BIOORGANIC CHEMISTRY 2, 322–336 (1973)

Chemical Synthesis of a Tetradecamer in the Deoxyribonucleic Acid Series[1]

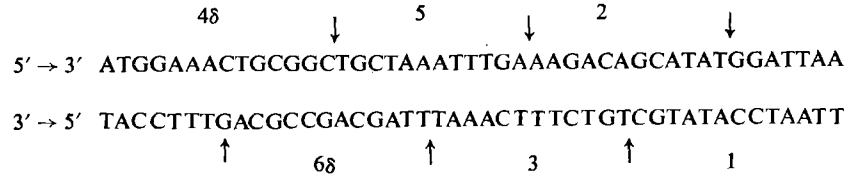
M. S. POONIAN, E. F. NOWOSWIAT, L. TOBIAS, AND A. L. NUSSBAUM

Chemical Research Department, Hoffmann-La Roche Inc., Nutley, New Jersey 07110

Received February 26, 1973

The chemical synthesis of a tetradecadeoxyribonucleotide, d-EtSp(A-T-G-G-A-A-A-C-T-G-C-G-G-C), is described. This oligomer, designated Fragment 4δ, constitutes the 5′-terminus of the *plus* strand of a projected duplex coding for S-Peptide$_{2-14}$ derived from Ribonuclease A. The Fragment was constructed by block condensation via a phosphorothioate anchor. Complications due to inadvertent phosphotriester condensations are discussed. Arguments justifying the sequence selection are presented.

Previous papers from this laboratory have dealt with the chemical synthesis of oligomers designed to be joined enzymically to form a stretch of double-helical DNA coding for a modified S-peptide (S-peptide$_{2-14}$) (*1*). Utilizing the strategy of Khorana (*2*), which entails a "leap frog" approach where chemical synthesis alternates with enzymatic joining, a number of fragments have been synthesized, and several have been



SCHEME 1. Structure of double-stranded DNA segment coding for S-peptide$_{2-14}$. The arrows indicate the ends of chemically synthesized fragments. The latter are referred to in the text by numerals and Greek letters.

joined together: Scheme 1 shows the overall design. Fragments 1 (*3*), 2 (*4*), 3 (*5*), and 5 (*6*) have been reported; the smaller fragments constituting the 3′-termini have also been constructed: their synthesis will be reported in a future publication. In this paper, we describe the chemical synthesis of Fragment 4δ; it will be seen that this constitutes the completion of the fragments in the *plus* (or "nonsense") strand[2] of the duplex.

[1] This publication constitutes Paper X in a series entitled Nucleoside S-Alkyl Phosphorothioates. For Paper IX, see Ref. (*5*).

[2] The *plus* (or "nonsense") strand is identical (except for the substitution of T for U) to a hypothetical RNA messenger transcribed from the *minus* (or "sense") strand. *Plus* and *minus* strands bear, of course, the usual Watson-Crick relationship.

The synthesis of this fragment, a tetradecamer constituting the 5′-terminus of its strand, followed the procedures previously outlined: the 5′-phosphate of the growing oligomer was protected by an S-ethyl substituent which survives all manipulations employed during the various condensations. Since the synthesis of sizeable amounts of *minus* strand is envisaged to result from repeated enzymatic copying of the plus strand template (in conjunction with an appropriate primer), and since the presence of the 5′-terminal S-alkyl group does not interfere with enzymatic joining at a different site (*5*), the removal of that protecting group was unnecessary for the present.

EXPERIMENTAL

Materials and Methods

Reagent grade pyridine (Fisher Scientific) was distilled from potassium hydroxide and stored over molecular sieves. This sample of pyridine was used in all phosphodiester condensation reactions and for obtaining solutions of protected oligomers for precipitation by introduction into dry ether. However, column effluents containing triethylammonium bicarbonate buffer were concentrated by coevaporation with undistilled reagent grade pyridine. Mesitylenesulfonyl chloride (Aldrich Chemical Company) was recrystallized twice from *n*-pentane and stored over P_2O_5 in a desiccator. Reagent grade triethylamine (Matheson, Coleman, and Bell) was used without further purification for preparation of triethylammonium bicarbonate buffers. Nucleoside and nucleotide monomers were purchased from Zellstofffabrik Waldhof, Mannheim, Germany. Thiophosphoryl chloride, a product of Alfa Inorganics, was distilled just prior to use in the preparation of trilithium phosphorothioate (*7*).

Phosphodiester condensation reactions were performed in the dark at room temperature in dry pyridine medium. Reaction time ranged between 2 and $3\frac{1}{2}$ hr. The amount of the condensing reagent, i.e., mesitylenesulfonyl chloride, varied between 0.7 and 1.0 equivalent per phosphate charge. Selective 3′-deacetylation was achieved by treatment of a 25% pyridine solution of the protected oligonucleotides in 1 *M* NaOH for 15 min at 0°C. The hydrolysis reaction was terminated by neutralizing the alkaline solution with the pyridium form of Dowex-50 ion exchange resin. The components of the condensation reaction mixture were usually (unless indicated otherwise) separated on a column of DEAE-cellulose (DE-23, bicarbonate form) by elution with a gradient of triethylammonium bicarbonate buffer (pH 7–7.5) at 4°C. The emergence of a peak of nucleotidic material was monitored by uv spectroscopy. Peaks (figures and Table 1) were identified on the basis of uv spectrum (Table 2), paper and thin layer chromatography (Table 3), sulfur test and nucleotide composition analysis after venom phosphodiesterase digestion (Table 4). Homogeneous portions of the peaks were pooled and concentrated by coevaporation with pyridine (bath temperature, 30°C) under vacuum in a flash evaporator. Triethylammonium bicarbonate buffer thus having been removed, the residue was dissolved in dry pyridine and precipitated by dropping the solution gradually into anhydrous ether. The precipitate was isolated either by filtration or centrifugation and dried in a vacuum desiccator.

Preparation of Oligonucleotides

Synthesis of d-EtSp(bz*A-T*-ibu*G*), **2**.[3] A solution of protected dimer (*3*) d-EtSp-(*bz*A-T) (**1**, 16.7 g, 17.5 mmole), and mononucleotide d-p(*ibu*G) (*i*Bu) (29.4 g, 52.5 mmole) in dry pyridine (125 ml) was treated with anhydrous mesitylenesulfonyl chloride (21.4 g, 98 mmole) in the dark at room temperature for 3 hr. The reaction was cooled

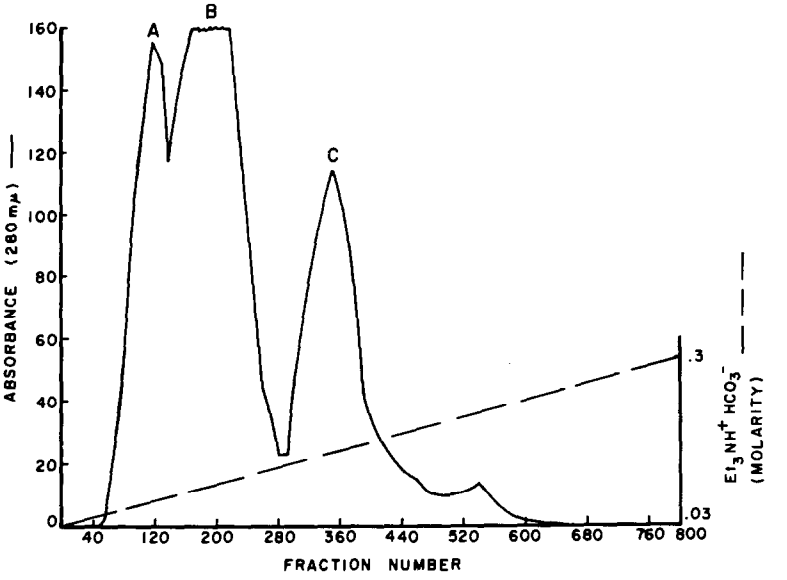
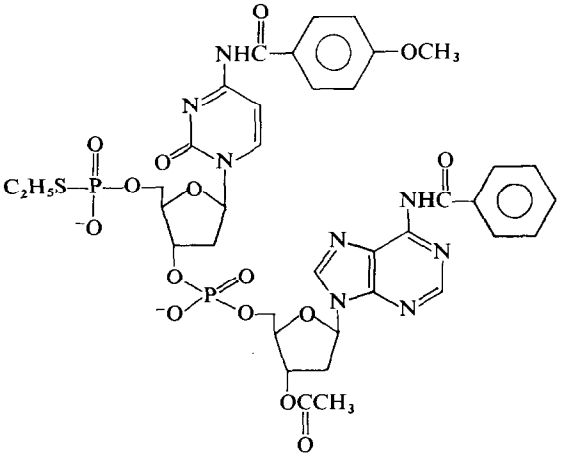


FIG. 1. DEAE-cellulose (DE-23) column chromatography of the components from condensation between a dimer d-EtSp(*bz*A-T) and mononucleotide d-p(*ibu*G) (*i*Bu). The solution of the partially hydrolyzed reaction products was charged onto a column (5 × 80 cm) and eluted with a linear gradient of triethylammonium bicarbonate, pH 7.0 (from 0.025 *M* to 0.3 *M* 8 liters each vessel). Fractions of 20 ml were collected.

[3] Abbreviations of nucleotides are in strict accordance with the 1970 IUPAC–IUB recommendations, *Biochemistry* **9**, 4022 (1970), special attention being directed to sections 3.1 and 3.2. In addition, prefix EtS denotes ethylthio: thus, d-EtSp(*an*C-*bz*A) (Ac) symbolizes



on ice and stopped by the addition of 98 mmoles of cold triethylamine and 223 ml of cold water. After keeping the quenched reaction at 4°C overnight, it was treated with 1 *M* NaOH in 25% pyridine solution for 15 min at 0°C. This procedure selectively hydrolyzes the 3'-acetyl blocking group. The components of the reaction mixture were then separated on a DEAE-cellulose (DE-23) column (bicarbonate form). The chromatographic profile and details are shown in Fig. 1 and its legend. The isolated trimer amounted to 4.44 mmoles (25.4% yield).

The anchor block, d-EtSp(bz*A*-*T*-ibu*G*-ibu*G*), **3**. The total amount (4.4 mmole) of trimer d-EtSp(*bz*A-T-*ibu*G), **2**, obtained in the above preparation was used in condensation with protected mononucleotide d-p(*ibu*G) (*i*Bu) (13.32 mmole, 3-fold excess).

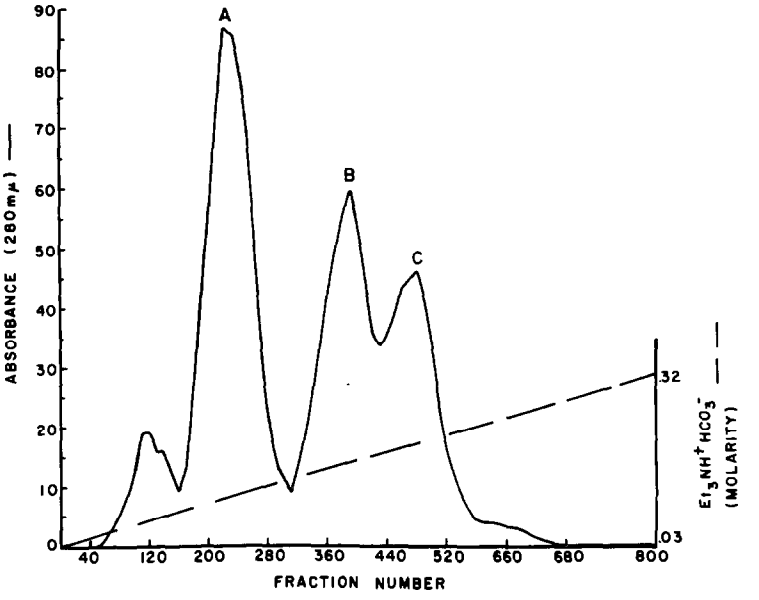


FIG. 2. DE-23 column chromatography of the condensation between a trimer d-EtSp(*bz*A-T-*ibu*G) and mononucleotide d-p(*ibu*G) (*i*Bu). The mixture was fractionated on a column (5 × 80 cm) under a triethylammonium bicarbonate (pH 7.0) linear gradient of 0.03 *M* to 0.32 *M* (8 liters each). Volume of each fraction was 20 ml.

The reaction was performed in a total of 50 ml dry pyridine with 27.97 mmole of mesitylenesulfonyl chloride for 3 hr. The selectively deacylated mixture was then fractionated on a DE-23 column (5 × 80 cm). The desired tetramer block was located between fractions 450 and 550 (Fig. 2). The isolated yield of the analytically pure product amounted to 1.36 mmole (30.9%).

Tetramer d-p(bz*A*-an*C*-*T*-ibu*G*), **6**. Trimer d-p(*bz*A-*an*C-T), 1.74 mmole, was cyanoethylated according to the usual procedure and 1.57 mmole of the resultant d-CNEtp(*bz*A-*an*C-T), was condensed with 11.6 mmole d-p(*ibu*G) (*i*Bu) in the presence of 27.9 mmole mesitylenesulfonyl chloride in 45 ml of anhydrous pyridine for 3 hr. Fractionation of the products after selective alkaline hydrolysis on a DE-23 (4 × 89 cm) column provided the desired tetrameric product between fractions 450 and 580 in 32.4% yield (0.51 mmole).

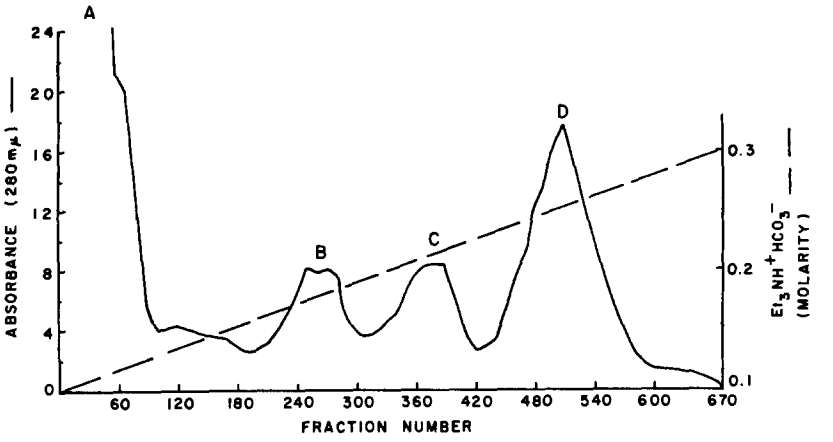


FIG. 3. DE-23 column chromatography on a mixture of condensation between a trimer d-CNEtp-(*bz*A-*an*C-T) and d-p(*ibu*G) (*i*Bu). Fractionation was achieved on a column (4 × 89 cm) under a convex gradient of triethylammonium bicarbonate (pH 7.0) from 0.1 *M* (5 liters) to 0.28 *M*. Fractions of 20 ml were collected.

Trimer d-CNEtp(an*C*-ibu*G*-ibu*G*). A sample of protected dimer d-CNEtp(*an*C-*ibu*G), 8 mmole, was mixed with a twofold excess of protected deoxyguanosine mononucleotide d-p(*ibu*G) (*i*Bu) and dissolved in 50 ml of anhydrous pyridine. Condensation between the two components was achieved by adding 48 mmole of mesitylenesulfonyl chloride. The reaction was allowed to proceed for 3½ hr. Quenching of the reaction mixture was followed by selective deacylation and decyanoethylation, and finally DEAE-cellulose chromatography provided the trimeric product in 17.5% yield (1.4 mmole). Further cyanoethylation of the material was achieved according to the usual procedure.

The terminal tetramer block, d-p(an*C*-ibu*G*-ibu*G*-an*C*), **8**. A mesitylenesulfonyl chloride (18.2 mmole) promoted condensation between cyanoethylated trimer, d-CNEtp(*an*C-*ibu*G-*ibu*G), 1.4 mmole, and d-p(*an*C) (Ac), 7.0 mmole in anhydrous pyridine (50 ml) for 2 hr provided the desired tetramer d-p(*an*C-*ibu*G-*ibu*G-*an*C) in 14% yield. After selective alkaline hydrolysis of the 5′ cyanoethyl and 3′-*O*-acetyl substituents, the desired product was separated from unreacted materials and side products on a DE-23 (5 × 65 cm) column (Fig. 4). Fractions 680–740 represent the region of the analytically pure tetramer on the chromatogram.

d-EtSp(bz*A*-*T*-ibu*G*-ibu*G*-bz*A*-bz*A*), **4**. A 3.5-fold excess of the protected dimer d-p(*bz*A-*bz*A) (Ac), 4.4 mmole, was allowed to react with anchor tetramer d-EtS-p(*bz*A-T-*ibu*G-*ibu*G), 1.25 mmole, in a total of 50 ml anhydrous pyridine solution. Mesitylenesulfonyl chloride (13.4 mmoles) was added, and the reaction stirred for 3½ hr in the dark at room temperature. After stopping the reaction and selective 3′-hydroxyl deblocking, the oligomeric components were fractionated on a DEAE-cellulose column (4 × 85 cm). The profile of the chromatogram is presented in Fig. 5. The pool of the nucleotidic material between fractions 400 and 560, although highly enriched in the desired hexamer, was contaminated with surviving tetramer. It was therefore concentrated to a small volume, adjusted to an appropriate conductivity,

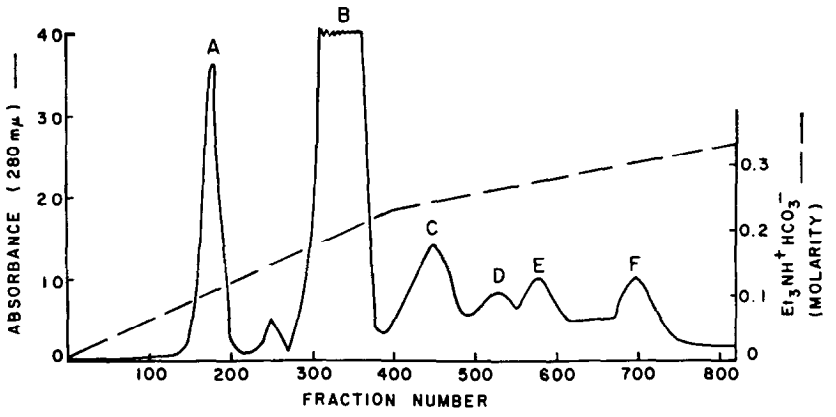


FIG. 4. DE-23 column chromatography of the mixture of products obtained from a condensation between trimer d-CNEtp(*an*C-ibuG-ibuG)ibuG) and d-p(anC) (Ac). The mixture was loaded on a (5 × 65 cm) column and eluted first with a triethylammonium bicarbonate linear gradient of 0 to 0.23 *M* (4 liters each). At fraction 400, the gradient was stepped up from 0.23 *M* (4 liters) to 0.33 *M* (4 liters). Fractions of 20 ml were collected throughout.

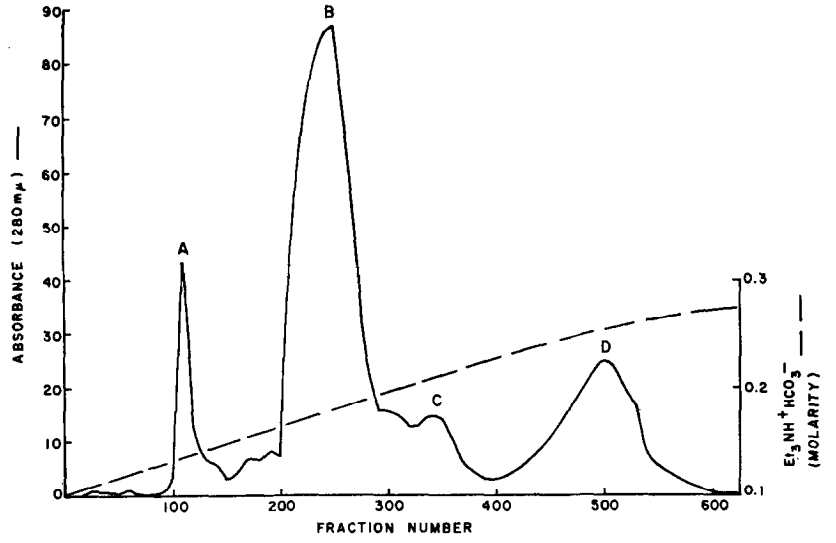


FIG. 5. DE-23 Column chromatography of the 4 + 2 condensation products, i.e., d-EtSp(*bz*A-T-*ibu*G-*ibu*G) + d-p(*bz*A-*bz*A) (Ac). Column dimensions were 4 × 85 cm and a convex gradient of triethylammonium bicarbonate (pH 7.0) between 0.1 *M* (8 liters) and 0.32 *M* was used for elution. The hexameric product obtained by pooling fractions 400 to 560 was further purified by rechromatography on a DE-23 (4 + 98 cm) column under a TEAB convex gradient between 0.1 *M* (8 liters) and 0.25 *M* (not shown). In both columns, fractions of 20 ml were collected.

and rechromatographed on a DE-23 column (4 × 98 cm) under shallower convex gradient [0.1 *M* TEAB (8-liter mixing vessel) to 0.25 *M*]. The overall yield of the pure hexamer thus obtained was 0.35 mmoles (35.5%).

Decamer, *d-EtSp*(bz*A*-*T*-ibu*G*-ibu*G*-bz*A*-bz*A*-bz*A*-an*C*-*T*-ibu*G*), 7. The sample of S-ethyl hexamer, d-EtSp(*bz*A-T-*ibu*G-*ibu*G-*bz*A-*bz*A) *4*, (200 μmoles) as obtained in the experiment described above, was dissolved in anhydrous pyridine along with the

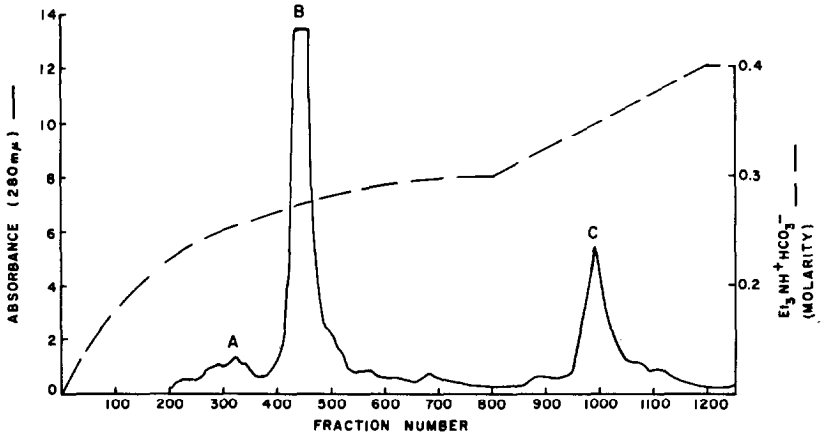


FIG. 6. DE-23 column (4.4 × 90 cm) chromatography of the condensation between hexamer d-EtSp(*bz*A-T-*ibu*G-*ibu*G-*bz*A-*bz*A) and tetramer d-p(*bz*A-*an*C-T-*ibu*G) (*i*Bu). A convex gradient of triethylammonium bicarbonate (pH 7) starting from 0.1 *M* (4 liters) to 0.3 *M* was used for elution in the first step. As shown, the second step of the gradient was linear between a molarity of 0.3 to 0.4 (4 liters each). Fractions of 20 ml were collected.

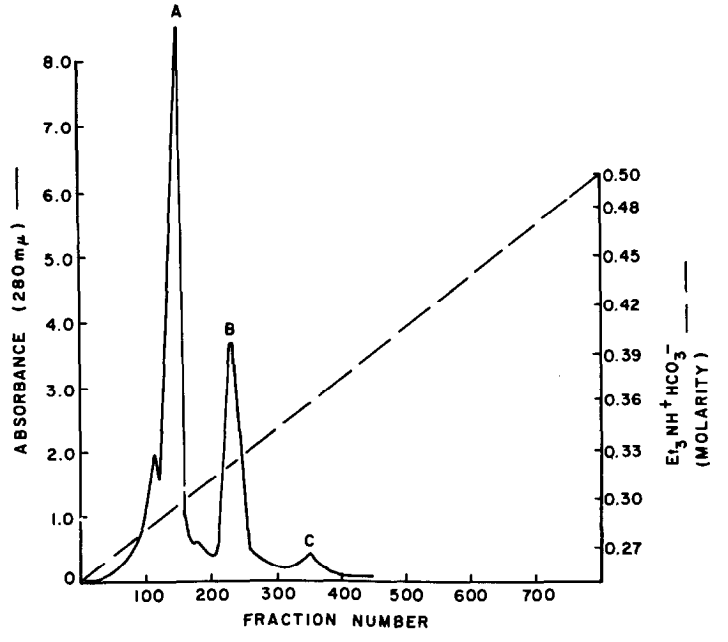


FIG. 7. DE-23 column (2.5 × 75 cm) chromatography of the decamer and tetramer condensation reaction: d-EtSp(*bz*A-T-*ibu*G-*ibu*G-*bz*A-*bz*A-*bz*A-*an*C-T-*ibu*G) + d-p(*an*C-*ibu*G-*ibu*G-*an*C) (Ac). The components were eluted with a linear gradient of triethylammonium bicarbonate (pH 7.0) for 0.25 *M* to 0.5 *M* (6 liters each). Fractions of 15 ml were collected.

protected tetramer block d-p(*bz*A-*an*C-T-*ibu*G) (Ac) (500 μmoles, 2.5-fold excess). The total reaction volume of 10 ml was then charged with condensing agent mesitylene sulfonyl chloride (3.7 mmoles). Under twist action shaker, the reaction was continued for 3½ hr. After the usual aqueous treatment followed by selective hydrolysis of the 3'-*O*-acetyl substituent, the mixture was loaded on a DEAE-cellulose column (4.4 × 90 cm), and fractionation was achieved by a two-step gradient (Fig. 6). The region between fractions 950 and 1050 analyzed to be pure decamer in 18.8% yield (37.6 μmoles).

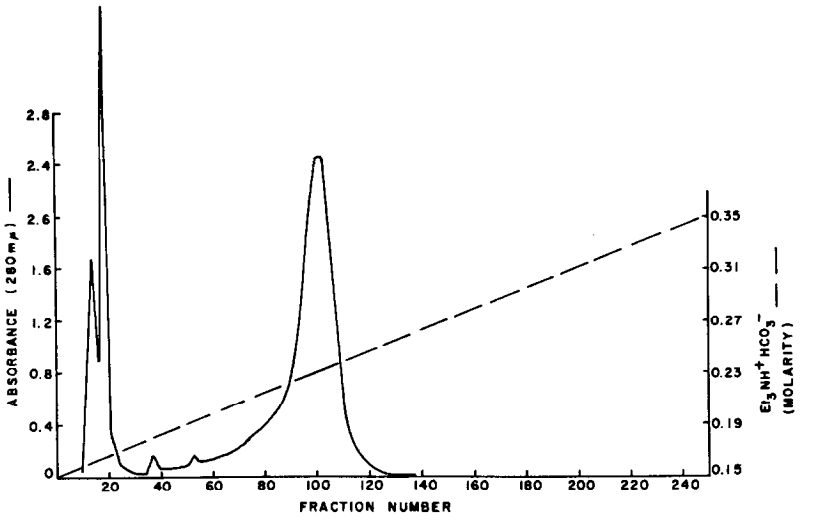


FIG. 8. DE-23 chromatography of completely deacylated S-ethyl tetradecamer, **10**, in 7 *M* urea and 0.2 *M* Tris (pH 7.4) buffer. The nucleotidic material was eluted under a linear gradient (0.15 *M* to 0.35 *M*) of NaCl in 7 *M* urea and 0.02 *M* Tris buffered at pH 7.4 (each vessel 500 ml). Fractions of 4 ml were collected at a column flow rate of 40 ml/hr.

Tetradecamer d-EtSp(A-T-G-G-A-A-A-C-T-G-C-G-G-C). The final condensation was performed in a solution of protected S-ethyldecamer (18.76 μM) and d-p(*an*C-*ibu*G-*ibu*G-*an*C) (Ac), (**8**, 93.8 μmoles) in anhydrous pyridine (5 ml) by adding mesitylenesulfonyl chloride (657 μmoles). A 3-hr reaction time at room temperature was allowed. After selective alkaline deacetylation, the components were separated on a DE-23 column (2.5 × 75 cm) as depicted in Fig. 7. The fractions between 300 and 380 (yielding 284 A_{280}, 9.5%) were pooled, concentrated, and hydrolyzed with concentrated ammonia at room temperatute for 48 hr. The ammonia solution was taken to dryness *in vacuo*, and the residue dissolved in 2 ml of water for charging onto a DE-52 column (0.7 × 120 cm) which had previously been equilibrated with a 7 *M* urea solution (pH 7.4) containing 0.02 *M* Tris and 0.15 *M* NaCl. A linear gradient of 0.15 *M* to 0.35 *M* NaCl in 0.02 *M* Tris and 7*M* urea (500 ml each) was used for elution (chromatographic profile, Fig. 8) at room temperature. The tetradecamer peak between fractions 70 and 120 was pooled (168 $A_{260\,nm}$) and desalted on a G-15 Sephadex column (2.5 × 90 cm) with 0.1 *M* triethylammonium bicarbonate. Recovery of the nucleotidic material from G-15 was 124 A_{260} (74%).

TABLE 1

IDENTIFICATION OF PEAKS IN COLUMN CHROMATOGRAPHS

Figure	Peak A	B	C	D	E	F
1	d-p(*ibu*G)	**1**	**2**			
2	d-p(*ibu*G)	**2**	**3**			
3	d-p(*ibu*G)	surviving d-CNEtp(*bz*A-*an*C-T-*ibu*G)	d-p(*bz*A-*an*C-T)	**6**		
4	$MesSO_3H$[a]	d-p(*an*C)	pyrophosphate of d-p(*an*C)	unidentified	d-p(*an*C-*ibu*G-*bu*G)	**8**
5	$MesSO_3H$[a]	**5**	pyrophosphate of **5**	**4** + **3**[b]		
6	unidentified	mixture of **4** and **6**	**7**			
7	**8**	**7**	**9**			

[a] Mesitylene sulfonic acid.
[b] Rechromatographed: see Experimental.

TABLE 2

ULTRAVIOLET SPECTRA OF THE OLIGOMERS[a]

Compound	λ_{max}	λ_{min}	260/280	300/280
d-EtSp(*bz*A-T-*ibu*G)	276 (sh 262)	230	1.0	0.32
d-EtSp(A-T-G)	257	227		
d-p(*an*C-*ibu*G-*ibu*G)	260 (sh 282)	270, 232	1.05	0.79
d-p(C-G-G)	250 (sh 270)	234		
d-p(*bz*A-*an*C-T-*ibu*G)	278 (sh 262)	233	0.86	0.58
d-p(A-C-T-G)	256	226		
d-EtSp(*bz*A-T-*ibu*G-*ibu*G)	265 (sh 275)	228	1.04	0.37
d-EtSp(A-T-G-G)	255	227		
d-p(*an*C-*ibu*G-*ibu*G-*an*C)	288 (sh 279)	258, 235	0.96	0.92
d-p(C-G-G-C)	251 (sh 268)	227		
d-EtSp(*bz*A-T-*ibu*G-*ibu*G-*bz*A-*bz*A)	278 (sh 259)	232	0.87	0.40
d-EtSp(A-T-G-G-A-A)	256	230		
d-EtSp(*bz*A-T-*ibu*G-*ibu*G-*bz*A-*bz*A-*bz*A-*an*C-T-*ibu*G)	277 (sh 260)	234	0.90	0.45
d-EtSp(A-T-G-G-A-A-A-C-T-G)	256	230		
d-EtSp(*bz*A-T-*ibu*G-*ibu*G-*bz*A-*bz*A-*bz*A-*an*C-T-*ibu*G-*an*C-*ibu*G-*ibu*G-*an*C)	280 (sh 260)	267, 240	0.95	0.61
d-EtSp(A-T-G-G-A-A-A-C-T-G-C-G-G-C)	257			

[a] Spectra in distilled water.

TABLE 3

R_f VALUES OF OLIGOMERS IN PC AND TLC

Oligomer	System[a,b]					
	A	B	C	D	E	F
d-EtSp(A-T-G)	0.47				0.68	
d-p(C-G-G)	0.50[c]					0.41
d-p(A-C-T-G)	0.30			0.53	0.37	
d-EtSp(A-T-G-G)	0.29	0.58[d]			0.61	
d-p(C-G-G-C)	0.43[c]					0.34
d-EtSp(A-T-G-G-A-A)	0.24		0.23	0.53	0.44	
d-EtSp(A-T-G-G-A-A-A-C-T-G)			0.18	0.19		
d-EtSp(A-T-G-G-A-A-A-C-T-G-C-G-G-C)			0.08			

[a] Solvent systems (descending, Whatman No. 1); (A) *n*-PrOH 55, H_2O 35, $NH_3$10; (B) EtOH 70, 1 *M* NH_4Ac (pH 7.5) 30; (C) *n*-PrOH 50, 0.1 *M* NH_4Ac (pH 7.5) 50; (D) *n*-PrOH 50, 0.05 *M* NH_4Ac (pH7.5) 50.

[b] TLC (silica gel): (E) EtOH 5, MeOH 3, 0.1 *M* SDS 3 (*22*); (F) EtOH 5, MeOH 3, 0.1 *M* NH_4Ac 3.

[c] With respect to d-pC.

[d] With respect to d-pG.

TABLE 4

NUCLEOTIDE COMPOSITION OF OLIGOMERS[a]

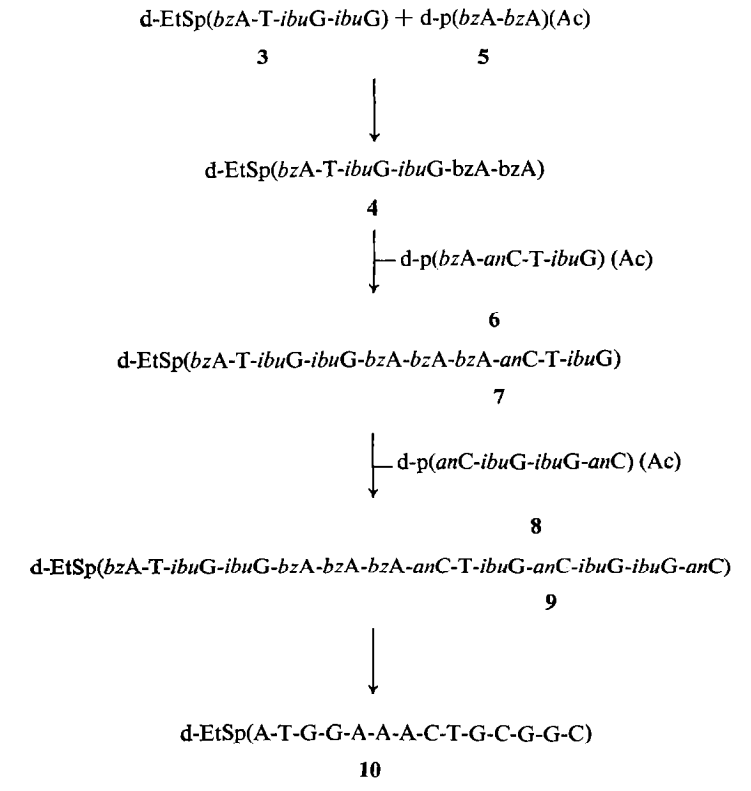
Compound	Monomer			
	d-pT	d-pC	d-pG	d-pA
d-EtSp(A-T-G)	32.3(33.3)	—	33.0(33.3)	34.6(33.3)
d-p(C-G-G)	—	35.2(33.3)	64.8(66.6)	—
d-p(A-C-T-G)	23.2(25)	25.3(25)	26.9(25)	25.4(25)
d-EtSp(A-T-G-G)	22.5(25)	—	25.5(25)	51.9(50)
d-p(C-G-G-C)	—	50.4(50)	49.5(50)	—
d-EtSp(A-T-G-G-A-A)	16.3(16.6)	—	32.2(33.3)	51.6(50)
d-EtSp(A-T-G-G-A-A-A-C-T-G)	19.1(20)	10.4(10)	28.5(30)	47.2(40)
d-EtSp(A-T-G-G-A-A-A-C-T-G-C-G-G-C)	15.5(14.3)	22.5(21.9	31.9(35.7)	29.3(28.5)

[a] Determined by digestion with snake venom phosphodiesterase and high pressure liquid chromatography (*20*). Data are expressed in mole percent, with calculated values in brackets.

RESULTS AND DISCUSSION

The chemical synthesis of oligomers in the nucleic acid series still trails similar efforts in the peptide field by a considerable margin. Much of the difficulty stems from the fact that the necessary condensation reactions do not proceed with nearly the efficiency of

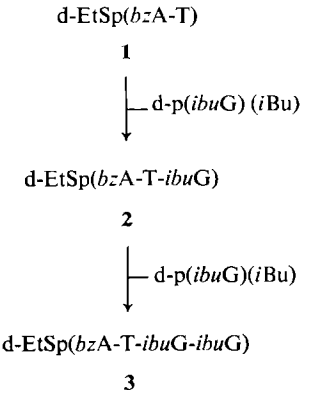
amino acid condensations and that the separation of complex mixtures becomes necessary. Nevertheless, much effort continues to be devoted to this end. Thus, the synthesis of a DNA sequence complementary to the tyrosine suppressor t-RNA of *E. coli* is progressing in Khorana's laboratory (*8*). On a more modest scale, the synthesis of an oligomer complementary to a cohesive end of coliphage lambda (*9*) and two oligomers containing a portion of the natural sequence of T4-lysozyme gene (*10*) have been reported.



SCHEME 2. Synthesis of the tetradecamer.

In this paper, we describe the synthesis of a tetradecamer in the deoxyribo series, which is to be joined into a larger array (*vide supra*, Scheme 1). This tetradecamer (termed Fragment 4δ) was constructed according to the general outline shown in Scheme 2. It will be seen that, as in our previous studies, the 5′-phosphate of the growing fragment is protected throughout by an S-ethyl substituent (*11*). A protecting group of this type must survive the various operations of coupling and unblocking of other functional groups during the synthesis (*12*) in order to avoid self-condensation. Since we envisage the use of this fragment as part of a template for an enzymatic copying process, compound **10**, the final product of the synthesis, is still in the form of a phosphorothioate, not incompatible with the proposed template function.

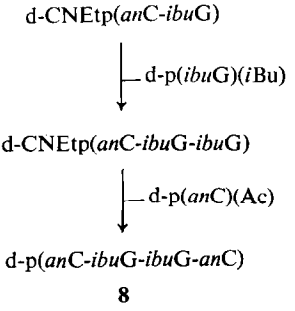
The 5′-terminal ("anchor") tetramer was obtained by two stepwise mononucleotide additions, as shown in Scheme 3. Elongation to a hexamer **4** was carried out by block-



SCHEME 3. Synthesis of the anchor block.

condensation with the appropriately protected preformed dimer **5** (*13*). Further condensation of this hexamer with tetramer 6 [synthesized by the condensation of d-p(*ibu*G) (*i*Bu) with d-CNEtp(*bz*A-*an*C-T) (*14*)] gave decamer **7**.

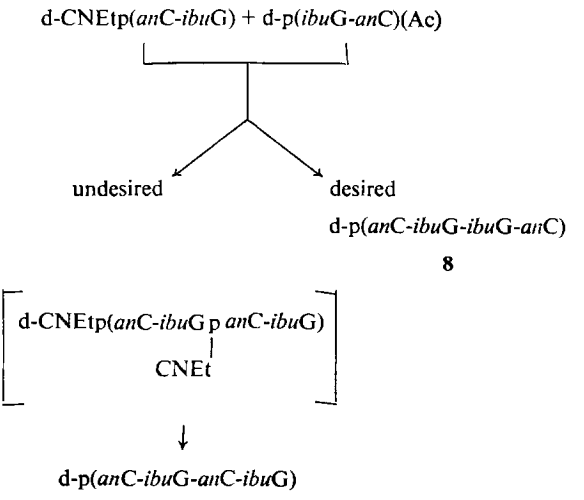
The 3′-terminal tetramer **8** was synthesized in similar fashion as the preceding, **6**, i.e., by two monomer addition to cyanoethylated d-p(*an*C-*ibu*G) (*15*) (see Scheme 4).



SCHEME 4. Synthesis of the terminal block.

A word of caution, illustrating a general problem, is in order here. In an earlier attempt to prepare this same tetramer by condensing d-p(*ibu*G-*an*C) (Ac) (*16*) with the preceding protected sequence isomer, a tetrameric fraction was obtained which had the appropriate mononucleotide composition and the proper 5′-end (nucleoside analysis after phosphomonoesterase and venom phosphodiesterase treatment). However, this material was found to be heterogeneous at its 3′-terminus; some 30% of the nucleoside obtained after phosphomonoesterase and spleen phosphodiesterase turned out to be deoxyguanosine, the rest being the expected deoxycytidine. It seems likely that this observation

finds its explanation in the series of events depicted in Scheme 5: in addition to the desired course, self-condensation of d-CNEtp(*an*C-*ibu*G) by a phosphotriester pathway gives rise to the material depicted in brackets, which generates the undesired sequence isomer of **8** upon partial hydrolysis. A similar observation has recently been made by Holý



SCHEME 5. Aborted tetramer synthesis.

and Šorm (*17*). It may well be that such triester condensations occur quite frequently in this type of work: the products are usually separable from the desired material and lost among the numerous condensation side products. It is only in a case as the present, where inseparable sequence isomers are produced, that the complication becomes fatal.

Condensation of genuine tetramer **8** with decamer **7** gave the protected tetradecamer **9**, which was deacylated to give the desired **10**. Final purification of the latter was carried out on DEAE-cellulose in 7 *M* urea–Tris buffer under a gradient of NaCl (*18*). The properties of this material, as well as those of the other oligomers first described in this paper, are summarized in the tables (see Experimental).

Some remarks concerning the process of sequence selection may be made at this point. The tetradecamer described in this paper constitutes the 5′-terminus of the *plus* (or "nonsense") strand of the proposed duplex (Scheme 1); it would therefore be identical with a corresponding stretch of messenger RNA copied from its complement. Table 5 aligns the triplets from this RNA sequence with the corresponding amino acids for which they are intended to code. These assignments are based on the genetic code which is in

TABLE 5

CODON VALIDATION

RNA corresponding to Fragment 4δ	AUG	GAA	ACU	GCG	GC
Cognate amino acids	N_{fo}Met	Glu	Thr	Ala	
References	(*22–25*)	(*22, 23, 26*)	(*22, 23, 26*)	(*23, 27*)	

turn derived on a wide variety of experiments, both *in vitro* and *in vivo*. However, the code is highly degenerate, and many different sequences could be assigned. Nevertheless, while the universality of the code would prevent errors, the actual translatability of a specific sequence in any given system (say, *E. coli*) is not assured (*6*, *14*). It is, therefore, highly desirable that the validity of these triplets in the intended target system (*E. coli*) can be documented. A convincing demonstration that the sequence selected is valid comes from the actual sequencing of the genes for known proteins of certain very small coliphages. These organisms presumably have too little RNA to specify their own translation machinery, as may be the case with more complex phages (*21*), so that the codons derived from such studies should be valid for the *E. coli* system. It is gratifying to note that such validation can be deduced from recent sequence studies: the references quoted in Table 5 substantiate the sequence selected.

ACKNOWLEDGMENTS

We are greatly indebted to our colleagues for their assistance and counsel. Special thanks are due to E. Heimer and G. Mack for starting materials, T. Gabriel and J. Michalewsky for monomer analysis, and A. deCzekala for technical assistance.

REFERENCES

1. For leading references to modified S-peptides, see, e.g., K. Hoffmann, R. Andreatta, F. M. Finn, J. Montibeller, G. Porcelli, and A. J. Quattrone, *Bioorg. Chem.* **1**, 66 (1971).
2. K. Agarwal, H. Buchi, M. Caruthers, N. Gupta, H. G. Khorana, K. Kleppe, A. Kumar, E. Ohtsuka, U. RajBhandary, J. van der Sande, V. Sgaramella, H. Weber, and T. Yamada, *Nature (London)* **227**, 27 (1970). This paper also serves as a leading reference to numerous Khorana publications for methodology.
3. A. F. Cook, E. P. Heimer, M. J. Holman, D. T. Maichuk, and A. L. Nussbaum, *J. Amer. Chem. Soc.* **94**, 1334 (1972).
4. M. S. Poonian, E. F. Nowoswiat, and A. L. Nussbaum, *J. Amer. Chem. Soc.* **94**, 3992 (1972).
5. C. L. Harvey, R. Wright, A. F. Cook, D. T. Maichuk, and A. L. Nussbaum, *Biochemistry* **12** (1973).
6. E. Heimer, M. Ahmad, S. Roy, A. Ramel, and A. L. Nussbaum, *J. Amer. Chem. Soc.* **94**, 1707 (1972).
7. S. Åkerfeldt, *Acta Chem. Scand.* **16**, 1897 (1962).
8. M. H. Caruthers, Proceedings of the Sixth Miles International Symposium on Molecular Biology, June 8–9, 1972, Baltimore, MD, in press, The Johns Hopkins University Press, 1973.
9. E. P. Heimer, M. Ahmad, and A. L. Nussbaum, *Biochem. Biophys. Res. Commun.* **48**, 348 (1972).
10. S. A. Narang, K. Itakura, C. P. Bahl, and Y. Y. Wigfield, *Biochem. Biophys. Res. Commun.* **49**, 445 (1972).
11. A. F. Cook, M. J. Holman, and A. L. Nussbaum, *J. Amer. Chem. Soc.* **91**, 6479 (1969), and papers referred to therein.
12. For some other blocking groups fulfilling these requirements, see F. Eckstein, *Chem. Ber.* **100**, 2228 (1967); E. Ohtsuka, M. Ubasawa, and M. Ikehara, *J. Amer. Chem. Soc.* **91**, 1537 (1969); S. A. Narang, O. S. Bhanot, J. Goodchild, R. H. Wightman, and S. K. Dheer, *J. Amer. Chem. Soc.* **94**, 6183 (1972).
13. R. K. Ralph and H. G. Khorana, *J. Amer. Chem. Soc.* **83**, 2926 (1961).
14. S. A. Narang and S. K. Dheer, *Biochemistry* **8**, 3443 (1969).

15. H. SCHALLER, G. WEIMANN, AND H. G. KHORANA, *J. Amer. Chem. Soc.* **85**, 3841 (1963).
16. H. SCHALLER AND H. G. KHORANA, *J. Amer. Chem. Soc.* **85**, 355 (1962).
17. A. HOLÝ AND F. ŠORM, *Coll. Czechoslov. Chem. Commun.* **36**, 3282 (1971). Even their observation is not without precedent: see G. WEIMANN AND H. G. KHORANA, *J. Amer. Chem. Soc.* **84**, 419 (1962).
18. G. M. TENER, *J. Amer. Chem. Soc.* **83**, 159 (1961).
19. E. F. NOWOSWIAT AND M. S. POONIAN, *J. Chromatogr.* **72**, 217 (1972).
20. T. F. GABRIEL AND J. MICHALEWSKY, *J. Chromatogr.* **67**, 309 (1972).
21. N. H. SHERBERG AND S. B. WEISS, *Proc. Nat. Acad. Sci. USA* **69**, 1114 (1972).
22. J. M. ADAMS, P. G. N. JEPPESEN, F. SANGER, AND B. G. BARRELL, *Nature (London)* **223**, 1009 (1969).
23. W. FIERS, R. CONTRERAS, R. DEWACHTER, G. HALGEMAN, J. MERREGAERT, W. MINJOU, AND A. VANDENBERGHE, *Biochimie* **53**, 495 (1970).
24. J. ARGETSINGER STEITZ, *Nature (London)* **224**, 957 (1969).
25. J. L. NICHOLS, *Nature (London)* **225**, 147 (1970).
26. P. G. N. JEPPESEN, *Biochem. J.* **124**, 357 (1971).
27. R. CONTRERAS, A. VANDENBERGHE, G. VOLCKERT, W. MINJOU, AND W. FIERS, *Letters* **24**, 339 (1972).