THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 251, No. 3, Issue of February 10, pp. 642-650, 1976
Printed in U.S.A.

Total Synthesis of the Structural Gene for the Precursor of a Tyrosine Suppressor Transfer RNA from *Escherichia coli*

8. ENZYMATIC JOINING OF THE CHEMICALLY SYNTHESIZED SEGMENTS TO FORM DNA DUPLEXES CORRESPONDING TO NUCLEOTIDE SEQUENCES 23-60 AND 23-66*

(Received for publication, March 19, 1975)

PETER C. LOEWEN,‡ ROBERT C. MILLER,§ AMOS PANET,¶ TAKAO SEKIYA, AND H. GOBIND KHORANA

From the Departments of Biology and Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, and the Institute for Enzyme Research of the University of Wisconsin, Madison, Wisconsin 53706

Polynucleotide ligase-catalyzed joining of the eight chemically synthesized deoxyribopolynucleotide segments (Fig. 1) comprising the nucleotide sequence 23-66 of the DNA corresponding to the *Escherichia coli* tyrosine tRNA precursor has been systematically investigated. Joining was studied using all possible combinations of 3, 4, and 5 and larger numbers of segments at a time. The extent of joining varied widely (0 to about 90%) in three component systems. The "self-structure" of some of the components evidently inhibited the joining. Addition of a fourth segment in general enhanced the extent of joining and optimal yields were obtained in systems containing six or more segments. A comparison of the T_4-induced ligase and the *E. coli* polynucleotide ligase for joining of the chemically synthesized segments showed the *E. coli* enzyme to be inferior to the T_4-induced ligase. Satisfactory syntheses of the duplexes [IIa] and [IIb] comprising, respectively, eight and seven segments were achieved in single steps. Of the two terminal segments carrying 5'-OH groups in the duplexes, only one (segment 7) was used in the prephosphorylated form. The duplexes were isolated pure and characterized by enzymatic degradations and by electrophoresis.

The plan for the total synthesis of the DNA duplex corresponding to the entire nucleotide sequence of the tyrosine transfer RNA precursor has been formulated in an accompanying paper (2). Chemical syntheses of the required polydeoxyribonucleotide segments as well as the enzymatic joining of segments 1 to 5 to form duplex [I] (nucleotides 1 to 26) of the total duplex have been described in the accompanying papers (1, 3-7). The present paper describes the enzymatic joining of the segments 6 to 13 and of segments 6 to 12 to form the duplexes [IIa] and [IIb] (Fig. 1), representing the nucleotide sequence 23-66 and 23-60, respectively, of the precursor RNA. The following papers describe the enzymatic joining to form duplex [III] (nucleotides 57 to 94) (8), duplex [IV] (nucleotides 90 to 126) (9), and the subsequent joining of the duplexes representing the four parts to form the total duplex (10).

Of the segments which form the subject of the present work, segments 5 and 6 have the same hexanucleotide sequence, d(T-T-C-G-A-A), at their 5'-ends and both possess a thymidine unit at the 3' terminus. The two segments could easily substitute for each other if both were to be simultaneously present in the ligase reactions. Joining reactions were, therefore, arranged so that the first set, duplex [I], comprised segments 1 to 5 and duplex [II] began with segment 6. Starting with this segment, a systematic study was carried out of the maximum number of segments which could be advantageously and unambiguously used in the synthesis of the proposed duplex [II]. It was found that duplex [II] could nicely contain eight components (up to segment 13); as discussed later, segment 14 could not be included in this set. One-step synthesis and characterization of the duplex [IIa] (segments 6 to 13) was, therefore, accomplished and is now described. During subsequent work on the phosphorylation of the 5'-OH end groups in duplexes representing different parts of the structural gene, difficulties were encountered in driving the phosphorylation reactions to completion. Phosphorylation was sometimes incomplete in those cases where the 5'-OH

* This work has been supported by grants from the National Cancer Institute of the National Institutes of Health, United States Public Health Service (CA05178 and CA11981), the National Science Foundation, Washington, D.C. (GB-7434X, GB-21053X, GB-36881X, and BMS73-06757), the American Cancer Society (NP-140), and by funds made available to the Massachusetts Institute of Technology by the Sloan Foundation.

This is Paper CXXXVIII in the series "Studies on Polynucleotides." The preceding paper is Ref. 1.

‡ Present address, Department of Microbiology, University of Manitoba Winnipeg, Manitoba, Canada.

§ Present address, Department of Microbiology, University of British Columbia, Vancouver, B.C., Canada.

¶ Present address; Department of Virology, Hebrew University, The Hadassah Medical School, P.O. Box 1172, Jerusalem, Israel.

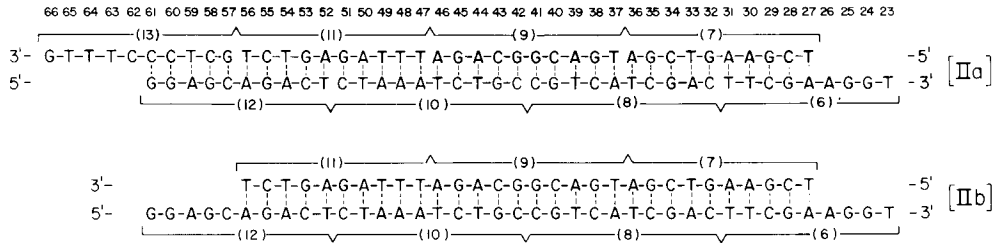

FIG. 1. Part [II] of the plan for the total synthesis of the gene for the tyrosine tRNA precursor. The deoxyribopolynucloetide segments synthesized chemically are shown between successive *carets*, the segment number being inserted into the *brackets*. The duplex [IIa] contained chemically synthesized segments 6 to 13 while duplex [IIb] contained segments 6 to 12.

end groups were present on the shorter strands of the duplexes. Phosphorylation of the 5'-OH groups carried by protruding single-stranded regions of the duplexes presented no difficulty. In view of these findings, the original plan for duplex [II] (see also 2) was amended and duplex [IIb] containing segments 6 to 12 was therefore synthesized. The preparation and characterization of this duplex is included in the present paper.

The yields in the joining of the short synthetic deoxyribopolynucleotides using the T_4-ligase have in general been less than quantitative and sometimes unsatisfactory. With the segments used in the present paper, extensive studies were carried out in attempts to understand the factors influencing the joining reactions. All possible combinations of 3, 4, 5, and larger numbers of segments at a time were tried in the joining reactions. The extent of joining varied widely in three-component reactions, the "self-structures" of some of the components evidently inhibiting the joining. The addition of a fourth component had, in general, a beneficial effect but optimal yields were obtained in systems containing 6 or more components. In further studies, the *Escherichia coli* ligase was compared with the T_4-induced ligase in parallel experiments. Although the *E. coli* enzyme had been tried at the outset (11) in joining experiments with synthetic deoxyribooligonucleotides, the study was not rigorous. However, the present work has confirmed the earlier conclusion that the T_4-ligase is superior to the *E. coli* enzyme in the joining of short deoxyribopolynucleotides.

MATERIALS AND METHODS

These were as described in the preceding paper (1).

RESULTS

Studies of Joining Reactions

Joining Reactions Using Three Deoxyribopolynucleotide Segments

The results obtained using all of the possible combinations of three segments (systems 1 to 7) are shown in Table I. Systems 1 and 2 failed to give any joining. System 3 gave 36%, system 4 gave 69%, while system 5 gave 92% joining. System 6 gave only a small amount of joining (23%) even after several additions of the enzyme. System 7 gave up to 52% joining but experiments to be discussed later indicated that this measurement probably included an unexpected joining. Thus, the extent of joining varied very widely in the three-component systems. As mentioned above, in one system (segments 10 to 12) the joining was practically quantitative. In contrast, strikingly, systems containing segments 6 or 7 gave no joining at all. Evidently, these segments, which contain partly self-complementary sequences d(T-T-C-G-A-A) in segment 6 and d(T-C-G-A) in segment 7; Fig. 2), form structures which do not permit the formation of the correct duplexes composed of the three segments.

TABLE I

Extent of ligase-catalyzed joining in systems containing different oligonucleotide segments

The reaction mixtures (usually 10 μl) were set up as described in the preceding paper. The concentration of the individual oligonucleotides was 2 μM and the ligase concentration was 400 units/ml. Kinetics were followed in all cases, the percentages shown being plateau values of phosphatase resistance.

System	Segments	Extent of Joining
		%
Containing three segments		
1	6, 7, 8	0
2	7, 8, 9	0
3	8, 9, 10	36
4	9, 10, 11	69
5	10, 11, 12	92
6	11, 12, 13	23
7	12, 13, 14	52
Containing four segments		
8	6, 7, 8, 9	34
9	7, 8, 9, 10	36
10	8, 9, 10, 11	39
11	9, 10, 11, 12	65
12	10, 11, 12, 13	77
13	11, 12, 13, 14	58
Containing five segments		
14	6, 7, 8, 9, 10	40
15	7, 8, 9, 10, 11	37
16	8, 9, 10, 11, 12	47
17	9, 10, 11, 12, 13	70
18	10, 11, 12, 13, 14 (2 × 14)	64 (72)
Containing more than five segments		
19	7, 8, 9, 10, 11, 12	56
20	6, 7, 8, 9, 10, 11, 12	48
21	6, 7, 8, 9, 10, 11, 12, 13	52
22	6, 7, 8, 9, 10, 11, 12, 13, 14	45

In an attempt to characterize one of the joined products further, the reaction mixture from the system 5 (segments 10 to 12) was applied to an Agarose 0.5m gel filtration column. The peak corresponding to the joined product contained only the decanonanucleotide formed from the joining of segments 10 and 12 and none of the expected duplexes of this decanonanucleotide with segment 11. The decanonanucleotide material is the deoxy equivalent of the tRNA anti-codon loop and it can evidently form a stable hairpin structure as shown in Fig. 3.

The formation of this stable structure accounts for the lack of isolation of the expected duplex composed of the three segments under conditions where it is usually possible to isolate such duplexes.

Joining Reactions Using Four Deoxyribopolynucleotide Segments

All of the possible four component systems containing segments 6 to 14 were studied and the results are shown in Table I (systems 8 to 13). In contrast to the wide variation observed in the three-component systems, substantial joining was observed in all of the four-component systems. Thus, system 8 (including segments 6 to 9), system 9 (segments 7 to 10), and system 10 (segments 8 to 11) all gave yields of 35 to 40%. System 11 (segments 9 to 12), system 12 (segments 10 to 13), and system 13 (segments 11 to 14) gave joining between 60 and 80%. As in the three-component systems, the four components containing segment(s) 6 and/or 7 gave lower yields. However, there was a marked effect of the addition of the fourth component (*e.g.* segment 9) such that the correct duplex could be formed to give yields of 35% or better.

Characterization of Joining Reactions in System 12 of Table I (Segments 10 to 13)—As described above, the three-component system containing segments 10 to 12 joined very well, while the system containing segments 11 to 13 joined only poorly. Therefore, it was of interest to study the rates of joining of the two strands in segments 10 to 13 (system 12 of Table I). In order to follow the kinetics of joining in the two strands, segment 10 was phosphorylated enzymatically using [γ-^{33}P]ATP, while segment 13 was phosphorylated using [γ-^{32}P]ATP. The results are shown in Fig. 4. Indeed, a difference in the rates of joining in the two strands was observed. The joining of ^{33}P-labeled segment 10 to segment 12 was relatively fast, being essentially complete within 4 hours. The joining of ^{32}P-labeled segment 13 to segment 11 had reached only 23% in 4 hours and required 24 hours to reach a plateau value. Also, there was a substantial difference in the extents of joining to form the two strands; segments 10 and 12 joined to the extent of 85%, while the joining of segments 13 to 11 was 62%.

For characterization, an aliquot of the reaction mixture was withdrawn after only 3 hours in order to examine the joined products at an intermediate time. This sample and another one from a completed reaction mixture were successively run through the same column of Agarose 0.5m and gave the elution patterns shown in Fig. 5. The pattern of the completed reaction (Fig. 5*A*) showed one major product (Peak I) which contained both ^{32}P and ^{33}P. This was characterized to be the expected joined product. There was a small amount of a second product containing ^{33}P only which was characterized as the product formed by the joining of segments 10 and 12. The unreacted segments were in Peak III and, as expected, there was a greater proportion of ^{32}P label in Peak III relative to ^{33}P.

In the partially completed reaction (Fig. 5*B*), three main peaks were observed. Peak I contained the four components and had both ^{32}P and ^{33}P. (The ratio of ^{33}P to ^{32}P in Peak I of Fig. 5*B* appears to be different from the ratio in Fig. 5*A* simply because of the different scales of radioactivity used.) Characterization of the peaks by enzymatic degradation is given in Table II. Peak II contained the decanonanucleotide (segments 10 + 12) and was present in much greater amount than in the completed reaction. This was in agreement with the greater extent of ^{33}P joining (70%) than of ^{32}P joining (20%) at 3 hours and with the observation described above that the decanonanucleotide separated away from the complementary segment 11.

Characterization of Duplex Consisting of Segments 11 to 14—The four-component system consisting of segments 11 to 14 (Fig. 6) was also studied further. Using all the segments 5′-phosphorylated (^{32}P), the kinetics of development of the phosphatase-resistant radioactivity are shown in the *inset* to Fig. 6. Gel filtration through an Agarose 0.5m column gave the elution pattern shown in the same figure. The shape of the peak corresponding to the product clearly suggested heterogeneity. Analyses of the fractions designated I and II (Fig. 6) gave the results shown in Table III. Thus, Fractions I and II were 63% and 50.5% phosphatase-resistant, respectively. From these results, Fraction II was concluded to be the expected product containing the four segments (Fig. 7*A*), while Fraction I was possibly larger in size. Nearest neighbor ana-

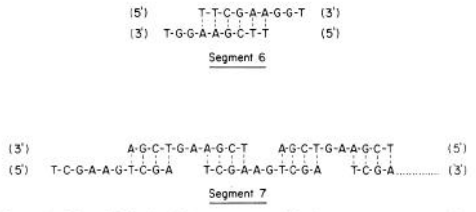

FIG. 2. Possible "self-structures" of segments 6 and 7.

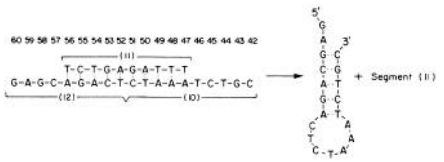

FIG. 3. Possible hairpin structure for the decanonanucleotide resulting from the joining of segment 10 to segment 12. The duplexed structure with segment 11 is shown at the *left*.

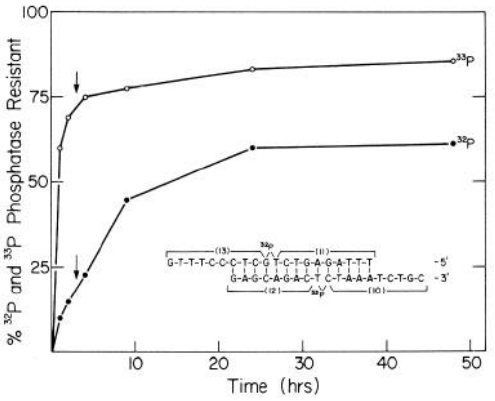

FIG. 4. Kinetics of the ligase-catalyzed joining reactions using segments 10 to 13. Segments 11 and 12 were unphosphorylated at 5′-end, segment 10 was phosphorylated at 5′-end with ^{33}P, and segment 13 with ^{32}P. The reaction was carried out at 5° in 50 mM Tris, pH 7.6 10 mM $MgCl_2$/10 mM dithiothreitol/33 μM ATP. The reaction mixture (400 μl) contained 5 μM concentration of each segment. It was boiled for 2 min and cooled to 5° slowly over a period of 5 hours before the addition of ligase to a concentration of 400 units/ml. A portion of 200 μl was removed from the reaction mixture after 3 hours and quenched by the addition of ethylenediaminetetraacetic acid to 20 mM and heating to 100°. For kinetics, samples were taken from the reaction mixture and assayed for phosphatase resistance by the DEAE-strip assay.

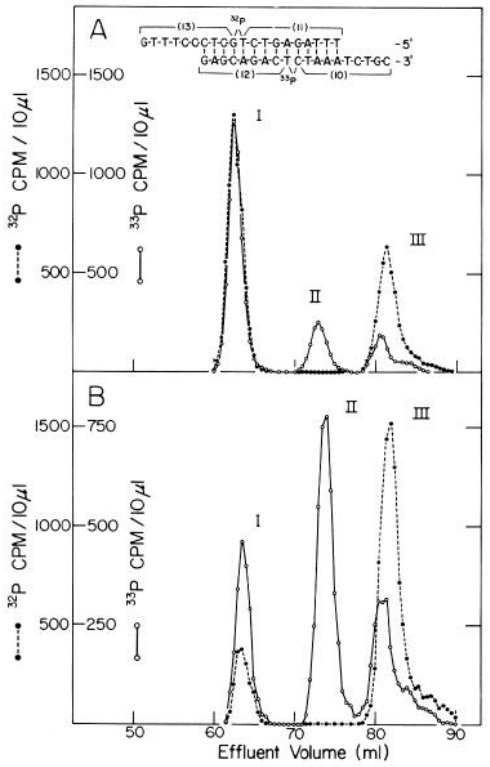

FIG. 5. Fractionation of 2 aliquots of the ligase reaction described in Fig. 4 containing [5′-^{33}P]segment 10, [5′-^{32}P]segment 13, and unphosphorylated segments 11 and 12. A column (1 × 150 cm) of Bio-Gel A-0.5 m (200- to 400-mesh) equilibrated with 50 mM triethylammonium bicarbonate was used. Fractions of 350 µl were collected. *A* shows the fractionation of the reaction mixture when joining was allowed to go for 48 hours. *B* is the fractionation of the aliquot removed after 3 hours of reaction. The characterization of *Peaks I* and *II* is described in Table II. *Peak III* contained the unreacted starting materials.

TABLE II

Characterization of products in enzymatic joining of segments 10, 11, 12, and 13 (Peaks I and II of Fig. 5)

Nearest neighbor (3′-nucleotide) and 5′-nucleotide analyses were performed as in the preceding paper.

3′-Nucleotide analysis (cpm)					
Peak	Label	dAp	dGp	dTp	dCp
I	^{32}P	0	5	1810	14
	^{33}P	3	16	1480	15
II	^{32}P	0	0	15	0
	^{33}P	3	6	1670	19

5′-Nucleotide analysis (cpm)					
Peak	Label	pdA	pdG	pdT	pdC
I	^{32}P	8	1700	42	0
	^{33}P	0	12	11	1030
II	^{32}P	0	7	3	0
	^{33}P	0	12	4	262

lysis revealed that Fraction I had a 2:1 ratio of radioactivity in dGp and dTp, while Fraction II had a 1.2:1 ratio of radioactivity in dGp and dTp. The expected four-component system would give a 1:1 ratio of radioactivity in dGp and dTp. Degradation to 5′-mononucleotides before a phosphatase treatment yielded a 3:1:1 ratio of counts in pdG, pdT, and pdC for Fraction I and a 2:1:1 ratio of counts in pdG, pdT, and pdC for Fraction II. After the phosphatase treatment, both fractions gave the bulk of the counts in pdG upon 5′-nucleotide analysis (Table III).

The sum of the above evidence indicated that Fraction I contained an additional segment with a labeled 5′-pdG unit joined to the 3′-dG end of segment 13. Of the two compo-

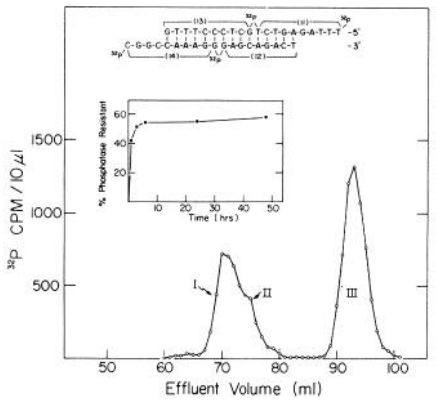

FIG. 6. The kinetics of joining and fractionation of the ligase reaction containing 5′-^{32}P-labeled segments 11, 12, 13, and 14. The kinetics for the formation of phosphatase-resistant radioactivity are shown in the *inset*. The reaction mixture (200 µl) contained 50 mM Tris, pH 7.6, 10 mM dithiothreitol, 10 mM $MgCl_2$, 33 µM ATP, and 2 µM concentration of each of the oligonucleotide segments. The segments were annealed by cooling from 95°–5° over a 5-hour period and the ligase was added to a concentration of 400 units/ml. The reaction mixture was passed through a column (1 × 100 cm) of Bio-Gel A-0.5 m (200- to 400-mesh). The eluant was 50 mM triethylammonium bicarbonate. The first peak containing Fractions I and II consisted of duplexes whose characterization is described in text, Table III, and in Fig. 7. *Peak III* contained unreacted starting material.

TABLE III

Characterization of products of joining of segments 11, 12, 13, and 14 (Peaks I and II of Fig. 6)

Phosphatase assay	
Peak	Phosphatase resistance
	%
I	62.6
II	50.5

3′-Nucleotide analysis after phosphatase (cpm)[a]				
Peak	dAp	dGp	dTp	dCp
I	10	331 (2.0)	161 (1.0)	6
II	15	193 (1.2)	164 (1.0)	11

5′-Nucleotide analysis after phosphatase (cpm)[a]				
Peak	pdA	pdG	pdT	dpC
I	76	1060	81	40
II	46	1360	50	38

5′-Nucleotide analysis before phosphatase (cpm)[a]				
Peak	pdA	pdG	pdT	pdC
I	17	817 (3.0)	260 (0.9)	273 (1.0)
II	22	502 (2.1)	239 (1.0)	223 (0.9)

[a] The numbers in parentheses are the experimental molar ratios.

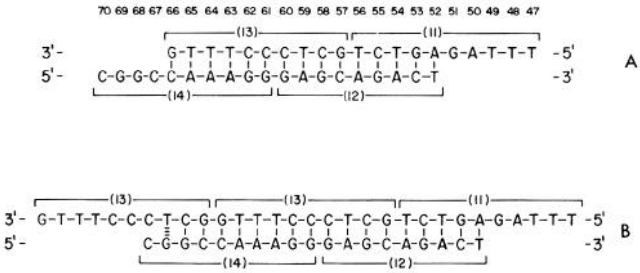

FIG. 7. *A* is the duplex expected to be formed from the joining of segments 11, 12, 13, and 14. *B* shows the unexpected duplex formed in this system (experiment of Fig. 6). Segment 13 can form two G-C base-pairs and one G-T base-pair with segment 14 allowing it to join to the 3'-end of segment 13 in the duplex (*A*) shown in the *top*.

nents, segments 12 and 13, which carried [^{32}P]pdG end groups, segment 13 can form a reasonably stable anti-parallel base-paired structure at its 5'-sequence with the protruding end of segment 14. Segment 12 can only form one base-pair in this region. It is, therefore, concluded that Fraction I contains 2 molecules of segment 13 joined end to end as shown in Fig. 7*B*. Additional support for this conclusion was obtained in an experiment in which only segments 13 and 14 were present. Up to 15% of joining was observed after a 48-hour reaction time and the 3'-nucleotide analysis of the joined product showed the bulk of the radioactivity to be in dGp. Indications of the self-joining of segment 13 were also obtained in systems 18 and 22 (Table I) which involved segments 11 to 14 and segments 6 to 14, respectively. As seen in Table I, an increase in the amount of segment 13 in the reaction mixture of system 18 resulted in an increase in the extent of phosphatase resistance. Further, a 3'-nucleotide analysis of the products formed in system 22 showed excessive (50 to 60%) radioactivity in dGp.

Systems Containing Five and Six Segments

Systems 14 and 15 (Table I) starting, respectively, with segments 6 and 7 gave relatively low yields (about 40%), while the systems 16 to 18 gave markedly higher yields (up to 70%). The six-component system, which also contained segment 7 (system 19) gave a joining of 56%. This represented a substantial increase over the above-mentioned five-component system which lacked segment 12.

The five-component system consisting of segments 6 to 10 was studied further for isolation and characterization of the resulting duplex. The reaction was carried out using segments 6, 8, 9, and 10, all of which carried 5'-^{32}P-phosphate group and segment 7 which was unphosphorylated at its 5'-end. The kinetics of joining, which reached a plateau at 35% phosphatase resistance, and the separation of the products are shown in Fig. 8. One main peak of the joined product was found, in addition to the peak corresponding to the unreacted oligonucleotides. The characterization by 5'- and 3'-nucleotide analyses shown in Table IV verified that the main product corresponded to the expected duplex comprising the five components. Thus, the 3'-nucleotide analysis gave a 2.1:1 ratio of radioactivity in dCp and dAp, while 5'-nucleotide analysis after the phosphatase treatment gave a 2.2:1 ratio of counts in pdT and pdC. Before the phosphatase treatment, a 1:1 ratio of radioactivity was found in dTp and dCp.

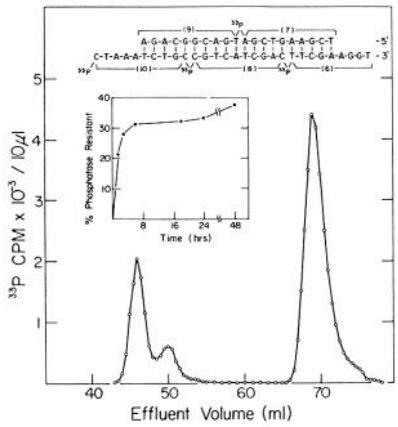

FIG. 8. Kinetics of joining and fractionation of the ligase reaction containing 5'-^{33}P-labeled segments 6, 8, 9, and 10 and unphosphorylated segment 7. The kinetics of joining are shown in the *inset*. The reaction mixture (200 µl) contained 50 mM Tris, pH 7.6, 10 mM dithiothreitol, 10 mM $MgCl_2$, 33 µM ATP, and 5 µM concentration of each of the oligonucleotide segments. The segments were annealed by cooling from 95°–5° for a 6-hour period and the ligase added to a concentration of 500 units/ml. The reaction mixture was passed through a column (1 × 90 cm) of Bio-Gel A-0.5 m (200- to 400-mesh). The eluant was 0.1 M triethylammonium bicarbonate. The first peak was the joined duplex and was characterized as in Table IV.

TABLE IV

Characterization of duplex from segments 6, 7, 8, 9, and 10 (Peak I of Fig. 8)

3'-Nucleotide analysis after phosphatase (cpm)[a]			
dAp	dGp	dTp	dCp
954 (1.0)	52	51	1980 (2.1)
5'-Nucleotide analysis after phosphatase (cpm)[a]			
pdA	pdG	pdT	pdC
163	154	3560 (2.2)	1620 (1.0)
5'-Nucleotide analysis before phosphatase (cpm)[a]			
pdA	pdG	pdT	pdC
236	334	9560 (1.0)	9530 (1.0)

[a] The numbers in parentheses after the counts per min are the experimental molar ratios.

Systems Containing Seven and Eight Segments

The experiments described above showed that the use of six components improved the joining. The seven-component system (segments 6 to 12) was tried next and this also gave satisfactory joining (50 to 60%) (Table I). A preparative scale experiment using the seven components is described later. Similarly, the eight-component system (segments 6 to 13) gave a very satisfactory yield of the joined product (Table I), although as described below, extra additions of the terminal segments (segments 6 and 13) had to be made to drive the joining to form the eight-segment duplex.

Lack of Inhibition of Joining of 5'-Phosphorylated Segment by Excess of Corresponding Unphosphorylated Segment

Of the data accumulated on the joining reactions, some data can be interpreted to mean that the joining reactions do not

go to completion because portions of the segments become frozen into wrong duplexed structures. If the formation and reassortment of the duplexes is a rate-determining step, then it would be interesting to examine the effect of the addition of an excess of a segment containing an unphosphorylated 5′-OH end to a reaction mixture containing and requiring the same segment in the 5′-phosphorylated form. As seen in Fig. 9, in the system containing segments 8 to 11, the addition of increasing amounts of unphosphorylated segment 8 had no effect on the rate or extent of the joining. Nearest neighbor analysis of the isolated joined products showed further that the presence of unphosphorylated segment 8 in the joining reaction had only a small effect on the ratio of joining in the two strands. Thus, a 5-fold excess of unphosphorylated segment 8 relative to the phosphorylated analog decreased the joining of the latter only by 15% (results not shown).

Comparison of Escherichia coli and T_4-Induced Ligases for Joining of Short Deoxyribopolynucleotides

Frequently, the joining of short deoxyribopolynucleotides using the T_4-induced ligase is not satisfactory (*e.g.* Table I). A comparison of this enzyme with that from *E. coli* (12)[1] was therefore made. In an initial experiment, activity of the two preparations was first compared for the end to end joining of [5′-^{32}P](dT)$_{10}$ in the presence of poly(dA) as the template (Fig. 10*A*). Since the *E. coli* enzyme was used at a 5-fold higher concentration than the T_4-ligase, the rate in this reaction was higher with the *E. coli* enzyme. However, both enzymes effectively brought about (dT)$_{10}$ polymerization.

Experiments with short synthetic deoxyribopolynucleotides (Fig. 10, *B* to *D*), however, showed that the *E. coli* enzyme was inferior and that the rate and extent of joining was uniformly higher with the T_4-ligase.

Synthesis of [II]

Duplex [II] Containing Segments 6 to 13 [IIa]

Three separate preparations at 8 nmol each were carried out. Segments 6 to 11 and 13 carried 5′-^{32}P-phosphate groups, segment 12 being unphosphorylated. While the small scale reactions are usually carried out at 1 to 5 μM concentrations of the individual segments, in the present large scale experiments, a 40 μM concentration of the oligonucleotides was used with no significant difference in the results (Fig. 11). In order to promote equivalent joining of segment 6 to the rest of the segments, a 50% excess of this segment was used since in a small scale experiment, contamination by the duplex containing segments 7 to 13 and lacking segment 6 was noted. As seen below, however, the presence of duplexes lacking terminal segments could not be eliminated at this stage.

The kinetics of joining in the eight-component system are shown in Fig. 11. A total of 51% phosphatase resistance was observed and this corresponded to 60% reaction because the 5′-^{32}P-phosphate group in segment 7 would not be affected. The products obtained on separation are shown in the same figure. The pattern showed the absence of duplexes of substantially smaller sizes. The main product peak showed 86.5% phosphatase resistance (theoretical for the eight-component product is 85.6%). Characterization by degradation to 3′- and 5′-nucleotides is given in Table V. Thus, the ratio of radioactivity in pdG to pdT to pdC was 1:4:2 before a phosphatase treatment, while after the phosphatase treatment the ratio was 1:3:2. Degradation to 3′-nucleotides gave a 1:1:1 ratio of radioactivity in dAp to dTp to dCp (Table V). However, electrophoresis of the product in Peak I on polyacrylamide gels under denaturing conditions indicated 10 to 15% contamination of one or more seven-component systems (Fig. 12). The shorter duplex can be seen in this figure as traveling ahead of the main eight-component product. The pattern obtained in Fig. 12 could imply that portions of the duplexes lacked segment 6 or segment 13. Incubation of the product in Fig. 11 with an excess of each of segments 6 and 13 in the presence of ligase eliminated most of the shorter duplexes as evidenced by the gel pattern shown in Fig. 12*b*.

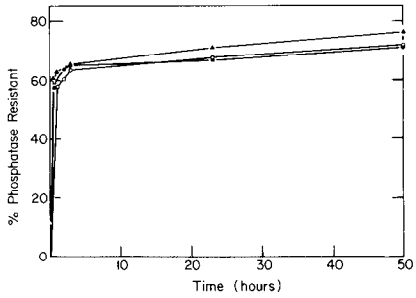

FIG. 9. Effect of addition of the same segment in the 5′-unphosphorylated form on the joining of the same segment in the 5′-phosphorylated form. The system used contained the four segments 8 to 11 with segments 8 and 11 containing 5′-^{32}P-phosphate group. The concentration of each of the segments was 2.5 μM, other conditions being standard. The kinetics of joining are shown: Δ——Δ, without unphosphorylated segment 8; ●——●, with 5 μM unphosphorylated segment 8; ○——○, with 12.5 μM unphosphorylated segment 8.

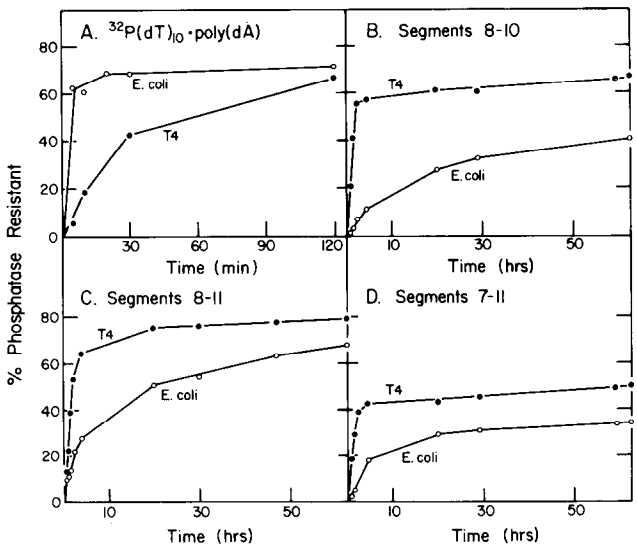

FIG. 10. Comparison of *Escherichia coli* and T_4-ligases in the joining of short deoxypolynucleotides. T_4-ligase reactions (10 μl) were performed as described under "Materials and Methods" with 2 μM segments concentration and 0.9 unit of T_4-ligase (●——●) as determined by the [5′-^{32}P](dT)$_{10}$ assay (13). Reaction mixtures with *E. coli* ligase in a volume of 10 μl contained: 20 mM Tris-HCl, pH 8.0; 7 mM $MgCl_2$; 20 μM NAD; 1 mM NH_4Cl; 1 mM EDTA; and *E. coli* ligase (4.5 units as determined by the [5′-^{32}P](dT)$_{10}$ assay (13)). *A*, [5′-^{32}P](dT)$_{10}$ (3 μM) and poly(dA) (10 μM); *B*, [5′-^{32}P]segment 8 and unphosphorylated segments 9 and 10; *C*, [5′-^{32}P]segments 8 and 11 unphosphorylated segments 9 and 10; *D*, [5′-^{32}P]segments 8, 9, and 11, and unphosphorylated segments 7 and 10.

[1] We are grateful to Professor I. R. Lehman of Stanford University Biochemistry Department for generously providing us with this enzyme.

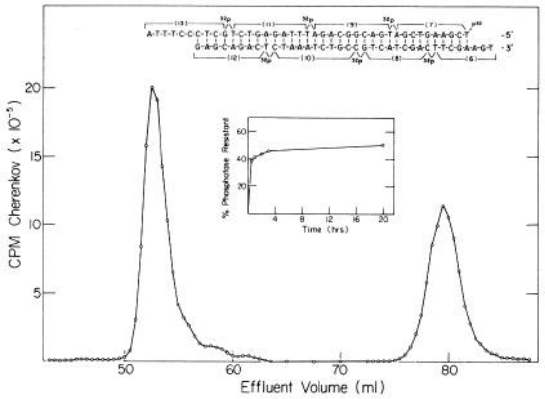

FIG. 11. Preparation of duplex [II] containing segments 6 to 13. Reactions containing 5′-^{32}P-labeled segments 6, 7, 8, 9, 10, 11, and 13 were used in 5′-phosphorylated (^{32}P) form while segment 12 was unphosphorylated. The kinetics of the formation of phosphatase-resistant radioactivity are shown in the *inset*. The reaction mixture contained 100 mM Tris, pH 7.6, 10 mM dithiothreitol, 10 mM $MgCl_2$, 500 μM ATP, and 40 μM concentration of each one of the oligonucleotides except for segment 6 which was present at 60 μM. The segments were annealed by slow-cooling from 95°–5° over a 4-hour period and the ligase was added to a concentration of 400 units/ml. After 20 hours, the reaction mixture was passed through a column (1 × 150 cm) of Bio-Gel A-0.5m (200- to 400-mesh). The eluant was 50 mM triethylammonium bicarbonate. The first peak contained the joined product and it was characterized as in Table V. The second peak contained the unreacted starting materials.

TABLE V

Characterization of duplex (Peak I of Fig. 11) from large scale joining of segments 6, 7, 8, 9, 10, 11, 12, and 13

3′-Nucleotide analysis after phosphatase (cpm)[a]			
dAp	dGp	dTp	dCp
10,900 (1.0)	796	10,950 (1.0)	10,260 (0.9)
5′-Nucleotide analysis after phosphatase (cpm)[a]			
pdA	pdG	pdT	pdC
77	5,820 (1.0)	18,470 (3.2)	11,810 (2.0)
5′-Nucleotide analysis before phosphatase (cpm)[a]			
pdA	pdG	pdT	pdC
74	4,450 (1.0)	15,050 (3.5)	9,370 (2.1)

[a] The numbers in parentheses after the counts per min are the experimental molar ratios.

Phosphorylation of Terminal 5′-OH Group in [IIa]

In order to further characterize [IIa], its single protruding 5′-OH group carried by the terminal guanosine nucleotide in segment 12 was treated with the kinase and [γ-^{32}P]ATP. Concentrations of the duplex below 2 μM were used as well for the denaturation of the duplex before phosphorylation. Variable and sometimes poor yields in the kinase reaction were obtained. The optimum conditions for the phosphorylation of the duplex (the first peak of Fig. 11) were found to be: 5 μM, duplex; 10 mM $MgCl_2$; 50 mM Tris, pH 7.6; and 200 units of the kinase/ml. The duplex was incubated at 37° for 15 min before the addition of enzyme and the reaction was allowed to proceed for 2 to 3 hours at 37°. After isolation of the duplex away from the excess of [γ-^{32}P]ATP, an aliquot was subjected to gel electrophoresis as shown in Fig. 12*d*. The slow moving band was undenatured duplex, while the major band corresponded to the bottom strand of size 38 nucleotides. A weak radioactive band also appeared at size 40. This probably arose by the exchange of the terminal 5′-phosphate of segment 7 with phosphate of [γ-^{32}P]ATP (14). In support of this conclusion, 5′-nucleotide analysis of the phosphorylated duplex gave a ratio of counts in pdG and pdT of 48,808 to 4,224. No other radioactive bands were visible, indicating that there were no other 5′-OH or 5′-phosphate end groups present in this preparation of [IIa].

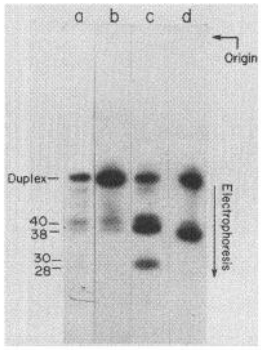

FIG. 12. Gel electrophoresis patterns of various preparations of duplex [IIa]. The samples were denatured in 100% formamide at 37° and subjected to electrophoresis on 12% polyacrylamide gels in 7 M urea. *Channel a* contains duplex [IIa] contaminated with seven-component systems. *Channel b* contains duplex [IIa] after further treatment with segments 6 and 13. *Channel c* contains duplex [IIa] and the contaminating seven-component systems after phosphorylation with the kinase. *Channel d* contains duplex [IIa] phosphorylated after a further reaction with segments 6 and 13.

Preparation of Duplex [IIb] (Segments 6 to 12)

A large scale (6.8 nmol scale) preparation of the duplex [IIb] containing 5′-^{32}P-phosphorylated segments (6 to 11) and unphosphorylated segment 12 is shown in Figs. 13 and 14. The conditions are shown in the legend, again about 50% excess of segment 12 being used. Since the joining of end segments seemed to be slower, analysis of the duplexes formed was performed by removing aliquots at different time intervals and subjecting them to polyacrylamide gel electrophoresis. The kinetics of the disappearance of different starting materials and of the intermediate products as well as of the formation of the required duplex [II] are shown in Figs. 13 and 14. The results plotted in the latter figure were obtained by cutting out appropriate sections of the gel and counting for Cerenkov radiation.

Large scale separation of the duplex [II] obtained in the above experiment was also carried out by preparative gel electrophoresis using 15% polyacrylamide gel. The major band (duplex [II]) and the bands corresponding to the duplex lacking segment 6 or segments 6 and 7 (Fig. 14) were eluted by powdering and stirring the appropriate sections with 2 M triethylammonium bicarbonate overnight. The supernatant solution and washes (3× with water) after centrifugation were concentrated and the concentrate passed through a column of Sephadex G-50 (0.9 × 24 cm) to remove salts. The product was eluted with 0.1 M triethylammonium bicarbonate.

The purified products (duplex [IIb]) was characterized by enzymatic degradations as shown in Table VI. The analyses for the distribution of radioactivity by the standard methods were all in general agreement with the expected values, except for the low value of the phosphatase-sensitive radio-

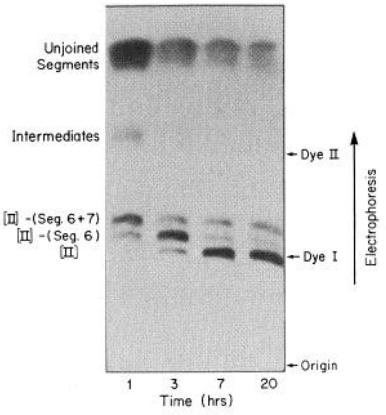

FIG. 13 (*left*). Kinetics of the joining of segments 6 to 12 (duplex [IIb]) as assayed by electrophoresis on a polyacrylamide gel. The reaction mixture (300 μl) contained [5′-^{32}P]segments 6 to 11 (23 to 25 μM each) and unphosphorylated segment 12 (33 μM) and 20 mM Tris, pH 7.6 and 10 mM $MgCl_2$. The reaction mixture was heated to 100° for 2 min, then cooled to room temperature in 15 min, then to 5°. The ligase (580 units/ml), dithiothreitol (10 mM), and ATP (200 μM) were added and the reaction mixture kept at the latter temperature. After 20 hours incubation, the mixture was again heated to 100° for 2 min, then cooled to 5°, and the same amounts of the ligase and dithiothreitol and one-fourth of the above amount of ATP were added. The mixture was incubated at 5° for another 22 hours. Aliquots (1 μl) were removed from the reaction mixture at different time intervals and subjected to the electrophoresis on a 15% polyacrylamide gel (20 × 20 cm). Radioactive bands, in order of increasing mobility were: duplex [II], duplex [II]-segment 6, duplex [II-(segments 6 + 7)], intermediates and unreacted segments. These were all cut out and counted by Cerenkov method.

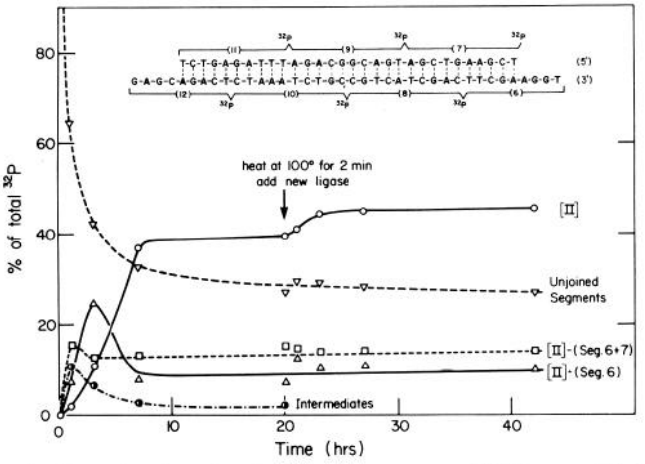

FIG. 14 (*right*). The results obtained from the experiment performed in Fig. 13 are plotted in this figure.

TABLE VI

Characterization of the duplex [IIb] (segments 6 to 12) (experiment of Figs. 13 and 14)

	Phosphatase resistance				
	Resistant (3890 cpm)			Sensitive (471 cpm)	
Molar ratio	5			0.6[a]	
	3′-Nucleotides				
	dAp	dGp	dTp	pdTp	dCp
Solvent I[b] (cpm)	2670	91	1270	4080 (pdTp + dCp)	
Molar ratio	2.0		1.0	3.0	
Solvent II[b] (cpm)	2340	242	1570	1560	2867
Molar ratio	1.5		1.0	1.0	1.8
	5′-Nucleotides				
	pdA	pdG	pdT	pdC	
Solvent I (cpm)	22	85	4100	2140	
Molar ratio			3.8	2.0	

[a] The 5′-phosphate group on segment 7 ought to be sensitive to the phosphatase. It is likely that the lower value for the phosphatase-sensitive ^{32}P is caused by the activation of this phosphate group by the ligase to form AMP-pyrophosphate. This has been demonstrated to occur in our present and previous work.

activity. (This has been commented on in the table.)

Similar analyses were performed on the accompanying products (Fig. 13). These showed that these duplexes lacked, respectively, segment 6 or both segments 6 and 7.

DISCUSSION

The synthesis of the DNA duplex (44 nucleotides long) reported in this paper has involved the largest number of chemically synthesized polynucleotide segments used so far in a one-step joining reaction. Studies were performed of the joining of the various segments comprising the present eight-component system: the segments were used in all possible combinations and the number of segments used in each joining reaction was increased systematically starting with the three-component systems. The results showed, especially in the case of the three- and four-component reactions, dramatic influences of the structures of the individual segments on the extent of joining. Thus, the yields in the joining of different three-component systems varied between 0 and better than 90% (Table I). Adding a fourth component obviously aided alignment to form the required double-stranded complex and led to considerable joining where the three-component systems failed to show any joining. Thus, it was particularly interesting that while neither segments 6 to 8 (system 1, Table I) nor segments 7 to 9 (system 2, Table I) gave any joining, the four components (segments 6 to 9) together, which comprise the above two systems gave about 35% joining. Presumably, the strong self-structures of one or more of these components (Fig. 2) could be overcome in part by the formation of an adequately long ordered double-stranded complex.

In a similar fashion, a six-component system (system 19) which included segment 7 gave better joining than a five-component system (system 15) which also included segment 7 (56% *versus* 37%). Further, all of the larger systems containing six, seven, or eight components gave similar extents of joining (about 50 to 60%). This systematic study led to the result that all of the eight components comprising this part of the DNA could be joined in one step to form a DNA duplex of 44 nucleotide units. However, as has been pointed out in the beginning, difficulty was experienced in quantitatively phosphorylating the 5′-OH hydroxyl group in the terminal segment 12 and, therefore, the plan for the total synthesis of the structural gene was altered.[2] The synthesis of the duplex containing segments 6 to 12 (duplex [IIb]) was also carried out.

[2] The synthetic DNA containing segments 1 to 13 has been used in preference to the DNA containing segments 1 to 12 for a study of a ribooligonucleotide primer-dependent initiation of transcription (unpublished work of Dr. R. Contreras of this laboratory).

The large scale syntheses of duplexes [IIa] (Fig. 11) and of duplex [IIb], which was monitored kinetically by gel electrophoresis, showed that the joinings of the terminal segments was slower than those of the internal segments. This could be largely corrected by using an excess of the terminal segments and by using longer reaction periods. Nevertheless, contamination of the desired duplex by products lacking one or more segments was found, especially, in the case of duplex [IIa] where separation was on an Agarose gel column. Retreatment of duplex [IIa] isolated in Fig. 11 with an excess of segment 6 and segment 13 evidently added the lacking components. Similarly, when some of the above preparation of duplex [IIa] was joined to duplex [I] (10), the deficiency in the joined product (duplex [I + II]) was corrected by treating with an excess of segment 13. The other incomplete duplex lacking segment 6 would not join to duplex [I] and was left behind in the joining reaction. Thus, the above situations do not cause any serious complication in the actual DNA synthesis. The fact that precise and complete base-paired overlaps are required for joining of the various sections selects in favor of only the correct duplex joining to the adjacent one. Indeed, this is a fortunate aspect of the present method of DNA synthesis that if the intermediate-sized DNA duplexes are deficient in one or more segments, the resulting product(s) are unlikely to accumulate or be carried through to the final steps in synthesis.

Will it be feasible to carry out in one step the joining of a much larger number of short chemically synthesized segments such as those that will form a DNA duplex corresponding to the gene for a transfer RNA? The comment at this stage would be to emphasize the possibilities of "wrong" joinings which lead to imperfect Watson-Crick structures. Unless a very elaborate computer program is used to guide the segmentation of the required long duplex such that homologies between different segments are totally reduced to ineffective levels, mistakes and ambiguities are likely to be introduced. Thus, in the present work, as described above, the number of components in duplex [IIa] could not be increased beyond eight. Neither segments 5 and 6 nor segments 13 and 14 could be used in joining reactions in the presence of each other, although for different reasons. These situations may not always be avoidable or foreseeable in the joining reactions. However, it will always be desirable to perform as many joinings as possible in one step, provided the situation remains completely unambiguous or can be demonstrated to be so. In general, the overall strategy for DNA synthesis will thus continue to be the joining of a few to several chemically synthesized segments to form an intermediate-sized duplex, which can be fully characterized, and subsequent joining of the duplexes to form a large or total duplex.

Why do the joining reactions not go to completion? As experienced in the present series of papers and in the previous work (14), a wide spectrum of yields is obtained at plateau values in different joining reactions. Clearly, there is an influence of the structures or self-structures of the components involved and each system has its characteristic plateau value of joining. It might have been expected that the ordered covalently linked double-helical structure would win out at the expense of the alternative partially duplexed structures to give quantitative yields of the joined products. Instead, the joining systems seem to become "frozen" at intermediate values.

REFERENCES

1. Sekiya, T., Besmer, P., Takeya, T., and Khorana, H. G. (1976) *J. Biol. Chem.* **251,** 634–641
2. Khorana, H. G., Agarwal, K. L., Besmer, P., Büchi, H., Caruthers, M. H., Cashion, P. J., Fridkin, M., Jay, E., Kleppe, K., Kleppe, R., Kumar, A., Loewen, P. C., Miller, R. C., Minamoto, K., Panet, A., RajBhandary, U. L., Ramamoorthy, B., Sekiya, T., Takeya, T., and van de Sande, J. H. (1976) *J. Biol. Chem.* **251,** 565–570
3. van de Sande, J. H., Caruthers, M. H., Kumar, A., and Khorana, H. G. (1976) *J. Biol. Chem.* **251,** 571–586
4. Minamoto, K., Caruthers, M. H., Ramamoorthy, B., van de Sande, J. H., Sidorova, N., and Khorana, H. G. (1976) *J. Biol. Chem.* **251,** 587–598
5. Agarwal, K. L., Caruthers, M. H., Fridkin, M., Kumar, A., van de Sande, J. H., and Khorana, H. G. (1976) *J. Biol. Chem.* **251,** 599–608
6. Jay, E., Cashion, P. J., Fridkin, M., Ramamoorthy, B., Agarwal, K. L., Caruthers, M. H., and Khorana, H. G. (1976) *J. Biol. Chem.* **251,** 609–623
7. Agarwal, K. L., Caruthers, M. H., Büchi, H., van de Sande, J. H., and Khorana, H. G. (1976) *J. Biol. Chem.* **251,** 624–633
8. Panet, A., Kleppe, R., Kleppe, K., and Khorana, H. G. (1976) *J. Biol. Chem.* **251,** 651–657
9. Caruthers, M. H., Kleppe, R., Kleppe, K., and Khorana, H. G. (1976) *J. Biol. Chem.* **251,** 658–666
10. Kleppe, R., Sekiya, T., Loewen, P. C., Kleppe, K., Agarwal, K. L., Büchi, H., Besmer, P., Caruthers, M. H., Cashion, P. J., Fridkin, M., Jay, E., Kumar, A., Miller, R. C., Minamoto, K., Panet, A., RajBhandary, U. L., Ramamoorthy, B., Sidorova, N., Takeya, T., van de Sande, J. H., and Khorana, H. G. (1976) *J. Biol. Chem.* **251,** 667–675
11. Gupta, N. K., Ohtsuka, E., Sgaramella, V., Büchi, H., Kumar, A., Weber, H., and Khorana, H. G. (1968) *Proc. Natl. Acad. Sci. U.S.A.* **60,** 1338
12. Modrich, P., Anraku, Y., and Lehman, I. R. (1973) *J. Biol. Chem.* **248,** 7495
13. Gupta, N. K., Ohtsuka, E., Weber, H., Chang, S. H., and Khorana, H. G. (1968) *Proc. Natl. Acad. Sci. U.S.A.* **60,** 285
14. Van de Sande, J. H., Kleppe, K., and Khorana, H. G. (1973) *Biochemistry* **12,** 5050