Volume 11 Number 2 1983 Nucleic Acids Research

Solid phase synthesis of polynucleotides. VIII. Synthesis of mixed oligodeoxyribonucleotides by the phosphotriester solid phase method[1]

Yoshimasa Ike, Satoshi Ikuta, Moriyuki Sato, Ting Huang and Keiichi Itakura

Department of Molecular Genetics, City of Hope Research Institute, Duarte, CA 91010, USA

Received 16 November 1982; Accepted 20 December 1982

ABSTRACT

A solid phase method for the simultaneous synthesis of mixed oligonucleotides using a phosphotriester approach has been developed. For this synthesis, a mixture of mono or dimeric coupling units is used, and a slight difference in the reactivity of those units is found. However, this difference does not hamper the simultaneous, mixed oligonucleotide synthesis, and the sequence analysis of a product demonstrates the existence of all desired sequences in the final mixture.

INTRODUCTION

A method for the isolation of specific cloned DNA sequences has been developed using advantages of the hydridization properties of oligodeoxyribonucleotides (2,3,4). Our general approach is to chemically synthesize a mixture of oligonucleotides which represent all possible codon combinations of a short peptide sequence of a given protein, and to use the mixed oligonucleotides as hybridization probes. One oligonucleotide sequence within this mixture must be complementary to the DNA coding for the short peptide. This oligonucleotide, therefore, will form a perfectly base paired duplex with the DNA, whereas the other oligonucleotides in the mixture will form mismatched base paired duplexes. Under highly stringent conditions only the perfectly matched duplex forms, and the formation of the mismatched duplexes is virtually eliminated (5). This specific property of the oligonucleotide hybridization allows us to use a mixture of oligonucleotides as a specific hybridization probe for the screening of cloned DNA sequences.

We have reported a solid phase method for synthesizing several oligonucleotides simultaneously, using the phosphotriester approach (5). In the synthesis we used a mixture of trinucleotide coupling units that have a different base in the middle to ensure a similar reactivity of each trimer block. The recent simplification of the solid phase method permits molecular biologists to easily synthesize oligonucleotides using commercially available

0305-1048/83/1102-0477$2.00/0

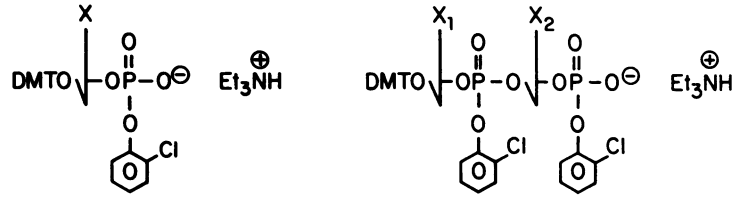
nucleotides, reagents, and solvents. We have re-examined the simultaneous synthesis of mixed oligonucleotide sequences using these commercially available mono and dinucleotides.

RESULTS AND DISCUSSION

For the simultaneous synthesis of mixed oligonucleotide sequences using mono or dimer coupling units, the reactivity of those different coupling units must be identical, or at least similar, to each other. If there was a significant difference in the reactivity, the final mixture would not contain all desired oligonucleotide sequences. Consequently, we first examined the reactivity of mono 1 and dimer 2 coupling units (Figure 1).

The trinucleotide (TTT) bound to the polystyrene resin reacted with a mixture of equimolar amounts of A (1a) and G (1b), T (1c) and C (1d), and all four mononucleotides, respectively. All of these resultant tetranucleotides were further reacted with the monomer coupling unit T (1c) to give pentanucleotides. After removal of all protecting groups by the published procedure (1), the mixtures were analyzed by high performance liquid chromatography (HPLC) on a μBondapak C_{18} column (Figures 2A, B, and C).

The relative reactivity of four kinds of mononucleotides 1 to form an internucleotidic phosphotriester bond was determined by comparison of the extinction coefficient (254 nm) and the peak area of the pentanucleotides (Table I). These results demonstrate that the reactivity between mononucleotides with pyrimidine bases (C and T) is identical, and that those with purine bases (A and G) have a similar reactivity. However, the reactivity of mononucleotides with purine bases is lower than that with pyrimidine bases when the coupling reaction mixture contains four mononucleotides at the same time. This data does not agree with the results that the mononucleotides have



1 2

X = A^{Bz} (a) G^{iBu} (b) C^{Bz} (c) T (d)

DMT = 4,4′ - dimethoxytrityl

FIGURE 1. Coupling units for synthesizing mixed oligonucleotide sequences.

the same reactivity in the phosphite-triester approach (6). Obviously, this discrepancy reflects quite a different reactivity of coupling units between phosphotriester and phosphite-triester methods. They are more reactive in the latter method and, therefore, no selectivity could be found.

The difference in the yields of the pentanucleotides is most likely due to

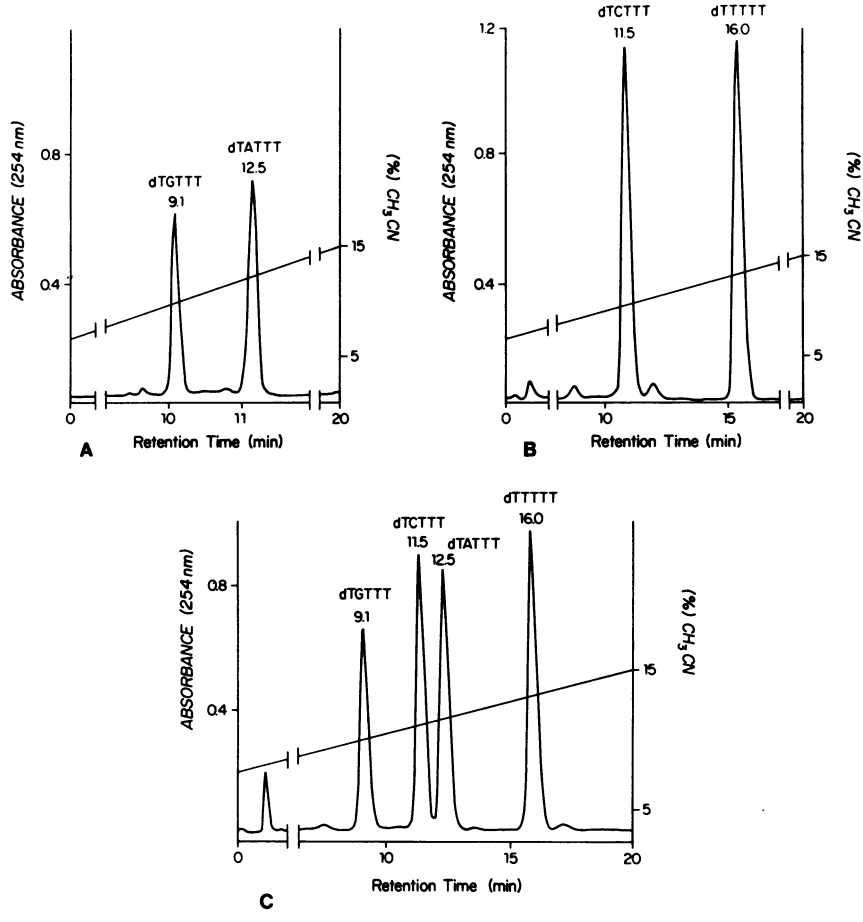


FIGURE 2. Separation of mixed pentanucleotides (dTXTTT) by HPLC. X = A, G, C, and T. A linear gradient of CH_3CN (4-7%) for 20 minutes was used for this analysis:

A. The pentanucleotides produced by competition between mononucleotides with adenine (1a) and guanine (1b) bases. The first peak, with a retention time of 9.1 minutes, corresponds to dTGTTT, and the second (RT = 12.5 minutes) corresponds to dTATTT.

B. Competition between cytosine (1c) and thymine (1d). The first peak (RT = 11.5 miutes) is dTCTTT and second peak (RT = 6.0 minutes) is dTTTTT.

C. Competition among four bases.

TABLE I. Relative Reactivity of Four Kinds of Mononucleotides.

Solid Support	Mixture of Mononu-cleotides	Product[a]	Extinction Coefficients[b] $\varepsilon_{254}(X10^4)$	Relative Peak Area[c]	Relative Reactivity
HO-TTT-P	A + G	TATTT	4.32	8.4	1.0
		TGTTT	4.17	7.2	0.9
	C + T	TCTTT	3.53	14.0	1.0
		TTTTT	3.60	14.3	1.0
	A+G+C+T	TATTT	4.32	10.4	0.7
		TGTTT	4.17	8.1	0.6
		TCTTT	3.53	11.1	1.0
		TTTTT	3.60	12.0	1.0

[a]Pentamers are analyzed after further coupling of tetramers with the mononucleotides (T).
[b]Base stacking effect is ignored for the estimation of extinction coefficient which is done simply by the addition of the extinction coefficient of each base.
[c]Relative peak areas are calculated from the peaks intergrated in HPLC profile.

the competitive coupling reaction of four different mononucleotides with the trinucleotide bound to the resin. The following coupling reaction between a single mononucleotide (T) and a mixture of four different tetranucleotides bound to the support should have no competition because an excess of the mononucleotide is used to drive the reaction to completion.

When dinucleotides are used for the simultaneous synthesis of a mixed oligonucleotide probe, a mixture of dinucleotides with a different base either at the 3'- or 5'-position is coupled to the growing chain of oligonucleotides.

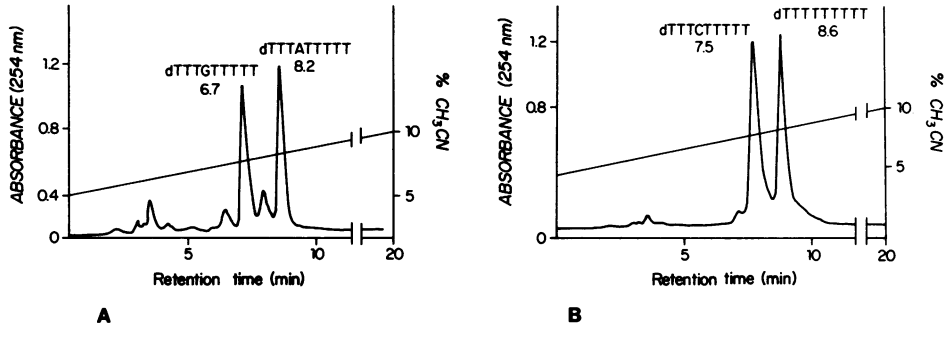


FIGURE 3. Separation of mixed nonanucleotides (dTTTXTTTTT) by HPLC X = A, G, C and T. A linear gradient of CH_3CN (5-10%) was used for this analysis:
A. The first peak is dTTTGTTTTT (RT = 6.7 minutes) and the second peak is dTTTATTTTT (RT = 8.2 minutes).
B. The first peak is dTTTCTTTTT (RT = 7.5 minutes) and the second peak is dTTTTTTTTT (RT = 8.6 minutes).

In the preliminary experiment, we compared the reactivity of dinucleotides with a different base at the 3'-position. As shown in Figures 3A and B, and Table II, the relative reactivity of dinucleotides with a different base at the 3'-position is similar in the pyrimidine and purine series, respectively.

Assuming that the reactivity of dinucleotides with a different base at the 5'-position should be identical, the following mixed sequences are synthesized:

Amino Acid Sequence: Met Asn Asn Asp Thr

mRNA Sequence: 5' A U G A A U A A U G A U A C X 3'
C C C
(X = A, G, C and U)

Mixed Probe: 3' T A C T T A T T A C T A T G 5'
G G G

Synthetic Strategy: 5' GT- $(\frac{AT}{GT})$-$(\frac{CA}{CG})$-TT-$(\frac{AT}{GT})$-TC-A-T-(Ps) 3'
T-(Ps) = Polystyrene resin with thymidine

The polystyrene solid support loaded with thymidine (20 mg, 0.1 mmole/g) is coupled first with the mononucleotide (1a, A), and then the following dinucleotides are successively used for the simultaneous synthesis of eight kinds of 14-mers; TC, a mixture of equimolar amounts of AT and GT, TT, a pair of CA and CG, a mixture of AT and GT, and finally, GT. After removal of all protecting groups except the DMT group, the mixture is analyzed by HPLC on a μBondapak C_{18} column to give two peaks (peaks I and II, Figure 4A). The DMT group is removed and each peak is further analyzed on the same column to resolve into three peaks, 1, 2, and 3, respectively (Figures 4B and C). Oligonucleotides in these peaks are phosphorylated at the 5'-hydroxyl group by T4 polynucleotide kinase with [γ-^{32}P]ATP, and their sequences are determined by the standard method (7) (Figure 5).

These sequence data reveal that the final mixture contains all eight

TABLE II. Relative Reactivity of Dinucleotides.

Solid Support	Mixture of dinucleotide	Product[a] TTTXTTTTT	Extinction Coefficient ε_{254} (X10^4)	Relative Peak Area	Relative Reactivity
HOTTTTT-P	TA + TG	TTTATTTTT	7.20	6.1	1.0
		TTTGTTTTT	7.05	5.5	0.9
	TC + TT	TTTCTTTTT	6.41	7.2	1.0
		TTTTTTTTT	6.48	7.4	1.0

[a]Nonamers are analyzed after further coupling of heptamers with the dinucleotide (TT).

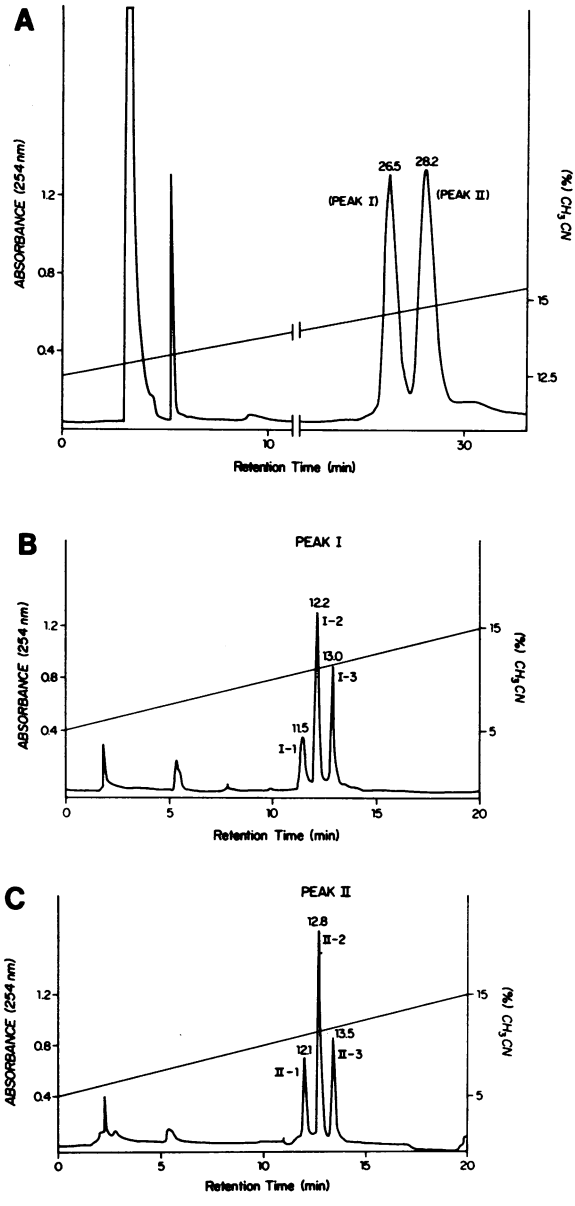


FIGURE 4. HPLC analysis of the mixed oligonucleotide probe.

A: Analysis of oligonucleotides with the DMT group. The separation was carried out by using a linear gradient of CH_3CN (12.5%-15%) for 40 minutes on two connecting columns. The DMT groups of peaks I and II were removed and the products were further analyzed.

B. and C. Analysis of oligonucleotides after removal of the DMT groups. A linear gradient of CH_3CN (5%-15%) for 20 minutes was used for each separation on one column.

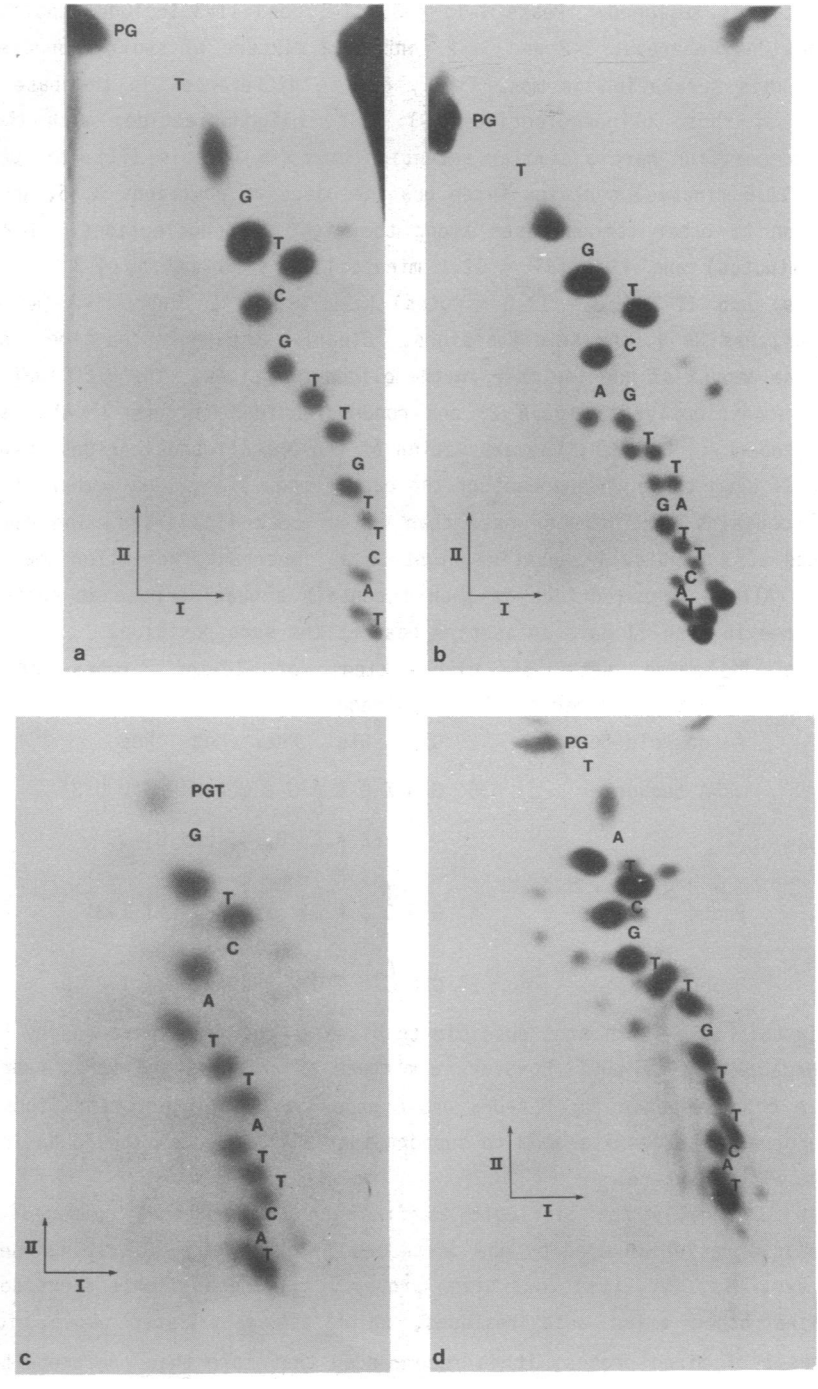
sequences as designed. Peaks I-1, I-3, II-1, and II-3 include single oligonucleotides, whereas, I-2 and II-2 contain a mixture of two oligonucleotides (8). This separation is most likely due to differences in the base composition of those oligonucleotides (9). The oligonucleotides with the same base composition have a similar retention time (RT) (Table III); the peak I-1 (RT = 11.5 minutes) contains three guanine bases at positions 3, 6, and 9, in addition to other common bases among the eight oligonucleotides. I-2 (RT = 12.2 minutes) and II-1 (RT = 12.1 minutes) have 2G and A, I-3 (RT = 13.0 minutes) and II-2 (RT = 12.8 minutes) have 1G and 2A, and II-3 (RT = 13.5 minutes) has 3A at the same positions. Clearly, the retention time increases with the amount of adenine base in the oligonucleotides. The HPLC analysis of the pentanucleotides (Figure 2) and nonanucleotides (Figure 3) also support this tendency. However, the separation of the DMT-oligonucleotides into peaks I and II seem to be independent of the base composition. One member of peak I (I-3) contains more adenine base than one of peak II (II-1). The base difference at a particular position might be an important factor for the separation. All oligonucleotides in peak I contain a guanine base at position 3, and those in peak II have an adenine base at the same position.

The following sets of eight kinds of 14-mer probes are also synthesized by the monomer coupling approach:

Amino Acid Sequence	Gly Glu Phe Met Phe
mRNA Sequence	5' G G X G A A/G U U U/C A U G U U U/C 3'
	(X = A, G, C, and U)
Probe I	5' G G T/C G A A/G T T T/C A T G T T 3'
Probe II	5' G G A/G G A A/G T T T/C A T G T T 3'

In this case, it is not possible to dissolve the final mixtures by HPLC on a μBondapak C_{18} column. However, a mixture of probes I and II is hybridized to the cDNA sequence for HLA-DRα chain under very stringent conditions (10). Therefore, these results seem to support that all sequences should be included in the final mixture.

It is usually not difficult to find peptide sequences composed of 4-5 amino acid residues coded by one or two codons (Met, Trp, Glu, Gln, Phe, Asp, Asn, Cys, His, Tyr, Lys) in a given protein. If the available short sequence contains other amino acid residues, which are not listed above, for the synthesis of mixed probes, it is recommended that more than one probe be pre-


PG
T
G
T
C
G
T
T
G
T
T
C
A
T
II
I
a
PG
T
G
T
C
A G
T
T
G A
T
T
C
A
T
II
I
b
PGT
G
T
C
A
T
T
A
T
T
C
A
T
II
I
c
PG
T
A
T
C
G
T
T
G
T
T
C
A
T
II
I
d

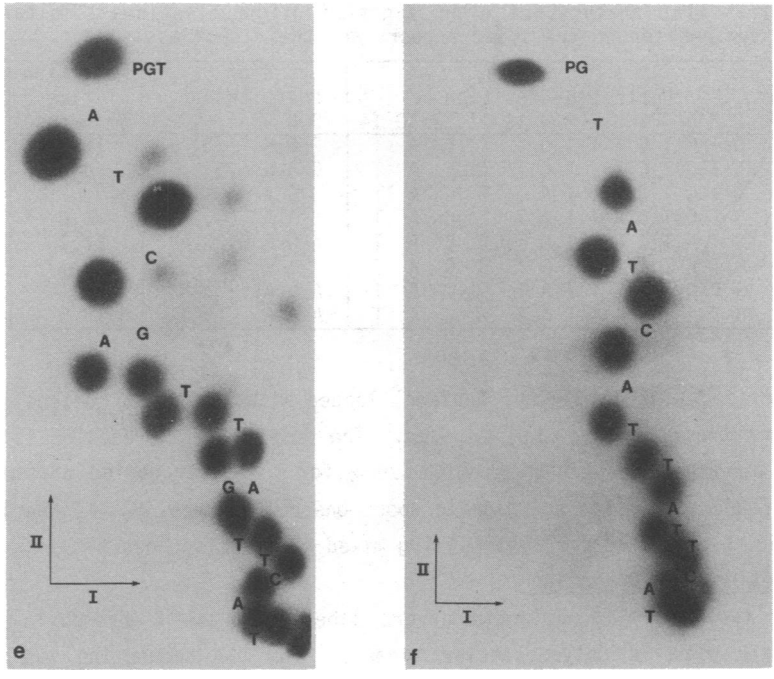


FIGURE 5. DNA sequence analysis of the mixed probe.
a) dGTGTCGTTGTTCAT (I-1).
b) A mixture of dGTGTCATTGTTCAT and dGTGTCGTTATTCAT (I-2).
c) dGTGTCATTATTCAT (I-3).
d) dGTATCGTTGTTCAT (II-1).
e) A mixture of dGTATCATTGTTCAT and dGTATCGTTATTCAT (II-2).
f) dGTATCATTATTCAT (II-3).

pared to ensure the existence of all sequences in the probe. Thus, each mixed probe is separately synthesized and purified, and then used for the hybridization after mixing together (10) or independently (2).

MATERIALS AND METHODS

Unless otherwise mentioned, materials and methods are the same as previously described (1,11). All reagents for the synthesis are commercially available (Bachem, Inc. and Vega Biologicals).

Assembly of Oligonucleotides

The solid phase method used is essentially the same as reported previously (1,11) with the following modifications: the concentration of the coupling units (mononucleotides 1 and dinucleotides 2) is kept at 0.1 M in pyridine; the coupling reaction time is 30 minutes at 37°C, and 20 mg of the polystyrene

TABLE III. HPLC analysis of eight kinds of oligonucleotides simultaneously synthesized on the solid support and their peak assignment.

Peak	Nucleotide Position 1 2 3 4 5 6 7 8 9 10 11 12 13 14	Bases of Positions at 3 6 9	Retention time (minutes)
I-1	G T G T C T T T G T T C A T	G G G (3G)	11.5
I-2	G T G T C A T T C T T C A T	G A G (2G.A)	12.2
	G T G T C G T T A T T C A T	G G A	
I-3	G T C T C A T T A T T C A T	G A A (G.2A)	13.0
II-1	G T A T C G T T G T T C A T	A G G (2G.A)	12.1
II-2	G T A T C A T T G T T C A T	A A G (G.2A)	12.8
	G T A T C G T T A T T C A T	A G A	
II-3	G T A T C A T T A T T C A T	A A A (3A)	13.5

copolymerized with 1% divinyl benzene, loaded with the first nucleoside (~100 μmole per gram of the resin) is used. The detritylation reaction is carried out by repeating the treatment with $ZnBr_2$ for a shorter period of time. One cycle of the nucleotide addition is about one hour (Table IV). Table IV shows the general procedure for synthesizing mixed oligonucleotides.

^{32}P-Labeling and Sequencing

The isolated oligonucleotides were labeled at the 5'-hydroxyl group by [γ-^{32}P]ATP with T4 polynucleotide kinase. The DNA sequencing was done as

TABLE IV. Synthetic Operation.

STEP	REAGENT OR SOLVENT	REACTION TIME	NO. OF OPERATIONS
1	1M $ZnBr_2/CH_2Cl_2$-IPA	2 min.	2 or 4[a]
2	0.5M $Et_3N^+HAcO^-$/DMF	10 sec.	3
3	Pyridine	10 sec.	2
4	THF	10 sec.	2
5	Vacuum	5 min.	1
6	Coupling Mixture (0.1M)[b] at 37°C	30 min.	1
7	Pyridine	10 sec.	1
8	$(CH_3CO)_2O$ in Pyridine	5 min.	1
9	Pyridine	10 sec.	2
10	CH_2Cl_2-IPA	10 sec.	3

[a]Operation time is dependent upon the base of the nucleotide at the 5' end from which the DMT group is removed. In case of adenine and guanine, two treatments are enough and for cytosine and thymine four operations are required.

[b]A mixture of equimolar amounts of mono- or dinucleotides was used to make the coupling solution 0.1 M for the mixed synthesis.

published (7).

HPLC Analysis

HPLC analysis was performed on a μBondapak C_{18} column (3.9 mm X 30 cm) using the Model 6000A solvent delivery system regulated by a Model 660 solvent programmer (Waters Assoc.). The column temperature was maintained at 55°C in a constant temperature oven (Spectra-Physics). A linear gradient of CH_3CN in 0.01 M ethylenediammonium acetate (EDAA) buffer (pH 7.5) was used for the elution of oligonucleotides. The flow rate of the buffer remained at 2 ml/min.

Oligonucleotide samples for the analysis were prepared as follows. After the synthesis and removal of the protecting groups, except the 5'-hydroxyl function, the DMT oligonucleotides were dissolved in 0.05 M triethylammonium bicarbonate (TEAB) buffer, extracted with ether 5 times, and applied to the HPLC system to separate the DMT-oligonucleotides from truncated non-DMT-oligonucleotides. In the case of pentamers and nonamers, resolution of the DMT-oligonucleotides into components was not attempted.

ACKNOWLEDGMENT

This research is supported by USPHS grants GM28651 and GM31259 (K.I.). K.I. is a member of the Cancer Research Center (CA16434) at the City of Hope Research Institute.

REFERENCES

1. No. VII in the series: Tan, Z-K., Ikuta, S., Huang, T., Dugaiczyk, A. and Itakura, K. (1982) XLVII Cold Spring Harbor Quant. Biol. Symp., in press.
2. Suggs, S.V., Wallace, R.B., Hirose, T., Kawashima, E.H. and Itakura, K. (1981) *Proc. Natl. Acad. Sci. U.S.A.*, 78, 6613-6617.
3. Reyes, A.A., Johnson, M.J., Schold, M., Ito, H., Ike, Y., Morin, C., Itakura, K. and Wallace, R.B. (1981) *Immunogenetics*, 14, 383-392.
4. Itakura, K., Miyake, T., Kawashima E.H., Ike, Y., Ito, H., Morin, C., Reyes, A.A., Johnson, M.J., Schold, M., and Wallace, R.B. (1982) Recombinant DNA, Proceedings of the Third Cleveland Symposium on Macromolecules, A.G. Walton (Ed.); Elsevier Scientific, Amsterdam. pp. 273-289.
5. Wallace, R.B., Johnson, M.J., Hirose, T., Miyake, T., Kamashima, E.H. and Itakura, K. (1981) *Nucl. Acids Res.*, 9, 879-894.
6. Elmblad, A., Josephson, S. and Palm, G. (1982) *Nucl. Acids Res.*, 10, 3291-3301.
7. Tu, C-P.D. and Wu, R. (1980) *Methods in Enzymology*, 65, 620-638.

8. The ratio of these oligonucleotides in those peaks is variable in different experiments because the separation of peaks I and II is not perfect, and direct analysis of the mixed sequences after removal of the DMT group is not possible.
9. Fritz, H-J., Belagaje, R., Brown, E.L. Fritz, R.H., Jones, R.A., Lees, R.G. and Khorana, H.G. (1978) *Biochemistry*, 17, 1257-1267.
10. Our unpublished results.
11. Ito, H., Ike, Y., Ikuta, S. and Itakura, K. (1982) *Nucl. Acids Res.*, 10, 1755-1769.