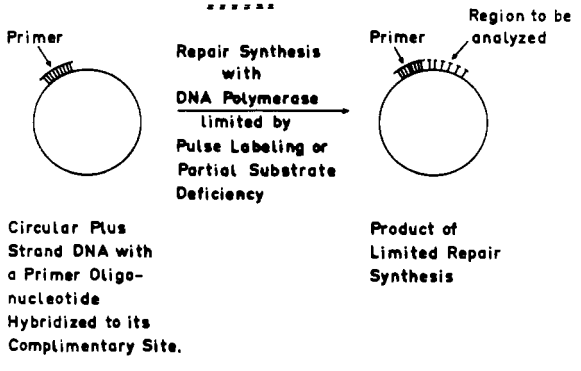
Synthesis of Four Undecanucleotides Complementary to a Region of the Coat Protein Cistron of Phage fd†

Herbert Schott, Dietrich Fischer, and Hans Kössel*

ABSTRACT: From the amino acid sequence AlaTrpMetVal reported for the coat protein of phage fd (Asbeck, F., Beyreuther, K., Köhler, H., Wettstein, G. v., and Braunitzer, G. (1969), *Hoppe-Seylers Z. Physiol. Chem. 350*, 1047) the sequences dA-C-C-A-T-C-C-A (**1**) and dA-C-C-A-T-C-C-A-N-G-C (**2**) (N = A, C, G, or T) can be deduced to occur in the minus-strand DNA of the phage. Owing to the degeneracy of the alanine codons only the ninth position (N) of the undecanucleotide should be ambiguous. Synthesis of the octanucleotide (**1**) and of the four possible undecanucleotides (**2**) is described in this article. A general approach for the sequence determination of single-stranded DNA by the use of specific oligonucleotide primers is proposed, which is based on specific hybridization followed by limited repair synthesis or by pulse labeling during the polymerization reaction in the presence of DNA polymerase.

While powerful methods for the sequence determination of proteins and polyribonucleotides could be developed during the past decades, elucidation of nucleotide sequences from DNA has been achieved with limited success only (for review, see Wu *et al.*, 1972). Recently it has been proposed independently by two other laboratories and by us (Wu *et al.*, 1972; Fischer *et al.*, 1972), to make use of the primer dependency of DNA polymerase in order to obtain specific starting points for this enzyme with specific primer oligonucleotides hybridized to the single-stranded DNA to be sequenced. Thus, from the observation that circular single-stranded DNA, isolated from phage ϕX174, serves as efficient template for DNA polymerase only when oligonucleotide primers of random sequence are provided (Goulian, 1968), it seems likely that a single oligonucleotide primer complimentary to only one distinct region of a given single-stranded DNA would create a single specific starting point for DNA polymerase as schematically represented in Chart I.

Chart I



Information about the sequences occurring at the enzymatically extended 3′ end of the primer (corresponding to the 5′ direction of the complimentary template) should then be deduced from data obtained from pulse-labeling experiments and/or repair synthesis partially limited by the presence of only one, two, or three (instead of all four) of the corresponding deoxyribonucleoside triphosphates. The latter is in general accordance with the methodology developed for sequence determination of the sticky ends of phage λ DNA (Wu and Taylor, 1971). Additional sequence information should be gained from the nearest-neighbor analysis data if the newly synthesized chains are still short enough to give unequivocal results. DNA sequencing by this approach therefore seems feasible as soon as short sequences of single-stranded DNA molecules are already known, to which the complimentary oligonucleotide primers can be synthesized chemically.

In order to verify this possibility, the problem of finding such oligonucleotide sequences has to be solved first. Owing to the degeneracy of the genetic code (Khorana *et al.*, 1966; Nirenberg *et al.*, 1966), unambiguous nucleotide sequences of mRNA and DNA usually cannot be deduced from known amino acid sequences. As an exception, however, only one triplet codes for each of the two amino acids methionine and tryptophan and the ambiguity of corresponding nucleotide sequences is therefore eliminated, where these two amino acids occur adjacent to each other in known amino acid sequences. If unambiguous positions of neighboring codons are included even larger portions of the corresponding DNA can be deduced unequivocally. Thus, from the amino acid sequence AlaTrpMetVal, reported for the coat protein of phage fd (Asbeck *et al.*, 1972), the corresponding octanucleotide sequence dA-C-C-A-T-C-C-A[1] (**1**) can be deduced to occur in the DNA of this phage; as it is known that the minus-strand DNA serves as template for mRNA synthesis (Sugiura *et al.*, 1969), this sequence must occur within the minus strand of fd DNA

†From the Institut für Biologie III (Genetik, Molekularbiologie und Biophysik), Universität Freiburg, Freiburg, West Germany. *Received February 26, 1973*. This is the second paper of a series describing the synthesis of phage-specific DNA fragments. The first paper is by Schott and Kössel (1973). We are indebted to the Deutsche Forschungsgemeinschaft for support of the present work by grants and by awarding a research fellowship to H. S.

[1] The system of abbreviations used is principally as has been suggested by the IUPAC–IUB commission published in *Eur. J. Biochem. 15*, 203 (1970). In this paper the prefix d (for deoxy) in all cases refers to the entire nucleoside residues of the oligonucleotide chains described; for clarity therefore parentheses and hyphens usually following the prefix d are always omitted; thus parentheses are only used to mark hydroxyl protecting groups. TPS refers to 2,4,6-triisopropylbenzenesulfonyl chloride. One A_{260} unit is defined as the amount of nucleotide giving an absorbance of 1 at 260 nm when dissolved in 1 ml of solvent and measured in a 1-cm light-path quartz cell. Different wavelengths are indicated by the respective subscripts.

(Chart II).[2] If the alanine codons with a degeneracy in the third position are to be included, a corresponding undecanucleotide sequence dA-C-C-A-T-C-C-A-N-G-C (2) (N = A, C, G, or T) can be predicted, in which only one position (designated by N) is ambiguous.

Chart II[2]

H_2N - - - - -	Ala Trp Met Val - - - - -	COOH	Coat Protein
5′ - - - - -	GCMUGGAUGGU• - - - - -	3′	mRNA
3′ - - - - -	CGNACCTACCA• - - - - -	5′	Minus Strand DNA
5′ - - - - -	GCMTGGATGGT• - - - - -	3′	Plus Strand DNA

Therefore the octanucleotide (**1**) and one of the four undecanucleotides (**2**) should be able to from fully complimentary hybrids with the plus-strand DNA isolated from phage fd. As circular single-stranded DNA from this phage consists of about 6000 nucleotide residues (Marvin and Hoffmann-Berling, 1963; Eigner and Doty, 1969) and as 4^8 (~66,000) different octanucleotides, respectively, 4^{11} (~4×10^6) different undecanucleotides, are theoretically possible, multiple occurrence of these sequences in the complimentary minus-strand DNA is improbable (from the figures mentioned a statistical probability of $6000:4^8 \sim 1/10$ for the octanucleotide and of $6000:4^{11} \sim 1/700$ can be calculated). The probability that there would be more than one region in the plus-strand DNA, to which the octanucleotide (**1**) or the undecanucleotide(s) (**2**) can be fully hybridized seems therefore very limited. Hence the use of these oligonucleotides as primers for DNA polymerase with fd DNA plus strand as template should result in a unique specific starting point, necessary for sequence determination as outlined above. The chain length, in addition, of eight or eleven nucleotides seems also just long enough to allow the formation of comparatively stable hybrids with the complimentary region of the plus-strand DNA (Wells *et al.*, 1967; Niyogi, 1969), especially as 50 and <54% G·C pairs are expected for the hybrid with the octanucleotide (**1**) and with the undecanucleotides(s) (**2**), respectively. Experiments with the octanucleotide (**1**) have already confirmed these expectations (Fischer *et al.*, 1972) and have thereby allowed the determination of 50 nucleotide residues following the 3′ end of the octanucleotide, when it was used as a primer for DNA polymerase in the presence of f1 DNA as template (Sanger *et al.*, 1973). Although the nucleotide sequence thus derived does not correspond to the amino acid sequence of the coat protein adjacent to the AlaTrpMetVal sequence, as one would have expected, the DNA sequencing approach outlined here seems nevertheless valid in general. Similar experiments using also the undecanucleotides (**2**) as specific primers are therefore in progress.

Synthesis of the Octanucleotide dA-C-C-A-T-C-C-A (**1**) *and of the Undecanucleotides dA-C-C-A-T-C-C-A-N-G-C* (**2**) (*N = A, C, G, T*). The general methodology of connecting preformed oligonucleotide blocks by condensation reactions in the presence of triisopropylbenzenesulfonyl chloride (Kössel *et al.*, 1967a,b; Ohtsuka and Khorana, 1967; Hachmann and Khorana, 1969) was applied for the organic chemical synthesis of the oligomers. As indicated in Chart III the octanucleotide sequence dA-C-C-A-T-C-C-A was first synthesized in the protected from by using two kinds of protected dinucleotide blocks for stepwise addition onto the 3′ end of a protected dinucleoside phosphate. In four individual reactions the protected octanucleotide was then condensed with one of the four different protected trinucleotides dpN′-ibuG-anC-(Ac)[3] to yield the four undecanucleotides in protected form. Finally the protected octanucleotide and the protected undecanucleotides were converted to the free forms (**1** and **2**) by ammonia and mild acid treatment. Reaction conditions and yields of the individual condensation steps are summarized in Table I. As far as the optimal conditions of the reactions concerned, previous observation have been confirmed (Kössel *et al.*, 1967a,b).

CHART III:[3] Synthesis of the Octanucleotide dA-C-C-A-T-C-C-A and of the Undecanucleotides dA-C-C-A-T-C-C-A-N-G-C.

Characterization of the Products by End-Group Labeling and Partial Sequence Determination. In earlier synthetic work one major tool for the characterization of the final products consisted in the determination of the base composition and of the terminal nucleoside residues by means of degradation with phospodiesterases from snake venom and spleen. Several A_{260} units are necessary in order to perform this technique with reasonable accuracy as is also demonstrated in the present work (see Experimental Section and Table III). In addition to this, a recently developed technique for the labeling of the 3′ ends of deoxyoligonucleotides with ^{32}P radioactive ribonucleotide residues in the presence of terminal deoxynucleotidyl transferase (Kössel and Roychoudhury, 1971) was used for the characterization of the oligomers described in this publication. This technique in combination with partial degradation by snake venom phosphodiesterase also allows partial sequence determination (Roychoudhury *et al.*, 1971). The expected sequences for the oligonucleotides synthesized could be confirmed by this technique (see Experimental

[2] M and N refer to the four possible complementary standard nucleotides corresponding to the degenerate position of the alanine codon. The minus-strand sequences, the syntheses of which are described in this publication, are underline by arrows in the 5′ to 3′ direction.

[3] N refers to the four deoxynucleosides A, C, G, and T, respectively. N′ designates deoxythymidine and the three *N*-acyl-deoxynucleosides bzA, anC, and ibuG, respectively.

TABLE I: Summary of Conditions and Yields in the Synthesis of Protected Oligonucleotides.

Condensation Products	3′-OH-Bearing Components (mmol)	5′-Phosphate-Bearing Components (mmol)	TPS (mmol)	Pyridine (ml)	Time (hr)	Yields	
						%	A_{300} Units[a] (%)
d(MeOTr)bzA-anC-anC-bzA-T-anC (I)	d(MeOTr)bzA-anC-anC-bzA	dpT-anC(Ac)					
	0.18	0.60	5	4	8	70	9900
	0.13	0.30	1.2	3	5	40	4100
d(MeOTr)bzA-anC-anC-bzA-T-anC-anC-bzA (II)	I	dpanC-bzA(Ac)					
	0.12	0.60	2.5	4	10	40	5200
d(MeOTr)bzA-anC-anC-bzA-T-anC-anC-bzA-T-ibuG-anC	II	dpT-ibuG-anC(Ac)					
	0.007	0.035	0.3	1.5	10	5–10[b]	170 (18)
d(MeOTr)bzA-anC-anC-bzA-T-anC-anC-bzA-anC-ibuG-anC	II	dpanC-ibuG-anC(Ac)					
	0.007	0.015	0.3	1.5	10	5–10[b]	200 (18)
d(MeOTr)bzA-anC-anC-bzA-T-anC-anC-bzA-bzA-ibuG-anC	II	dpbzA-ibuG-anC(Ac)					
	0.007	0.050	0.3	1.5	10	5–10[b]	210 (21)
d(MeOTr)bzA-anC-anC-bzA-T-anC-anC-bzA-ibuG-ibuG-anC	II	dpibuG-ibuG-anC(Ac)					
	0.007	0.040	0.3	1.5	10	5–10[b]	220 (23)

[a] Before paper chromatographic purification. [b] Owing to losses during further extensive paper chromatographic purification approximate yields only can be given.

Section and Roychoudhury *et al.*, 1971). It is noteworthy that amounts only in the range 10^{-3} to 5×10^{-2} A_{260} units were required, when the technique was carried out on microscale and when the two-dimensional fingerprint method (Brownlee and Sanger, 1969) was applied for the separation of the labeled degradation products (Kössel *et al.*, 1973). From a comparison of the two fingerprint patterns shown in Figures 4 and 5 (see Experimental Section) it is evident that the method is sensitive enough for the discrimination of undecanucleotides differing only in one position.

Experimental Section

The general methods for condensation reactions, for the preparation of protected mononucleotides, dinucleoside monophosphate, and dinucleotides, for removal of protecting groups, and for characterization of the purified compounds by paper chromatography and for enzymic degradation were used essentially according to the general methodology developed by Khorana and his coworkers (Kössel *et al.*, 1967a,b; Ohtsuka and Khorana, 1967; Hachmann and Khorana, 1969; Agarwal *et al.*, 1972) with modifications as described previously (Schott and Kössel, 1973). A summary of conditions and yields in the syntheses of protected oligonucleotides is given in Table I. In order to separate tritylated oligonucleotides from trityl-free oligomers by preparative DEAE-cellulose chromatography, a combination of two subsequent buffer gradients was applied as previously (Schott and Kössel, 1973), the first containing 20% methanol, the second containing 50% ethanol. For identification and characterization of the compounds from various fractions, paper chromatography was performed in the following solvent systems with the descending technique: solvent A, ethanol–1 M ammonium acetate (pH 7.5) (7:3, v/v); solvent B, 1-propanol–concentrated ammonia–water (55:10:35, v/v); solvent C, isobutyric acid–concentrated ammonia–water (66:1:33, v/v). A summary of R_F values of the various protected and unprotected oligonucleotides is given in Table II. Base compositions of the unprotected oligonucleotides as determined by enzymatic digestion in the presence of venom or spleen phosphodiesterase are given in Table III. In all cases less than 10% impurities were found in the unprotected oligonucleotides. For the calculation of yields and of spectral ratios molar extinction values were employed as previously (Schott and Kössel, 1973) neglecting hypochromicity. A summary of observed and calculated ultraviolet (uv) characteristics of the various compounds is given in Table IV. Preparation of the protected di-, tri-, and tetranucleotide blocks was carried out essentially as described previously (Schott and Kössel, 1973).

Preparation of d(MeOTr)bzA-anC-anC-bzA-T-anC. An anhydrous mixture of 10,000 A_{300} units of d(MeOTr)bzA-anC-anC-bzA (0.18 mmol) and of 14,000 A_{300} units of dpT-anC (Ac) (0.6 mmol) in dry pyridine (4 ml) was reacted with TPS (5 mmol) for 8 hr at room temperature. After stopping the reaction and removal of the 3′-*O*-acetyl groups the resulting mixture (150 ml) was subjected to DEAE-cellulose column chromatography according to Figure 1. The column (4.5 cm × 50 cm) was first washed with 0.04 M $Et_3NH_2CO_3$ in 20% methanol until the eluates were free from pyridine. Then elution with a linear gradient was started using 1 l. of 0.04 M $Et_3NH_2CO_3$ in 20% methanol in the mixing vessel and 1 l. of 0.35 M $Et_3NH_2CO_3$ in 20% methanol in the reservoir. After passing 1.3 l. of the gradient, fractions of 15 ml were collected every 5 min. After elution with the remaining gradient mixtures the column was washed with 600 ml of 0.04 M $Et_3NH_2CO_3$ in 20% methanol and then eluted with a second gradient using 1 l. of 0.04 M $Et_3NH_2CO_3$ in 50% ethanol in the mixing vessel and 1 l. of 0.35 M $Et_3NH_2CO_3$ in 50% ethanol in the reservoir. Identification of the two main peaks indicated that peak A and B contained dpT-anC and the desired hexanucleotide

TABLE II: Paper Chromatographic Characterization of Protected and Unprotected Oligonucleotides; R_F Values (I); R_F Values Relative to pT (II); R_F Values Relative to d(A-C-C-A-T-C-C-A) (III).

Compound	Solvent B			Solvent C
	I	II	III	II
d(MeOTr)bzA-anC-anC-bzA-T-anC		0.94[a]		
d(MeOTr)bzA-anC-anC-bzA-T-anC-anC-bzA	0.82			
d(MeOTr)bzA-anC-anC-bzA-T-anC-anC-bzA-T-ibuG-anC	0.74			
d(MeOTr)bzA-anC-anC-bzA-T-anC-anC-bzA-anC-ibuG-anC	0.63			
d(MeOTr)bzA-anC-anC-bzA-T-anC-anC-bzA-ibuG-ibuG-anC	0.69			
d(MeOTr)bzA-anC-anC-bzA-T-anC-anC-bzA-bzA-ibuG-anC	0.62			
d(MeOTr) A-C-C-A-T-C	0.52	1.24		
d(MeOTr)A-C-C-A-T-C-C-A	0.34	0.91		
d(MeOTr)A-C-C-A-T-C-C-A-T-G-C	0.22	0.53		
d(MeOTr)A-C-C-A-T-C-C-A-C-G-C	0.11	0.27		
d(MeOTr)A-C-C-A-T-C-C-A-G-G-C	0.11	0.30		
d(MeOTr)A-C-C-A-T-C-C-A-A-G-C	0.16	0.38		
dA-C-C-A-T-C	0.23	0.43		
dA-C-C-A-T-C-C-A	0.19	0.23		1.13
dA-C-C-A-T-C-C-A-T-G-C			0.25	0.50
dA-C-C-A-T-C-C-A-C-G-C			0.33	0.73
dA-C-C-A-T-C-C-A-G-G-C			0.18	0.47
dA-C-C-A-T-C-C-A-A-G-C			0.40	0.48

[a] Solvent A.

d(MeOTr)bzA-anC-anC-bzA-T-anC, respectively. Material was pooled as shown by the vertical dotted lines.

Preparation of d(MeOTr)bzA-anC-anC-bzA-T-anC-anC-bzA. An anhydrous mixture of 9500 A_{300} units of d(MeOTr)bzA-anC-anC-bzA-T-anC (0.12 mmol) and of 16,900 A_{300} units of dpanC-bzA(Ac) (0.6 mmol) in dry pyridine (4 ml) was reacted with TPS (2.5 mmol) for 10 hr at room temperature. After stopping the reaction and removal of 3′-*O*-acetyl group the resulting mixture (200 ml) was subjected to DEAE-cellulose column chromatography according to Figure 2a.

The column (4.5 cm × 50 cm) was first washed with 0.04 M $Et_3NH_2CO_3$ in 20% methanol until the eluates were free from pyridine. Then elution with a linear gradient was started using 1 l. of 0.04 M $Et_3NH_2CO_3$ in 20% methanol in the mixing vessel am 1 l. of 0.35 M $Et_3NH_2CO_3$ in 20% methanol in the reser-

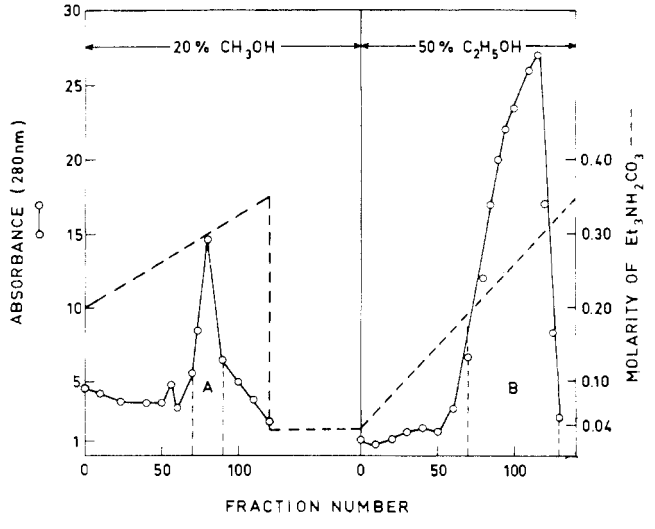


FIGURE 1: Chromatographic spearation of the reaction products resulting from condensation of d(MeOTr)bzA-anC-anC-bzA with dpT-anC(Ac) on a DEAE-cellulose column. Peak B contained the desired hexanucleotide d(MeOTr)bzA-anC-anC-bzA-T-anC. For explanation, see text.

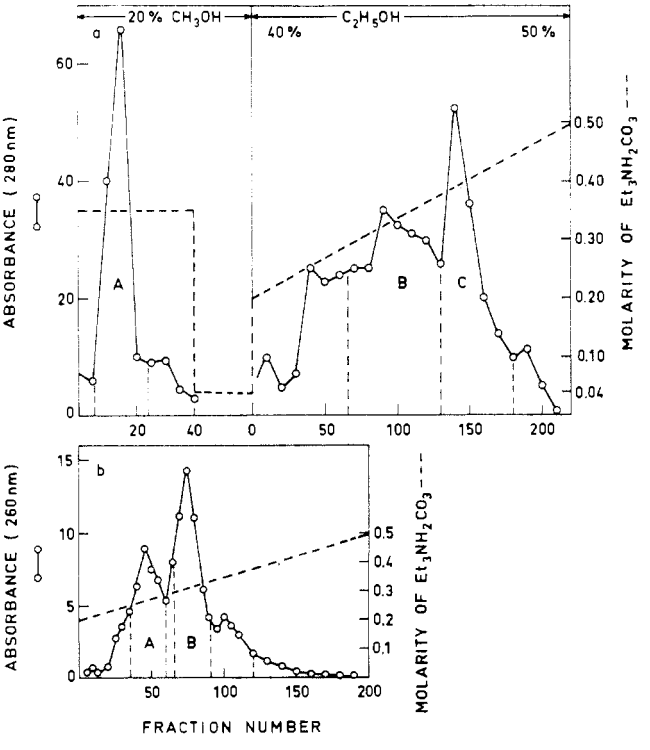


FIGURE 2: Chromatographic separation of the reaction products resulting from condensation of d(MeOTr)bzA-anC-anC-bzA-T-anC with dpanC-bzA(Ac) on a DEAE-cellulose column (a). Part b shows the elution profile of a rechromatography of the material from sections B and C of part a. Peak B in part b contained the desired octanucleotide d(MeOTr)bzA-anC-anC-bzA-T-anC-anC-bzA.

TABLE III: Characterization of Oligonucleotides by Enzymic Degradation.

Compound	Enzyme	A_{260} Units Degraded	Degradation Products	Theoretical Ratio	Ratio Observed
dA-C-C-A-T-C[a]	Spleen	5.0	Ap,Tp,Cp,C	2:1:2:1	2.0:0.88:2.1:1.1
dA-C-C-A-T-C-C-A[a]	Spleen	5.0	Ap,Tp,Cp,A	2:1:4:1	1.73:0.93:4.82:1.00
dA-C-C-A-T-C-C-A-T-G-C	Venom	6.9	pA,pT,pC,pG,A	2:2:5:1:1	2.2:2:5.27:0.9:1
dA-C-C-A-T-C-C-A-A-G-C	Venom	6.3	pA,pT,pC,pG,A	3:1:5:1:1	2.84:1.02:5.3:1:1.08
dA-C-C-A-T-C-C-A-G-G-C	Venom	9.5	pA,pT,pC,pG,A	2:1:5:2:1	2.25:0.91:5:2:1
dA-C-C-A-T-C-C-A-C-G-C	Venom	10.5	pA,pT,pC,pG,A	2:1:6:1:1	2.0:0.85:5.96:1:0.97

[a] This compound was also characterized previously by partial sequence determination (Roychoudhury *et al.*, 1971).

voir. After passage of this gradient, the column was eluted with 1 l. of 0.35 M $Et_3NH_2CO_3$ in 20% methanol and fractions of 15 ml were collected every 4 min. The column was then washed with 350 ml of 0.04 M $Et_3NH_2CO_3$ containing 20% methanol and with 230 ml of 0.04 M $Et_3NH_2CO_3$ containing 40% ethanol. A second linear gradient was then started using 2 l. of 0.2 M $Et_3NH_2CO_3$ in 40% ethanol in the mixing vessel and 2 l. of 0.5 M $Et_3NH_2CO_3$ in 50% ethanol in the reservoir. Fraction of ~20 ml were collected every 7 min. Identification of three main peaks indicated that peak A, B, and C contained dpanC-bzA, d(MeOTr)bzA-anC-anC-bzA-T-anC, and d(MeOTr)bzA-anC-anC-bzA-T-anC-anC-bzA, respectively. The latter two compounds were considerably contaminated with side products; therefore rechromatography of the combined material from peak B and C was carried out according to Figure 2b. The column (3 × 50 cm) was first eluted with 2 l. of 0.2 M $Et_3NH_2CO_3$ in 50% ethanol in the mixing vessel and 2 l. of 0.50 M $Et_3NH_2CO_3$ in 50% ethanol in the reservoir. Fraction of 20 ml were collected every 5 min. Identification of the two main peaks indicated that peak A and B contained d(MeOTr)bzA-anC-anC-bzA-T-anC and the desired octanucleotide d(MeOTr)bzA-anC-anC-bzA-T-anC-anC-bzA, respectively. Material was pooled as shown by vertical dotted lines.

TABLE IV: Calculated and Observed Ultraviolet Characteristics of Protected and Unprotected Oligonucleotides.[a]

Compound	λ_{max} (nm)	λ_{min} (nm)	Absorbance Ratios 250:260 Calcd	250:260 Obsd	260:280 Calcd	260:280 Obsd	280:300 Calcd	280:300 Obsd
d(MeOTr)bzA-anC-anC-bzA-T-anC	282	248			0.72	0.73	1.22	1.20
d(MeOTr)bzA-anC-anC-bzA-T-anC-anC-bzA	212 280	243			0.70	0.82	1.23	1.32
d(MeOTr)bzA-anC-anC-bzA-T-anC-anC-bzA-T-ibuG-anC	217 279	238			0.78	0.82	1.25	1.34
d(MeOTr)bzA-anC-anC-bzA-T-anC-anC-bzA-ibuG-ibuG-anC	214 260 276	237 270			0.85	1.00	1.17	1.40
d(MeOTr)bzA-anC-anC-bzA-T-anC-anC-bzA-anC-ibuG-anC	210 280	243			0.75	0.84	1.15	1.31
d(MeOTr)bzA-anC-anC-bzA-T-anC-anC-bzA-bzA-ibuG-anC	214 275	238			0.68	0.99	1.28	1.42
dA-C-C-A-T-C	261	210 230	0.79	0.83	1.88	2.03		
d(MeOTr)A-C-C-A-T-C	262	208 228	0.79	0.87	1.88	1.88		
d-A-C-C-A-T-C-C-A	261	231	0.79	0.85	1.99	2.08		
dA-C-C-A-T-C-C-A-T-G-C	208 261	231	0.82	0.87	1.75	1.86		
dA-C-C-A-T-C-C-A-A-G-C	212 261	229	0.82	0.86	1.99	1.92		
dA-C-C-A-T-C-C-A-G-G-C	210 258	227	0.86	0.96	1.75	1.78		
dA-C-C-A-T-C-C-A-C-G-C	214 261	228	0.83	0.90	1.71	1.75		

[a] Measurements were taken for the protected oligonucleotides in 0.04 M TEAB (pH 7.2–7.6) containing 20% methanol. The unprotected oligonucleotides were measured in 0.1 M Tris-HCl buffer (pH 7.0) without alcohol.

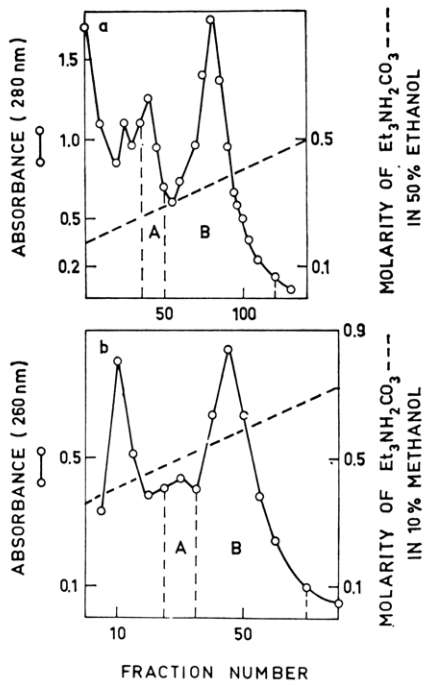


FIGURE 3: Chromatographic separation of the reaction products resulting from condensation of d(MeOTr)bzA-anC-anC-bzA-T-anC-anC-bzA with dpbzA-ibuG-anC(Ac) on a DEAE-cellulose column (a). Part b shows the elution profile of the ammonia and acid-treated oligonucleotides from section B of part a. Peak B of part b contained the desired undecanucleotide (2) (N = A). For explanation, see text.

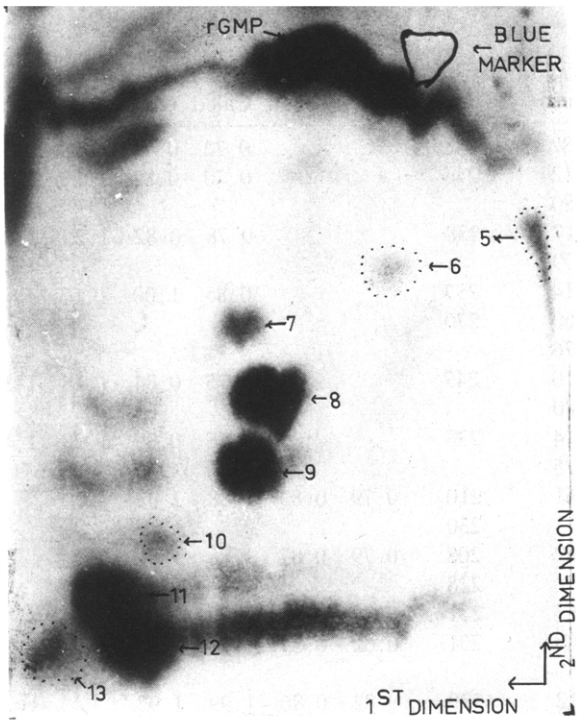


FIGURE 4: Partial sequence analysis of dA-C-C-A-T-T-C-A-T-G-C. The oligomer mixture resulting from partial digestion of the undecanucleotide with snake venom phosphodiesterase was subjected to the 3′-terminal labeling with [α-^{32}P]rGTP in the presence of terminal deoxynucleotidyl transferase. Two-dimensional separation of the labeled oligonucleotide mixture was then carried out by ionophoresis on cellulose acetate at pH 3.5 in the first dimension and by thin-layer chromatography on DEAE-cellulose in the second dimension. For explanation, see text.

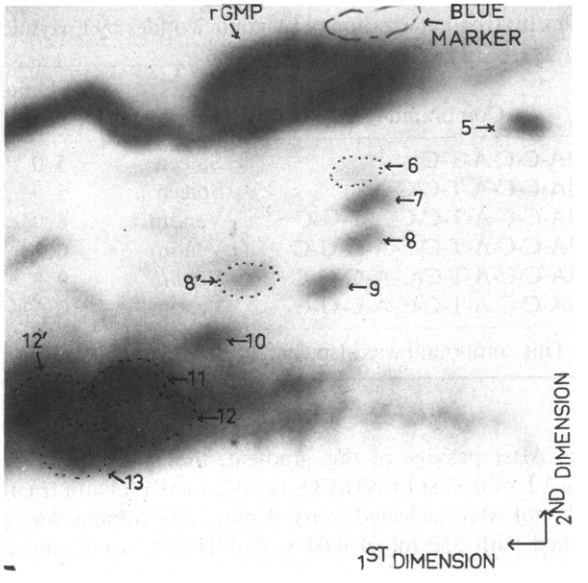


FIGURE 5: Partial sequence analysis of dA-C-C-A-T-C-C-A-T-G-C. The same procedure as described in the legend to Figure 4 was applied. For explanation, see text.

Preparation of Protected Undecanucleotides. Anhydrous mixtures containing 750 A_{300} units of protected octanucleotide d(MeOTr)bzA-anC-anC-bzA-T-anC-anC-bzA (0.007 mmol) and 700–1500 A_{300} units of the respective protected trinucleotide (0.02–0.05 mmol) in dry pyridine (1.5 ml) were reacted with TPS (0.3 mmol) for 10 hr at room temperature. A summary of conditions and yields in the synthesis of protected undecanucleotides is given in Table I. After stopping the reaction the resulting mixtures were subjected to DEAE-cellulose column chromatography, which is shown in Figure 3 for one of the four undecanucleotides. The individual columns (3 cm × 50 cm) were first eluted with a linear buffer gradient using 2 l. of 0.04 M $Et_3NH_2CO_3$ in 50% ethanol in the mixing vessel and 2 l. of 0.5 M $Et_3NH_2CO_3$ in 50% ethanol in the reservoir. After elution of pyridine, fractions of 15–20 ml were collected every 17 min. Characterization of the respective sections A and B indicated that sections A contained mostly trityl-negative products where as trityl-positive products were predominant in sections B (see Figure 3a). Sections B were pooled as shown by the vertical dotted lines and the N-protecting groups and the monomethoxytrityl groups were removed according to the general procedure. The unprotected products were then rechromatographed on second DEAE-cellulose columns (3 cm × 30) using 1 l. of 0.35 M $Et_3NH_2CO_3$ in 20% methanol in the mixing vessel and 1 l. of 0.9 M $Et_3NH_2CO_3$ in 10% methanol in the reservoir. Fractions of 17 ml were collected every 6 min. Sections A contained mainly the unprotected octanucleotide dA-C-C-A-T-C-C-A. Sections B in all cases contained more than 60% of the respective undecanucleotide (Figure 3b) and were pooled as shown by the vertical dotted line. Final purification of the undecanucleotides was achieved by preparative paper chromatography in solvent B and C.

Characterization of Oligonucleotides by Partial Sequence Determination. Partial sequence determination of the undecanucleotides was carried out essentially as described previously (Kössel *et al.*, 1973). dA-C-C-A-T-T-C-A-T-G-C (0.14 A_{260} unit; 127 pmol) (Schott and Kössel, 1973) was partially degraded with 0.04 μg of snake venom phospho-

diesterase in a final volume of 6 μl under the conditions described (Kössel *et al.*, 1973). After inactivation of the phosphodiesterase by heat 14 pmol of [α-^{32}P]GTP (1000 Ci/mol) and 0.8 μg of terminal deoxynucleotidyl transferase were added (final volume 10 μl). Further processing of the reaction was carried out as previously (Kössel *et al.*, 1973), the resulting fingerprint pattern shown in Figure 4 confirms the expected sequence according to the fingerprint rules (Sanger *et al.*, 1973). In a control experiment, in which the treatment with snake venom phosphodiesterase was omitted, only spots 12 and 13 were observed as primer-dependent products, which correspond to d(A-C-C-A-T-T-C-A-T-G-C)-[^{32}P]pGr and d(A-C-C-A-T-T-C-A-T-G-C)-[^{32}P]pGr[^{32}P]pGr, respectively. Material from the marked spots after elution from the corresponding regions of the cellulose thin layer was digested with spleen phosphodiesterase to determine the nearest neighbor of the radioactive guanylic acid residues added at the 3′ termini. The resulting 3′-mononucleotides were then separated by ionophoresis at pH 3.5 as described previously (Kössel and Roychoudhury, 1971; Kössel *et al.*, 1973).

Thus the original 3′ termini of material from spots 5, 6, 7, 8, 9, 10, 11, and 12 of Figure 4 were identified as A, T, T, C, A, T, G, and C, respectively, which is in agreement with the expected sequence. In the case of dA-C-C-A-T-C-C-A-T-G-C- identical conditions were applied except that 0.024 A_{260} unit of the undecanucleotide was reacted and that 42 pmol of [α-^{32}P]rGTP was added for the terminal addition reaction. The resulting fingerprint is shown in Figure 5. Again material from all primer-dependent spots was subjected to nearest-neighbor analysis and yielded A, C, C, A, G, and C residues for spots 5, 7, 8, 9, 11, and 12, respectively. Owing to very little radioactivity material from spots 6 and 10 could not be assigned to a distinct degradation product. The fingerprint pattern as well as the result of the nearest-neighbor analysis (with the two exceptions) confirm the expected sequence and in comparison to Figure 4 allow the discrimination of a single nucleotide exchange.

Analogous fingerprint patterns and nearest-neighbor results were obtained with the other oligomers described in this publication.

Acknowledgment

We thank Dr. G. Hobom for stimulating discussions during the course of the work. We are also indebted to Drs. Härle and G. Hobom for helpful criticism of the manuscript. Valuable technical help is credited to Ms. S. Baars and D. v. Nell.

References

Agarwal, K. L., Yamazaki, A., Cashion, P. J., and Khorana, H. G. (1972), *Angew. Chem., Int. Ed. Engl. 11*, 451.

Asbeck, F., Beyreuther, K., Köhler, H., Wettstein, G. v., Braunitzer, G. (1969), *Hoppe-Seylers Z. Physiol. Chem. 350*, 1047.

Brownlee, G. G., and Sanger, F. (1969), *Eur. J. Biochem. 11*, 395.

Eigner, J., and Doty, P. (1969), *J. Mol. Biol. 12*, 549.

Fischer, D., Donelson, J., Ling, V., and Kössel, H. (1972), *Hoppe-Seylers Z. Physiol. Chem. 353*, 704.

Goulian, M. (1968), *Proc. Nat. Acad. Sci. U. S. 61*, 284.

Hachmann, J., and Khorana, H. G. (1969), *J. Amer. Chem. Soc. 91*, 2749.

Khorana, H. G., Büchi, H., Ghosh, H., Gupta, N., Jacob, T. M., Kössel, H., Morgan, R., Narang, S. A., Ohtsuka, E., and Wells, R. D. (1966), *Cold Spring Harbor Symp. Quant. Biol. 31*, 39.

Kössel, H., Büchi, H., and Khorana, H. G. (1967b), *J. Amer. Chem. Soc. 89*, 2185.

Kössel, H., Moon, M. W., and Khorana, H. G. (1967a), *J. Amer. Chem. Soc. 89*, 2148.

Kössel, H., and Roychoudhury, R. (1971), *Eur. J. Biochem. 22*, 271.

Kössel, H., Roychoudhury, R., Fischer, D., and Otto, A. (1973), *Methods Enzymol.* (in press).

Marvin, A., and Hoffmann-Berling, H. (1963), *Nature (London) 197*, 517.

Nirenberg, M., Caskey, T., Marshall, R., Brimacombe, R., Kellog, D., Doctor, B., Hatfield, D., Levin, J., Rottman, Pestka, S., Wilcox, M., and Anderson, F. (1966), *Cold Spring Harbor Symp. Quant. Biol. 31*, 11.

Niyogi, S. K. (1969), *J. Biol. Chem. 244*, 1576.

Ohtsuka, E., and Khorana, H. G. (1967), *J. Amer. Chem. Soc. 89*, 2195.

Roychoudhury, R., Fischer, D., and Kössel, H. (1971), *Biochem. Biophys. Res. Commun. 45*, 430.

Sanger, F., Donelson, J., Coulson, A., Kössel, H., and Fischer, D. (1973), *Proc. Nat. Acad. Sci. U. S. 70*, 1209.

Schott, H., and Kössel, H. (1973), *J. Amer. Chem. Soc. 95*, 3778.

Sugiura, M., Okamoto, R., and Takanami, M. (1969), *J. Mol. Biol. 43*, 299.

Wells, R. D., Büchi, H., Kössel, H., Ohtsuka, E., Khorana, H. G. (1967), *J. Mol. Biol. 27*, 265.

Wu, R., Donelson, J., Padmanabhan, R., and Hamilton, R. (1972), *Bull. Inst. Pasteur Paris 70*, 203.

Wu, R., and Taylor, R. (1971), *J. Mol. Biol. 57*, 491.