) was kept at room temp. for 4 hr. Methanol (twice the volume of acetic anhydride used) and, after 1 hr at room temp., water (same volume as pyridine used) was added, and the mixture was left overnight at room temp. After addition of an equal amount of pyridine the solution was concentrated in a vacuum and rendered anhydrous by repeated evaporation of added dry pyridine. The concentrated pyridine solution was added dropwise to anhydrous ether (25 to 100 vol.). The precipitate was collected by centrifugation, washed three times with ether, and dried in a vacuum. It was stored in a desiccator over calcium chloride at  $-20^{\circ}\text{C}$  and reprecipitated from pyridine/ether just before use. The 3'-O-acetyl compounds were checked for purity by paper chromatography in solvent B. Yield was usually near quantitative.

#### *Cyanoethylation of 5'-phosphate end-groups*

Mono- and oligonucleotide intermediates which required protection of their 5'-phosphate group were cyanoethylated by the following general method. To an anhydrous mixture of the nucleotide component (as the pyridinium salt), cyanoethanol (50-fold molar excess), pyridinium Dowex-50 ion-exchange resin (2 to 5 g), pyridine (twice the volume of cyanoethanol) and DCC (tenfold molar excess) were added, and the suspension was shaken in the dark at room temp for 16 hr. Water (3 vol.) was added, the mixture was filtered, and the filtrate was extracted with cyclohexane or ether. After addition of pyridine (1 vol.) the aqueous solution was kept at room temp. overnight, finally concentrated in a vacuum and made anhydrous by repeated evaporation of added dry pyridine. The concentrated pyridine solution was added dropwise to anhydrous ether (25 to 100 vol.), and either a solid precipitate or an oily suspension resulted. The nucleotidic material was recovered by centrifugation, dissolved in dry pyridine and reprecipitated from ether. The second precipitation usually gave a dry, well-sedimenting precipitate, which was centrifuged, washed three times with ether and dried in a vacuum. It was stored in a desiccator over calcium chloride at  $-20^{\circ}\text{C}$  and reprecipitated from pyridine/ether just before use. Paper electrophoresis of the products showed invariably the absence of starting materials, but revealed the presence of several slower-moving compounds due to partial cyanoethylation of phosphodiester groups. This heterogeneity did not affect the use of the cyanoethylation products. Yields were 75 to 95%.

*General method of condensation and of chromatography. Using dicyclohexylcarbodiimide*

Several dinucleotide blocks were prepared using DCC as condensing reagent according to the following general procedure. To an anhydrous mixture of the pyridinium salts of the two suitably protected mononucleotide components (the compound carrying a free 5'-phosphate group being in excess) and a small amount of pyridinium Dowex-50 ion-exchange resin in pyridine, DCC (five to ten-fold molar excess over the 5'-phosphate component) was added and the suspension was shaken in the dark at room temp. for 3 to 5 days. Water (3 vol.) was added, the mixture was filtered and the filtrate was extracted with cyclohexane or ether. An equal volume of pyridine was added to the aqueous phase and the solution was kept at room temp. overnight. After cooling to 0°C, an equal volume of 2 M-sodium hydroxide was added and the mixture allowed to stand at room temp. for 25 min†. If the addition of sodium hydroxide did not result in a clear solution after 2 min, small amounts of ethanol were added. The solution was then poured into an excess of pyridinium Dowex-50 ion-exchange resin with cooling and shaking and after the pH had become neutral, the resin was removed by filtration. The resin was washed thoroughly with 20% aqueous pyridine, and the filtrate and washings were pooled and applied to a DEAE-cellulose anion-exchange column (bicarbonate form) pre-equilibrated with dilute TEAB buffer (pH 7.5) containing 10 to 40% ethanol. After application of the mixture, the column was washed with the equilibrating buffer until pyridine had been eluted completely. Elution of the nucleotidic compounds was then carried out using appropriate concentration gradients of salt and ethanol. To prevent the pH of the eluting buffer from rising during chromatography, conical flasks of identical shape were used for the mixing vessel and the reservoir, and a stream of carbon dioxide was continuously passed over the surface of the buffer in the mixing vessel. All column procedures were carried out at 4°C.

*General method of condensation**using tri-isopropyl benzenesulfonyl chloride or mesitylene-sulfonyl chloride*

Both reaction components (the compound carrying the free 5'-phosphate group being in excess) were dissolved in pyridine, and the mixture was rendered anhydrous by repeated evaporation of added dry pyridine in a vacuum, each time the flask being opened to the dry atmosphere of a dry box containing phosphorus pentoxide. During the last evaporation of pyridine, the minimal amount of the solvent necessary for complete solubilization of the reaction components was allowed to remain, and TPS or MS was then added inside the dry box. The reaction mixture was kept at room temp. for 1.5 to 4 hr with exclusion of moisture. It was then cooled in dry ice-ethanol mixture and di-isopropylethylamine (2 m-moles per m-mole of sulfonyl chloride used) was added as a 1 M-solution in pyridine, followed by water (volume equal to total volume of mixture). The solution was left at room temp. overnight. After cooling to 0°C, 2 M-sodium hydroxide (enough to give an over-all concentration of 1 M in the reaction mixture, taking into account all anions present) was added. In order to obtain a clear solution, small quantities of ethanol had to be added sometimes. In cases where only *O*-acetyl groups were to be hydrolyzed, the alkaline mixture was kept at 0°C for 10 min; for the removal of cyanoethyl groups the solution was left at room temp. for 25 min (paper CVIII). Neutralization and column chromatography on DEAE-cellulose were carried out as described for condensations using DCC. In order to achieve a satisfactory separation, the reaction mixture which was applied to the column had to be diluted enough to render the over-all anion molarity lower than 0.03 M. In most cases, the crude reaction mixtures were analyzed before column chromatography by paper chromatography using solvent B, or by chromatography on DEAE-paper using solvent G.

*Isolation of the protected oligonucleotides*

Appropriate fractions from columns containing the desired protected oligonucleotide were pooled and, after addition of an equal volume of pyridine, the solvent was removed by evaporation at room temp. in a vacuum. TEAB was also removed by continued evaporation of pyridine and the residue was rendered anhydrous by further evaporation of added dry

† More recent work by Dr Kumar (paper CVIII) showed that the removal of the cyanoethyl group was actually complete within 5 min in an ice bath.

pyridine. The concentrated pyridine solution was added to anhydrous ether (25 to 100 vol.), the precipitate centrifuged, washed three times with ether and dried in a vacuum. The dry compounds were stored in a desiccator over calcium chloride at  $-20^{\circ}\text{C}$ .

#### *Removal of protecting groups*

The *N*-acyl groups of protected oligonucleotides were removed by one of the following methods. (a) Compounds (usually portions of 20 to 40 o.d. units) containing *N*-benzoyldeoxyguanosine residues but no *N*-anisoyldeoxycytidine residues were treated, after complete evaporation of the solvent, with a 1 : 1 mixture (v/v) of *n*-butylamine and methanol (1 ml.) at room temp. for 2 days. (b) Compounds containing no *N*-benzoyldeoxyguanosine residues were treated with conc. ammonia (1 ml.) at room temp. for 2 days. (c) Compounds containing both *N*-benzoyldeoxyguanosine residues and *N*-anisoyldeoxycytidine residues were treated first with conc. ammonia (1 ml.) at room temp. for 2 days, then, after evaporation of the solvent in a vacuum, with a 1 : 1 mixture (v/v) of *n*-butylamine and methanol (1 ml.) at room temp. for 2 days. During evaporation of the conc. ammonia in a vacuum, pH was controlled by addition of a few drops of 1 M-TEAB buffer and frequent additions of pyridine. The compounds then were chromatographed in solvent C and, if they did not contain the monomethoxytrityl group, in solvent D. The monomethoxytrityl group was removed by treating the compound with 80% acetic acid for 45 min or with a mixture of acetic acid/pyridine/water, (14 : 1 : 3, by vol.), for 24 hr at room temp. After evaporation of the solvent in a vacuum at room temp., final purification of the compounds was achieved by paper chromatography in solvents C and D. Elution of the compounds from paper was carried out using 0.1 M-ammonia.

#### *N-Benzoyl-5'-O-mono-p-methoxytrityldeoxyguanosine, d-MMT<sup>+</sup>Gr<sup>Bz</sup>*

To a suspension of dry deoxyguanosine (10.0 g, 35 m-moles)<sup>†</sup> in anhydrous pyridine (150 ml.), benzoyl chloride (50 ml., 440 m-moles) was added with cooling, and the mixture was shaken for 1 hr at room temp. The resulting solution was poured into ice-water containing 5% pyridine (600 ml.) and the product extracted with chloroform (4 × 250 ml.). After back-washing with water (200 ml.), the chloroform layers were combined, filtered through anhydrous sodium sulfate and concentrated in a vacuum. The residue was dissolved in pyridine (200 ml.) and the solution treated with a mixture of ethanol and 2 M-sodium hydroxide (1 : 1, v/v; 800 ml.) at  $0^{\circ}\text{C}$  for 20 min. An excess of pyridinium Dowex-50 ion-exchange resin (500 ml.) was then added to neutralize the alkali, the resin was removed by filtration and washed twice with pyridine/ethanol (1 : 1, v/v, 300 ml.). The filtrate and washings were concentrated in a vacuum to a volume of about 250 ml. The residual solution was shaken in a round-bottom flask with ether (600 ml.) at  $0^{\circ}\text{C}$ , and the organic phase was syphoned off, the same operation being repeated 4 times. During the last operation, the product separated from the aqueous solution as a colorless, gelatinous precipitate. Traces of ether were removed by suction using a water pump, and the suspension, which had a pH of 7.5, was warmed to  $60^{\circ}\text{C}$ . After addition of ethanol (20 ml.), a clear solution resulted, which was cooled very slowly to  $0^{\circ}\text{C}$ . *N*-benzoyldeoxyguanosine separated as a crystalline precipitate and was collected by filtration. On heating, the dry crystals started to sinter above  $105^{\circ}\text{C}$ , but did not display a melting point up to  $250^{\circ}\text{C}$ , where slow decomposition occurred. On paper chromatography in solvents A ( $R_F = 0.54$ ) and E ( $R_F = 0.69$ ) the product was homogeneous except for faint tailing in both systems, probably due to some *N*-benzoylguanosine. The u.v. absorption characteristics of the compound are:  $\lambda_{\text{max}}$  297 nm ( $\epsilon = 12,100$ ), 265 nm ( $\epsilon = 11,800$ ), 258 nm ( $\epsilon = 11,000$ ), 238 nm ( $\epsilon = 14,000$ ) in ethanol;  $\lambda_{\text{max}}$  280 nm ( $\epsilon = 13,300$ ), 259 nm ( $\epsilon = 14,100$ ), 240 nm (shoulder,  $\epsilon = 12,200$ ) in water (pH 6.0). The yield was 9.75 g (75%).

A mixture of dry *N*-benzoyldeoxyguanosine (1.0 g, 2.7 m-moles) and mono-*p*-methoxytrityl chloride (1.2 g, 3.9 m-moles) in anhydrous pyridine (8 ml.) was stored at room temp. overnight. After addition of ethanol (1 ml.), the solution was further kept for 1 hr at room temp, then concentrated in a vacuum. The residue was poured on ice-water (50 ml.) and extracted with chloroform (2 × 100 ml.). After back-washing with ice-water (50 ml.), the organic phases were combined, filtered through anhydrous sodium sulfate and concentrated.

<sup>†</sup> The sample contained 3.3% guanosine as an impurity.

The residue was made anhydrous by repeated evaporation of added small portions of a mixture of benzene and chloroform in a vacuum and finally taken up in chloroform (10 ml.). Thin-layer chromatography (Silicagel Merck HF254; solvent: chloroform/ethanol, 5:1) at this stage showed two major  $\text{HClO}_4$ -positive spots at  $R_F = 0.8$  (mono-*p*-methoxytrityl-ethyl ether) and 0.6 (*N*-benzoyl-5'-*O*-mono-*p*-methoxytrityldeoxyguanosine), a faint  $\text{HClO}_4$ -positive spot at  $R_F = 0.4$ , as well as traces of  $\text{HClO}_4$ -positive and -negative, u.v.-absorbing materials near the origin. The above chloroform solution was applied to a Silicagel column (4 cm  $\times$  40 cm, Merck Silicagel H) and the elution carried out with chloroform containing 7% ethanol. The first 150 ml. of the eluate were discarded, then 15-ml. fractions were collected. The desired compound could be detected on the column as a slowly migrating, turbid band. Thin-layer chromatography showed fractions 56 to 65 (830 mg) to contain the pure product, whereas fractions 66 to 81 (590 mg) had to be recycled, since they contained increasing amounts of the slower-moving ( $R_F = 0.4$ ) impurity. (Analysis by paper chromatography of this impurity, after removal of the protecting groups, in solvent A, showed it to be the guanosine derivative.) The fractions containing pure product were pooled and evaporated in a vacuum to dryness, the residue was dissolved in chloroform (15 ml., a small amount of ethanol had sometimes to be added to improve solubility), and the solution was precipitated from pentane (300 ml.). Centrifugation, washing (twice) and drying in a vacuum yielded pure *N*-benzoyl-5'-*O*-mono-*p*-methoxytrityldeoxyguanosine as a fine, white powder with m.p. 145 to 155°C. The compound gave a single spot on paper chromatography in solvent A ( $R_F = 0.82$ ) and B ( $R_F = 0.30$ ). The u.v. absorption characteristics were:  $\lambda_{\text{max}}$  298 nm ( $\epsilon = 13,300$ ), 265 nm ( $\epsilon = 13,400$ ), 233 nm ( $\epsilon = 27,800$ ), shoulders at 290 nm and 255 nm (in ethanol). A portion of the product was deprotected by successive butylamine and acetic acid treatments. Paper chromatography in solvents A and E showed the resulting deoxyguanosine to be virtually pure and to contain less than 1% of contaminating guanosine. The average yield of pure d-MMTrG<sup>Bz</sup> was 70%.

*The protected trinucleotide, d-MMTrG<sup>Bz</sup>pA<sup>Bz</sup>pA<sup>Bz</sup>*

An anhydrous mixture of *N*-benzoyl-5'-*O*-mono-*p*-methoxytrityldeoxyguanosine, d-MMTrG<sup>Bz</sup> (2.1 m-moles), pyridinium *N*-benzoyl-3'-*O*-acetyldeoxyadenosine-5'-phosphate, dpA<sup>Bz</sup>-OAc (4.5 m-moles), pyridinium Dowex-50 ion-exchange resin (8 g) and DCC (4.15 g, 20 m-moles) in pyridine (15 ml.) was shaken in the dark at room temp. for 110 hr. Water (20 ml.) was added and the suspension was filtered, the residue being washed with 50% aqueous pyridine (two 5-ml. portions). The filtrate was extracted with cyclohexane (2  $\times$  100 ml.), pyridine (15 ml.) was added to the aqueous phase, and the solution was left at room temp. overnight. The mixture was then made to 100 ml. with 50% aqueous pyridine and extracted with ether (one 100-ml. and two 50-ml. portions), the combined ether extracts being back-washed with water (10 ml.). The aqueous phases were pooled and, after addition of an equal volume of pyridine, concentrated to dryness. The residue was dissolved in chloroform (100 ml.), and the solution was shaken gently with water (20 ml.). The aqueous layer was washed with chloroform (two portions of 100 ml.), and the combined organic phases were back-washed with water (10 ml.). After the addition of pyridine (100 ml.), the chloroform solution was concentrated in a vacuum to a vol. of approx. 30 ml. As judged by thin-layer chromatography, this preparation of the protected dinucleoside phosphate (d-MMTr-G<sup>Bz</sup>pA<sup>Bz</sup>) contained only minute amounts of the protected nucleoside d-MMTrG<sup>Bz</sup>, whereas paper chromatography in solvent B showed the absence of the protected mononucleotide, dpA<sup>Bz</sup>-OAc. Paper chromatography in solvent B of the combined aqueous phases revealed that they contained practically pure dpA<sup>Bz</sup>-OAc, whereas the ether fraction contained some protected nucleoside, d-MMTrG<sup>Bz</sup>, along with small quantities of unidentified fast-travelling side products. Paper chromatography in solvents A and B, of all three fractions after removal of the protecting groups, confirmed this distribution of the compounds. The solution of the protected dinucleoside phosphate in pyridine (30 ml.) was diluted with water to 45 ml., cooled to 0°C, and cold 2 M-sodium hydroxide (50 ml.) was added. Small amounts of water and pyridine were added in order to obtain a homogeneous reaction mixture, which was kept at 0°C for 5 min. An excess of pyridinium Dowex-50 ion-exchange resin (80 ml.) was added with cooling and shaking, and

after neutralization the resin was removed by filtration and washed thoroughly with 50% pyridine. Filtrate and washings were concentrated and the product was isolated as usual. This preparation was used for the next step without further purification. The yield was 68,000 o.d.<sub>280</sub> units.

A mixture of pyridinium d-MMTrG<sup>Bz</sup>pA<sup>Bz</sup> (60,500 o.d.<sub>280</sub> units) and pyridinium dpA<sup>Bz</sup>-OAc (140,000 o.d.<sub>280</sub> units, 7.7 m-moles) was allowed to react with TPS (8.0 g, 26.4 m-moles) in pyridine (50 ml.) with exclusion of moisture for 6 hr at room temp. The solution was cooled in dry ice-ethanol mixture, and di-isopropylethylamine (54 m-moles) in pyridine (54 ml.) was added, followed by water (100 ml.). The mixture was left at room temp. overnight, then concentrated to a vol. of 40 ml. Water (200 ml.) was added, and the solution was extracted with ethyl acetate (2 × 200 ml.), the organic phases being back-washed with water (50 ml.). The pooled aqueous phases (pH 6.8) were extracted with chloroform (4 × 400 ml.), the organic extracts combined without back-washing and, after addition of pyridine (200 ml.), concentrated in a vacuum to 100 ml. Paper chromatography in solvent B showed the chloroform fractions to contain mainly the desired trinucleotide, with traces of dinucleotide and mononucleotide present. The ethyl acetate fraction contained mainly tri-isopropyl benzene sulfonic acid, as well as some fast-moving side products, whereas the aqueous phase consisted mainly of the mononucleotide with only traces of di- and trinucleotide being present. The same pattern of distribution was obtained by paper chromatography of the three fractions in solvents A and B after removal of the protecting groups. To the solution of the crude protected trinucleotide in pyridine (100 ml.), cold 2 M-sodium hydroxide (200 ml.) was added at 0°C, followed by ice-water (80 ml.) and cold ethanol (30 ml.), whereupon a clear solution resulted. After 10 min at 0°C, pyridinium Dowex-50 ion-exchange resin (500 ml.) was added with cooling and the suspension was shaken vigorously. The resin was removed by filtration and washed thoroughly with 25% aqueous pyridine, and the combined filtrate and washings (total vol. 3 l.) were applied to a column (4 cm × 120 cm) of DEAE-cellulose (bicarbonate) pre-equilibrated with 0.01 M-TEAB in 30% ethanol at 4°C. After application of the nucleotide mixture, the column was washed with the same buffer (1.9 l.). Elution was started using a linear salt concentration gradient with 30% ethanol (3 l.) in the reservoir. Elution was then continued with 0.15 M-TEAB in 30% ethanol (3 l.) in the mixing vessel and 0.35 M-TEAB in 50% ethanol (3 l.) in the reservoir, followed by 0.35 M-TEAB in 50% ethanol (3 l.) in the mixing vessel and 0.45 M-TEAB in 50% ethanol (3 l.) in the reservoir. Various fractions of the resulting main peak (eluting salt concentration around 0.34 M) were examined by paper chromatography in solvent A, and the fractions containing the pure trinucleotide were pooled and concentrated at room temp. in a vacuum, with added pyridine. Fractions from the edges of the main peak containing a considerable amount of impure product were concentrated separately and rechromatographed. For this, the residue was dissolved in 50% pyridine (100 ml.) and applied to the same DEAE-cellulose column, after it had been washed with 0.5 M-TEAB in 50% ethanol (5 l.) and equilibrated with 0.20 M-TEAB in 35% ethanol. Elution was carried out using a constant buffer concentration identical to that used for equilibration of the column. The u.v.-absorbing material was eluted as in a broad peak, and fractions containing the pure trinucleotide were pooled and concentrated. The concentrate was combined with the bulk of the pure material and the trinucleotide was isolated by precipitation from pyridine/ether. The product was homogeneous on paper chromatography in solvents B and C after *n*-butylamine treatment and after subsequent acetic acid treatment. The yield was 42,000 o.d.<sub>280</sub> units (0.86 m-mole, 46% calculated from pure d-MMTrG<sup>Bz</sup>).

#### *The dinucleotide, dpC<sup>An</sup>pC<sup>An</sup>*

To an anhydrous pyridine solution (40 ml.) of pyridinium d-CE-pC<sup>An</sup> (90,000 o.d.<sub>302</sub> units, 4.0 m-moles) and pyridinium dpC<sup>An</sup>-OAc (114,000 o.d.<sub>302</sub> units, 5.1 m-moles) dry pyridinium Dowex-50 ion-exchange resin (12 g) and DCC (8.3 g, 40 m-moles) were added, and the suspension was shaken in the dark for 90 hr at room temp. After work-up, the reaction products were separated by chromatography on a DEAE-cellulose column (4 cm × 120 cm) pre-equilibrated with 0.01 M-TEAB in 10% ethanol. Elution was performed using a linear salt concentration gradient from 0.10 to 0.25 M-TEAB in 10% ethanol (total vol. 16 l.). The dinucleotide was eluted at a salt concentration of about

0.22 M. The product was isolated as usual. It was homogeneous on paper chromatography both before (solvent F) and after removal of protecting groups (solvents B and C). The yield was 103,000 o.d.<sub>302</sub> units (2.3 m-moles, 57%).

*The dinucleotide, dpC<sup>An</sup>pT*

An anhydrous mixture of pyridinium d-CE-pC<sup>An</sup> (153,000 o.d.<sub>302</sub> units, 7.6 m-moles), pyridinium dpT-OAc (101,000 o.d.<sub>267</sub> units, 10.6 m-moles), pyridinium Dowex-50 ion-exchange resin (3 g), and DCC (12 g, 58 m-moles) in pyridine (40 ml.) was shaken in the dark at room temp. for 75 hr. After the usual work-up, DEAE-cellulose chromatography was carried out using a column (6 cm × 110 cm) pre-equilibrated with 0.05 M-TEAB in 20% ethanol, and for elution a linear concentration gradient was applied from 0.05 to 0.30 M-TEAB in 20% ethanol (total volume 20 l.). The dinucleotide was eluted at a salt concentration of approximately 0.2 M. The product, after the usual isolation procedure, was homogeneous on paper chromatography in solvent B and, after removal of the protecting groups, in solvents C and D. The yield was 108,000 o.d.<sub>277</sub> units (4.6 m-moles, 61%).

*The dinucleotide, dpG<sup>iB</sup>pA<sup>Bz</sup>*

Pyridinium d-CE-pG<sup>iB</sup> (139,000 o.d.<sub>259</sub> units, 8.4 m-moles), and pyridinium dpA<sup>Bz</sup>-OAc (171,000 o.d.<sub>280</sub> units, 9.4 m-moles) were allowed to react with TPS (8.0 g, 26 m-moles) in pyridine (30 ml.) with exclusion of moisture for 4 hr at room temp. After the usual work-up the reaction mixture was applied to a DEAE-cellulose column (6 cm × 110 cm) pre-equilibrated with 0.05 M-TEAB in 20% ethanol. For elution a linear salt gradient from 0.05 to 0.20 M-TEAB in 20% ethanol was used (total volume 18 l.). The dinucleotide peak was eluted at a salt concentration of approximately 0.14 M. The product was isolated as usual. Paper chromatography in solvent B showed a strong main spot and several trace impurities ( $R_F = 0.23, 0.34$  and  $0.66$ ). Removal of protecting groups and paper chromatography in solvents C and D showed traces of u.v.-absorbing materials ( $R_F = 0.21, 0.37$  and  $0.51$  for solvent C;  $0.15, 0.24, 0.51$  and  $0.73$  in solvent D) in addition to the main product. Since dpGpA was not separable from dpG in any of the systems tested, the dinucleotide spot was eluted from paper and treated with alkaline phosphatase. The resulting dephosphorylated material travelled as a single spot, indistinguishable from authentic dGpA (solvent B). The yield of dpG<sup>iB</sup>pA<sup>Bz</sup> was 111,000 o.d.<sub>280</sub> units (3.7 m-moles, 44%).

*The dinucleotide, dpC<sup>An</sup>pA<sup>Bz</sup>*

An anhydrous pyridine solution (30 ml.) of pyridinium d-CE-pC<sup>An</sup> (113,000 o.d.<sub>302</sub> units, 5.0 m-moles) and pyridinium dpA<sup>Bz</sup>-OAc (135,000 o.d.<sub>280</sub> units, 7.4 m-moles) was treated with TPS (6.1 g, 20 m-moles) at room temp. for 2.25 hr with exclusion of moisture. After the usual work-up, the reaction products were applied to a DEAE-cellulose column (4 cm × 90 cm) pre-equilibrated with 0.05 M-TEAB in 20% ethanol. Elution was carried out using a linear salt gradient from 0.05 to 0.25 M-TEAB in 20% ethanol (12 l.), followed by a wash with 0.25 M-TEAB in 20% ethanol (2 l.). The product was eluted at a salt concentration of approximately 0.23 M. Paper chromatography of various fractions of the dinucleotide peak in solvent B showed the presence of considerable quantities of side products at the edges of the peak. The fractions containing pure product, from the central portion of the peak, were pooled and the dinucleotide isolated as usual. Removal of the protecting groups resulted in a material travelling as a single spot in solvents C and D. The yield of the protected dinucleotide was 83,000 o.d.<sub>285</sub> units (2.3 m-moles, 45%).

*The trinucleotide, dpG<sup>iB</sup>pG<sup>iB</sup>pA<sup>Bz</sup>*

An anhydrous mixture of pyridinium d-CE-pG<sup>iB</sup> (42,500 o.d.<sub>259</sub> units, 2.5 m-moles) and triethylammonium dpG<sup>iB</sup>pA<sup>Bz</sup>-OAc (86,000 o.d.<sub>280</sub> units, 2.9 m-moles) in pyridine (15 ml.) was treated with TPS (4.0 g, 13 m-moles) at room temp. for 3 hr 20 min with exclusion of moisture. After work-up, the reaction mixture was applied to a column (4 cm × 110 cm) of DEAE-cellulose pre-equilibrated with 0.5 M-TEAB in 40% ethanol. For elution, a linear salt gradient from 0.05 to 0.25 M-TEAB in 40% ethanol (total 24 l.) was used. The trinucleotide was eluted at a salt concentration of approximately 0.16 M. Paper chromatography in solvent B of several peak fractions showed that both the recovered dinucleotide,

dpG<sup>18</sup>pA<sup>Bz</sup> (46,000 o.d.<sub>280</sub> units, 53%), and the trinucleotide, dpG<sup>18</sup>pG<sup>18</sup>pA<sup>Bz</sup> (40,000 o.d.<sub>280</sub> units, 36%), were contaminated by considerable amounts of slower travelling side products. Fractions containing the trinucleotide were pooled, concentrated and rechromatographed on a DEAE-cellulose column of the same size, pre-equilibrated with 0.05 M-TEAB in 10% ethanol. For elution, a linear gradient was applied from 0.05 to 0.125 M-TEAB in 10% ethanol (4 l.), followed by a gradient from 0.125 to 0.20 M-TEAB in 10% ethanol (12 l.). The second half of the resulting peak consisted of practically pure product; the first half, however, still contained a certain amount of a side product ( $R_F = 0.24$ , solvent B). The contaminant, after ammonia treatment, was indistinguishable from dpGpA on paper chromatography in solvents C and D. It was therefore assumed to be dpGpA<sup>Bz</sup>, formed by partial hydrolysis of the *N*-isobutyryl group during alkaline treatment in work-up. This side product was finally removed by preparative paper chromatography in solvent B using Whatman no. 31 paper (acid washed, extra thick). Fractions containing the pure trinucleotide were pooled and the compound was isolated by the usual method. It was homogeneous on paper chromatography in solvent B and, after ammonia treatment, in solvents C and D. Final yield of the pure product was 29,000 o.d.<sub>280</sub> units (0.63 m-mole, 25%). The recovered impure dinucleotide, dpG<sup>18</sup>pA<sup>Bz</sup>, was combined with other fractions containing the same compound and rechromatographed on a DEAE-cellulose column.

*The trinucleotide, dpC<sup>An</sup>pA<sup>Bz</sup>pT*

To an anhydrous pyridine solution (10 ml.) of pyridinium d-CE-pC<sup>An</sup>pA<sup>Bz</sup> (60,000 o.d.<sub>285</sub> units, 1.6 m-moles) and pyridinium dpT-OAc (60,000 o.d.<sub>287</sub> units, 6.3 m-moles) TPS 3.6 g, 11.9 m-moles) was added and the mixture was kept at room temp. for 2 hr 50 min. After the work-up, the reaction products were applied to a DEAE-cellulose column (4 cm × 90 cm) pre-equilibrated with 0.05 M-TEAB in 40% ethanol. The linear concentration gradient used for elution was from 0.05 to 0.35 M-TEAB (40% ethanol, total vol. 18 l.). The product was eluted at a salt concentration of approx. 0.25 M. By paper chromatography in solvent B, the first part of the trinucleotide peak was shown to contain two contaminants ( $R_F = 0.13$  and 0.32). Fractions containing pure dpC<sup>An</sup>pA<sup>Bz</sup>pT were pooled and the compound isolated as usual. It was homogeneous on paper chromatography in solvent B and, after removal of protecting groups, in solvents C and D. The yield was 31,000 o.d.<sub>280</sub> units (0.75 m-mole, 46%).

*The tetranucleotide, dpG<sup>18</sup>pA<sup>Bz</sup>pC<sup>An</sup>pT*

An anhydrous mixture of pyridinium d-CE-pG<sup>18</sup>pA<sup>Bz</sup> (53,000 o.d.<sub>280</sub> units, 1.8 m-moles), triethylammonium dpC<sup>An</sup>pT-OAc (52,000 o.d.<sub>277</sub> units, 2.1 m-moles) and TPS (2.93 g, 9.7 m-moles) in pyridine (20 ml.) was kept at room temp for 3.25 hr. The usual work-up was followed by chromatography of the reaction products on a DEAE-cellulose column as shown in Fig. 3. The edges of the main peak (peak III) containing the tetranucleotide were contaminated by appreciable amounts of side products, as shown by removal of protecting groups and paper chromatography in solvents C and D. The fractions containing pure material were pooled and the tetranucleotide isolated. The product travelled as a single spot on paper chromatography in solvent B. The yield was 36,000 o.d.<sub>280</sub> units (0.66 m-mole, 37%). Furthermore, 6200 o.d.<sub>280</sub> units of dpG<sup>18</sup>pA<sup>Bz</sup> (0.2 m-mole, 12%) and 17,300 o.d.<sub>277</sub> units of dpC<sup>An</sup>pT (0.7 m-mole, 33%) could be recovered.

*The tetranucleotide, dpC<sup>An</sup>pTpC<sup>An</sup>pC<sup>An</sup>*

Pyridinium d-CE-pC<sup>An</sup>pT (32,000 o.d.<sub>277</sub> units, 1.3 m-moles) and triethylammonium dpC<sup>An</sup>pC<sup>An</sup>-OAc (87,000 o.d.<sub>302</sub> units, 1.93 m-moles) were allowed to react with TPS (2.58 g, 8.4 m-moles) in pyridine (15 ml.) for 3.75 hr at room temp. with exclusion of moisture. After the usual work-up, the reaction products were chromatographed on a DEAE-cellulose column as shown in Fig. 4. Paper chromatography in solvent B of various fractions of the tetranucleotide peak (peak III) indicated it to be pure except for the tail section of the peak. The pure tetranucleotide was isolated as usual. After removal of protecting groups, the compound was homogeneous on paper chromatography in solvents C and D. The yield was 40,000 o.d.<sub>302</sub> units (0.55 m-mole, 42%). Recovered dpC<sup>An</sup>pT amounted to 2100 o.d.<sub>277</sub> units (0.08 m-mole, 6.5%), and recovery of dpC<sup>An</sup>pC<sup>An</sup> was 23,000 o.d.<sub>302</sub> units (0.51 m-mole, 26%).

*The tetranucleotide, dpC<sup>An</sup>pA<sup>Bz</sup>pTpG<sup>iB</sup>*

To an anhydrous solution of pyridinium d-CE-pC<sup>An</sup>pA<sup>Bz</sup>pT (27,500 o.d.<sub>280</sub> units, 0.66 m-mole) and pyridinium dpG<sup>iB</sup>-OiB (31,200 o.d.<sub>259</sub> units, 1.87 m-moles) in pyridine (10 ml.), MS (1.11 g, 5.04 m-moles) was added and the reaction mixture was kept at room temp. for 1 hr 40 min. The usual work-up was followed by chromatography on DEAE-cellulose using a column (4 cm × 90 cm) pre-equilibrated with 0.05 M-TEAB in 20% ethanol. Elution was started using a linear salt gradient from 0.05 to 0.20 M-TEAB in 20% ethanol (3.3 l.), followed by a gradient from 0.20 to 0.35 M-TEAB in 20% ethanol (12 l.). The tetranucleotide was eluted at a salt concentration of approx. 0.32 M. Material on both sides of the tetranucleotide peak was shown by paper chromatography (solvent B) to contain side products. The fractions from the central portion of the peak were pooled and the product isolated as usual. The material was homogeneous on paper chromatography in solvents C and D after removal of protecting groups, except for a trace of a faster-travelling side product ( $R_F$  = 0.28, solvent C; 0.24, solvent D). The yield was 11,200 o.d.<sub>280</sub> units (0.21 m-mole, 34%).

*The <sup>3</sup>H-labeled tetranucleotide, dpC<sup>An</sup>pA<sup>Bz</sup>pTpG<sup>iB</sup>*

An anhydrous pyridine solution (0.5 ml.) of pyridinium d-CE-pC<sup>An</sup>pA<sup>Bz</sup>pT (1,570 o.d.<sub>280</sub> units, 37.5 μmoles), pyridinium d-p[8-<sup>3</sup>H]G<sup>iB</sup>-OiB (1125 o.d.<sub>259</sub> units, 68 μmoles, specific activity =  $8.5 \times 10^6$  cts/min/μmole), and MS (54.4 mg, 248.5 μmoles) was kept at room temp. for 2 hr with exclusion of moisture. After work-up, the reaction products were chromatographed on a DEAE-cellulose column as shown in Figure 5. Samples from several fractions of the resulting product peak (peak III) were treated with conc. ammonia and submitted to paper chromatography in solvents C and D. Scanning of the radioactive spots showed that only a central portion of the peak was free from contaminants. The fractions containing pure product were pooled and concentrated. The yield of the pure tetranucleotide was 612 o.d.<sub>280</sub> units (11.5 μmoles, 31%). The specific activity was  $6.8 \times 10^6$  cts/min/μmole. The preparation was combined with 11,700 o.d.<sub>280</sub> units of the non-radioactive product (20.2-fold dilution). The specific radioactivity of the diluted material was  $3.85 \times 10^5$  cts/min/μmole. Digestion of sample of this preparation by spleen phosphodiesterase (after removal of protecting groups and dephosphorylation by alkaline phosphatase) resulted in complete degradation of the tetranucleotide to mononucleotides and radioactive deoxyguanosine, as shown by paper chromatography in solvent A. Scanning of the paper chromatogram showed all the radioactivity to be located at the deoxyguanosine spot.

*The pentanucleotide, d-MMTTrG<sup>Bz</sup>pA<sup>Bz</sup>pA<sup>Bz</sup>pC<sup>An</sup>pC<sup>An</sup>*

Triethylammonium d-MMT-G<sup>Bz</sup>pA<sup>Bz</sup>pA<sup>Bz</sup> (30,000 o.d.<sub>280</sub> units, 0.62 m-mole) was condensed with triethylammonium dpC<sup>An</sup>pC<sup>An</sup>-OAc (100,000 o.d.<sub>302</sub> units, 2.2 m-moles) using TPS (2.0 g, 6.6 m-moles) in anhydrous pyridine (15 ml.). After a 5-hr reaction period at room temp., the mixture was worked up as usual. Column chromatography of the reaction products was carried out as shown in Fig. 6. Various fractions of the pentanucleotide peak (peak VI) were checked for purity by paper chromatography using solvent A. The fractions containing pure material were pooled and the product isolated as usual. The pentanucleotide was homogeneous on paper chromatography in solvents A and B. After removal of the N-protecting groups, the compound gave a single spot on paper chromatography in solvent C. Removal of the monomethoxytrityl group resulted in the free pentanucleotide, which was shown to be homogeneous on paper chromatography in solvents C and D. The yield was 35,700 o.d.<sub>284</sub> units (0.41 m-mole, 67%). Essentially pure dpC<sup>An</sup>pC<sup>An</sup> (30,000 o.d.<sub>302</sub> units, 0.66 m-mole, 30%) was also recovered.

*The octanucleotide, d-MMTTrG<sup>Bz</sup>pA<sup>Bz</sup>pA<sup>Bz</sup>pC<sup>An</sup>pC<sup>An</sup>pG<sup>iB</sup>pG<sup>iB</sup>pA<sup>Bz</sup>**Method A (via the hexanucleotide)*

An anhydrous pyridine solution (3 ml.) of triethylammonium d-MMTTrG<sup>Bz</sup>pA<sup>Bz</sup>pA<sup>Bz</sup>-pC<sup>An</sup>pC<sup>An</sup> (3050 o.d.<sub>284</sub> units, 35.6 μmoles) and pyridinium dpG<sup>iB</sup>-OiB (10,100 o.d.<sub>259</sub> units, 605 μmoles) was treated with TPS (550 mg, 1.8 m-moles) at room temp for 2-25 hr. After the usual work-up, the reaction products were applied to a DEAE-cellulose column

and chromatographed as shown in Fig. 7. The hexanucleotide was eluted as a broad peak (III), and the material gave a long,  $\text{HClO}_4$ -positive spot on paper chromatography in solvent A. Removal of the protecting groups followed by paper chromatography in solvents C and D showed the compound to be essentially pure, except for some traces of faster-moving material. The yield was 2310 o.d.<sub>284</sub> units (23.8 m-moles, 67%).

Triethylammonium d-MMTrG<sup>Bz</sup>pA<sup>Bz</sup>pA<sup>Bz</sup>pC<sup>An</sup>pC<sup>An</sup>pG<sup>IB</sup> (1860 o.d.<sub>284</sub> units, 18.1  $\mu$ moles) was condensed with triethylammonium dpG<sup>IB</sup>pA<sup>Bz</sup>-OAc (6100 o.d.<sub>280</sub> units, 203  $\mu$ moles) using TPS (180 mg, 590  $\mu$ moles) in anhydrous pyridine (1.5 ml.). After a reaction period of 2.5 hr at room temp., the mixture was worked up as usual. The column separation was carried out as shown in Fig. 6. The fractions containing the pure octanucleotide (peak IV) were pooled and the product was isolated as usual. Analysis by paper chromatography was carried out after removal of the *N*-protecting groups (solvent C) as well as after subsequent hydrolysis of the monomethoxytrityl group (solvents C and D). The compound proved to be homogeneous except for faint traces of faster-moving compounds. The yield of the pooled fractions was 900 o.d.<sub>280</sub> units (7.1  $\mu$ moles, 39%). 460 o.d.<sub>284</sub> units of the hexanucleotide (4.7  $\mu$ moles, 26%) and 3700 o.d.<sub>280</sub> units of dpG<sup>IB</sup>pA<sup>Bz</sup> (123  $\mu$ moles, 61%) were recovered.

### Method B

An anhydrous pyridine solution (10 ml.) of triethylammonium d-MMTrG<sup>Bz</sup>pA<sup>Bz</sup>pA<sup>Bz</sup>pC<sup>An</sup>pC<sup>An</sup> (15,400 o.d.<sub>284</sub> units, 0.178 m-mole), triethylammonium dpG<sup>IB</sup>pG<sup>IB</sup>pA<sup>Bz</sup>-OAc (20,000 o.d.<sub>280</sub> units, 0.46 m-mole) and TPS (930 mg, 3.1 m-moles) was kept at room temp. for 3 hr 20 min. The usual work-up was followed by column chromatography as shown in Fig. 9. The fractions containing peak IV were pooled and the octanucleotide was isolated as usual. After removal of the *N*-acyl groups, the compound was homogeneous on paper chromatography in solvent C, except for some faint tailing and a diffuse trace of fast-moving material. The same pattern was found on paper chromatography in solvents C and D after the removal of the monomethoxytrityl group from the *N*-deacylated compound. The paper chromatographic behavior of this octanucleotide preparation was indistinguishable from that obtained by method A. The yield was 5200 o.d.<sub>280</sub> units (41  $\mu$ moles, 23%). In addition, 11,400 o.d.<sub>280</sub> units of the trinucleotide (0.25 m-mole, 55%), and 6000 o.d.<sub>280</sub> units of the pentanucleotide (69.5  $\mu$ moles, 39%) were recovered.

#### *The decanucleotide, d-MMTrG<sup>Bz</sup>pA<sup>Bz</sup>pA<sup>Bz</sup>pC<sup>An</sup>pC<sup>An</sup>pG<sup>IB</sup>pG<sup>IB</sup>pA<sup>Bz</sup>pG<sup>IB</sup>pA<sup>Bz</sup>*

A mixture of the protected octanucleotide (triethylammonium salt, 770 o.d.<sub>280</sub> units, 6.06  $\mu$ moles) and triethylammonium dpG<sup>IB</sup>pA<sup>Bz</sup>-OAc (4710 o.d.<sub>280</sub> units, 157  $\mu$ moles) in anhydrous pyridine (1 ml.) was treated with TPS (180 mg, 590  $\mu$ moles) for 3 hr at room temp. with exclusion of moisture. After work-up the reaction products were separated by chromatography on a DEAE-cellulose column as shown in Fig. 10. The fractions containing peak IV were pooled and the decanucleotide was isolated as usual. Paper chromatography of the sample in solvent C, after removal of the *N*-acyl groups, showed a single spot, with faint traces of faster-moving material (mainly  $R_F = 0.25$  to 0.3). Removal of the monomethoxytrityl group by paper chromatography in solvents C and D again showed the compound to be essentially homogeneous. The yield was 515 o.d.<sub>280</sub> units (3.3  $\mu$ moles, 55%). 165 o.d.<sub>280</sub> units of the octanucleotide (1.3  $\mu$ moles, 21%) and 3260 o.d.<sub>280</sub> units of dpG<sup>IB</sup>pA<sup>Bz</sup> (109  $\mu$ moles, 69%) were also recovered.

#### *The dodecanucleotide, d-MMTrG<sup>Bz</sup>pA<sup>Bz</sup>pA<sup>Bz</sup>pC<sup>An</sup>pC<sup>An</sup>pG<sup>IB</sup>pG<sup>IB</sup>pA<sup>Bz</sup>pG<sup>IB</sup>pA<sup>Bz</sup>pC<sup>An</sup>pT*

To an anhydrous pyridine solution (7 ml.) of the protected octanucleotide (triethylammonium salt, 5800 o.d.<sub>280</sub> units, 45.7  $\mu$ moles) and triethylammonium dpG<sup>IB</sup>pA<sup>Bz</sup>pC<sup>An</sup>pT-OAc (21,800 o.d.<sub>280</sub> units, 0.40 m-mole), MS (510 mg, 2.32 m-moles) was added and the mixture was kept for 2.75 hr at room temp. with exclusion of moisture. The usual work-up was followed by column chromatography of the reaction products on DEAE-cellulose as shown in Fig. 11. Fractions containing peak IV were pooled, and the protected dodecanucleotide isolated. After removal of the *N*-protecting groups, the compound proved to be practically homogeneous on paper chromatography in solvent C, except for some

weak tailing and faint traces of faster-travelling material. A similar pattern was observed after removal of the monomethoxytrityl group from the *N*-deacylated compound and paper chromatography in solvents C and D, except that the totally unprotected compound travelled extremely slowly in these solvents (especially in solvent D) and the chromatogram had to be developed for 10 days. The yield of the protected compound was 2360 o.d.<sub>280</sub> units (13  $\mu$ moles, 28%). The recovery of the octanucleotide was 2130 o.d.<sub>280</sub> units (17  $\mu$ moles, 37%) and of the tetranucleotide block 14,600 o.d.<sub>280</sub> units (0.27 m-mole, 67%). For a further check on the purity of the unprotected dodecanucleotide, the compound (8.4 o.d.<sub>258</sub> units) was eluted from the paper chromatograms using 0.1 M-ammonia and, after evaporation of the solvent, was applied to a column (0.5 cm  $\times$  90 cm) of DEAE-cellulose (Whatman DE23, chloride form) which had been equilibrated at 4°C with a solution containing 7 M-urea, 0.05 M-sodium chloride and 0.02 M-Tris-HCl (pH 7.5). Elution was carried out using a linear salt gradient of sodium chloride (0.2 to 0.3 M; total volume 400 ml.) in the presence of 7 M-urea and 0.02 M-Tris-HCl. 3-ml. fractions were taken. The elution pattern obtained is given in Fig. 12. Only a single major peak of nucleotidic material was obtained. The peaks eluted early from the column were of non-nucleotidic material and probably consisted of impurities eluted from paper chromatographic strip.

*The hexadecanucleotide, d-MMTrG<sup>Bz</sup>pA<sup>Bz</sup>pA<sup>Bz</sup>pC<sup>An</sup>pC<sup>An</sup>pG<sup>IB</sup>pG<sup>IB</sup>pA<sup>Bz</sup>pG<sup>IB</sup>pA<sup>Bz</sup>pC<sup>An</sup>.  
pTpC<sup>An</sup>pTpC<sup>An</sup>pC<sup>An</sup>*

The protected dodecanucleotide (triethylammonium salt, 3600 o.d.<sub>280</sub> units, 20  $\mu$ moles) was condensed with triethylammonium dpC<sup>An</sup>pTpC<sup>An</sup>pC<sup>An</sup>-OAc (31,00 o.d.<sub>302</sub> units, 460  $\mu$ moles), using MS (478 mg, 2.18 m-moles) in anhydrous pyridine (6 ml.). After a reaction time of 2.5 hr at room temp., the mixture was worked up and separated on a DEAE-cellulose column as shown in Fig. 13. Fractions of peak IV were pooled and the hexadecanucleotide was isolated as usual. Paper chromatography in solvent C after removal of the *N*-protecting groups resulted in a single major spot, which showed some tailing and faint traces of faster-moving materials. The yield of the protected hexadecanucleotide was 1630 o.d.<sub>280</sub> units (6.8  $\mu$ moles, 34%). The recovered dodecanucleotide (1800 o.d.<sub>280</sub> units) was contaminated by the pyrophosphate of the tetranucleotide block (see Fig. 13). Recovery of the tetranucleotide was 22,400 o.d.<sub>302</sub> units (330  $\mu$ moles, 72%). For further characterization, the band on the paper chromatogram containing the MMTr-hexadecanucleotide (without *N*-protecting groups) was eluted from paper using 0.1 M-ammonia and, after evaporation of the solvent, the residue was treated with 80% aqueous acetic acid (1 ml.) at room temp. for 45 min. The solvent was evaporated and the residue was dissolved in water (1 ml.). The solution was extracted three times with ethyl acetate (2 ml.) and the aqueous phase was concentrated in a vacuum. The compound was applied to a DEAE-cellulose column in 7 M-urea, identical to that used for the dodecanucleotide, and eluted using a linear gradient of sodium chloride (0.23 to 0.33 M; total volume 400 ml.). Fractions of 3 ml. were collected. The elution profile is given in Fig. 14.

*The icosanucleotide, d-MMTrG<sup>Bz</sup>pA<sup>Bz</sup>pA<sup>Bz</sup>pC<sup>An</sup>pC<sup>An</sup>pG<sup>IB</sup>pG<sup>IB</sup>pA<sup>Bz</sup>pG<sup>IB</sup>pA<sup>Bz</sup>pC<sup>An</sup>.  
pTpC<sup>An</sup>pTpC<sup>An</sup>pC<sup>An</sup>pC<sup>An</sup>pA<sup>Bz</sup>pTpG<sup>IB</sup>*

An anhydrous pyridine solution (1.5 ml.) of the protected hexadecanucleotide (triethylammonium salt, 1540 o.d.<sub>280</sub> units, 6.4  $\mu$ moles), triethylammonium dpC<sup>An</sup>pA<sup>Bz</sup>pTpG<sup>IB</sup>-OAc (11,500 o.d.<sub>280</sub> units, 216  $\mu$ moles; labeled with <sup>3</sup>H at C<sub>8</sub> of the deoxyguanosine residue, specific activity,  $3.85 \times 10^5$  cts/min/ $\mu$ mole), and MS (220 mg, 1.0 m-mole) was kept at room temp for 2 hr 20 min. After the usual work-up the reaction products were separated on a DEAE-cellulose column as shown in Fig. 15. The specific activities of various fractions were determined. Fractions of peak IV were pooled and the icosanucleotide was isolated as usual. The specific activity of the product was  $3.24 \times 10^5$  cts/min/ $\mu$ mole.

The yield was 470 o.d.<sub>280</sub> units (1.6  $\mu$ moles, 25%). Recovery of the hexadecanucleotide was 460 o.d.<sub>280</sub> units (1.9  $\mu$ moles, 30%), of the tetranucleotide block 7550 o.d.<sub>280</sub> units (141  $\mu$ moles, 66%). A sample (approximately one third) of the protected icosanucleotide was treated successively with conc. ammonia and *n*-butylamine as described previously,

and then, without paper chromatographic purification, directly submitted to the 80% acetic acid treatment (45 min, room temp.). After evaporation of the solvent in a vacuum, the residue was dissolved in 1 ml. of 0.005 M-Tris-HCl (pH 8.0) and extracted with ethyl acetate (2 ml.) three times. The aqueous phase (containing 130 o.d.<sub>260</sub> units) was evaporated in a vacuum, the residue was dissolved in 0.3 ml. of 7 M-urea containing 0.02 M-Tris-HCl (pH 7.5) and the solution was applied to a column (0.5 cm × 80 cm) of DEAE-cellulose (Whatman DE32, chloride form) pre-equilibrated with 7 M-urea containing 0.02 M-Tris-HCl (pH 7.5) and 0.05 M-sodium chloride at room temp. Elution was carried out using a linear concentration gradient of sodium chloride (0.22 to 0.30 M; total volume 250 ml.). The elution pattern is shown in Fig. 16, the fractions collected had a volume of 0.8 ml. and 0.1-ml. portions were counted for radioactivity. The specific activity of the peak fractions was approximately  $3 \times 10^5$  cts/min/μmole. The peak fractions were pooled and passed through a column of Biogel P2 to remove urea and salt. 50 o.d.<sub>260</sub> units of the peak fractions and 22 o.d.<sub>260</sub> units of the head and tail fractions were recovered.

#### *Determination of the end groups of the icosanucleotide*

##### *5'-Terminus*

The synthetic icosanucleotide was phosphorylated at the 5'-end using polynucleotide kinase and [ $\gamma$ -<sup>32</sup>P]ATP according to a procedure described elsewhere (see accompanying papers). A sample of this preparation (62 pmoles = 15,000 cts/min) was digested with snake venom phosphodiesterase using standard conditions of incubation. The digest was applied on paper with appropriate unlabeled marker compounds and chromatographed in solvent D. After drying, the chromatogram was scanned for radioactivity using a Tracerlab radioactivity scanner.

##### *3'-Terminus*

The icosanucleotide (2.5 nmoles = 800 cts/min) was incubated in 33 mM-sodium glycinate (pH 9.2), 10 mM-CaCl<sub>2</sub> (total volume 0.06 ml.) with 5 μl. micrococcal nuclease at 37°C. After 2 hr an additional 5 μl. of enzyme was added and the incubation was continued for 2 hr. Addition of 1 M-acetic acid (2 μl.) adjusted the pH to about 6. 1 M-ammonium acetate, pH 5.9, was added (5 μl.), followed by 25 μg of spleen phosphodiesterase. The mixture was incubated for 7 hr at 37°C, then applied on paper together with appropriate markers and chromatographed in solvent A. The dried chromatogram was cut into 1-cm wide strips, which were counted in a toluene counting solution.

#### *Characterization of side products formed on n-butylamine treatment of compounds containing N-anisoyldeoxycytidine residues*

Various oligonucleotides containing *N*-anisoyldeoxycytidine residues (about 10 o.d. units each) were treated with a 1:1 mixture of *n*-butylamine and methanol (1 ml.) for 1 to 2 days at room temperature. The solution was then evaporated in a vacuum to dryness, and the residue was dissolved in 50% aqueous pyridine (0.05 ml.). On paper chromatography in solvent C, the unprotected oligonucleotide was always found to be contaminated by a faster-travelling side product. Furthermore, the anisic acid *n*-butylamide spot ( $R_F$  = 0.98, solvent C) was always accompanied by a slower-moving anisic acid derivative ( $R_F$  = 0.92, solvent C), which was shown by thin-layer chromatography (Silicagel Merck HF254; solvent: chloroform/ethanol, 9 : 1, v/v) to be anisic acid amide ( $R_F$  = 0.62, anisic acid) *n*-butylamide;  $R_F$  = 0.32, anisic acid amide). Enzymic degradation of the various side products using venom phosphodiesterase resulted in the formation, along with the normal nucleoside 5'-phosphates, of a fast-moving nucleotide identical with the side product obtained by *n*-butylamine treatment of d-pC<sup>An</sup>. The u.v. characteristics of this compound were  $\lambda_{max}$  271, 235 nm,  $\lambda_{min}$  251, 230 nm (neutral pH);  $\lambda_{max}$  282 nm,  $\lambda_{min}$  243 nm (acid pH). The amount of the nucleotide suggested that it had been formed by alteration of one of the deoxycytidine residues of the oligonucleotide.

The side-product formation did not occur when the *n*-butylamine treatment was replaced by a treatment with concentrated ammonia for 2 to 3 days. Properties of various side products are listed in Table 6.

## REFERENCES

- Gupta, N. K. & Khorana, H. G. (1968). *Proc. Nat. Acad. Sci., Wash.* **61**, 215.
- Gupta, N. K., Ohtsuka, E., Sgaramella, V., Büchi, H., Kumar, A., Weber, H. & Khorana, H. G. (1968). *Proc. Nat. Acad. Sci., Wash.* **60**, 1338.
- Hachmann, J. & Khorana, H. G. (1969). *J. Amer. Chem. Soc.* **91**, 2749.
- Holley, R. W., Apgar, J., Everett, G. A., Madison, J. T., Marquisee, M., Merrill, S. H., Penswick, J. R. & Zamir, A. (1965). *Science*, **147**, 1462.
- Jacob, T. M. & Khorana, H. G. (1964). *J. Amer. Chem. Soc.* **86**, 1630.
- Jacob, T. M. & Khorana, H. G. (1965). *J. Amer. Chem. Soc.* **87**, 368.
- Khorana, H. G. (1967). *Proc. 7th Int. Congr. Biochem.* p. 17.
- Khorana, H. G. (1968). *J. Pure & Applied Chem.* **17**, 349.
- Khorana, H. G., Büchi, H., Caruthers, M. H., Chang, S. H., Gupta, N. K., Kumar, A., Ohtsuka, E., Sgaramella, V. & Weber, H. (1968). *Cold Spr. Harb. Symp. Quant. Biol.* **33**, 35.
- Khorana, H. G., Turner, A. F. & Vizsolyi, J. P. (1961). *J. Amer. Chem. Soc.* **83**, 686.
- Kleppe, K., Ohtsuka, E., Kleppe, R., Molineux, I. & Khorana, H. G. (1971). *J. Mol. Biol.* **56**, 341.
- Kössel, H. R., Büchi, H. & Khorana, H. G. (1967). *J. Amer. Chem. Soc.* **89**, 2185.
- Kössel, H. R., Moon, M. W. & Khorana, H. G. (1967). *J. Amer. Chem. Soc.* **89**, 2148.
- Lohrmann, R. & Khorana, H. G. (1966). *J. Amer. Chem. Soc.* **88**, 829.
- Ohtsuka, E. & Khorana, H. G. (1967). *J. Amer. Chem. Soc.* **89**, 2195.
- RajBhandary, U. L., Chang, S. H., Stuart, A., Faulkner, R. D., Hoskinson, R. M. & Khorana, H. G. (1967). *Proc. Nat. Acad. Sci., Wash.* **57**, 751.
- Ralph, R. K., Connors, W. J., Schaller, H. & Khorana, H. G. (1963). *J. Amer. Chem. Soc.* **85**, 1983.
- Ralph, R. K. & Khorana, H. G. (1961). *J. Amer. Chem. Soc.* **83**, 2926.
- Richardson, C. C. (1965). *Proc. Nat. Acad. Sci., Wash.* **54**, 158.
- Schaller, H., & Khorana, H. G. (1963). *J. Amer. Chem. Soc.* **85**, 3841.
- Smith, M., Drummond, G. I. & Khorana, H. G. (1961). *J. Amer. Chem. Soc.* **83**, 698.
- Tomlinson, R. V. & Tener, G. M. (1962). *J. Amer. Chem. Soc.* **84**, 2644.
- Weimann, G. & Khorana, H. G. (1962). *J. Amer. Chem. Soc.* **84**, 419.