Construction of a double-stranded deoxyribonucleotide sequence of 45 base pairs designed to code for S-peptide$_{2-14}$ of bovine ribonuclease A

C.L. Harvey, K. Olson, A. de Czekala and A. L. Nussbaum

Chemical Research Department, Hoffmann-La Roche Inc., Nutley, NJ 07110, USA

Received 30 July 1975

ABSTRACT

An artificial DNA duplex, each strand consisting of 45 monomers, is constructed from chemically synthesized deoxyribonucleotides. The resulting bihelical polymer may code for a modified S-peptide of Ribonuclease A. This is the first synthetic duplex designed to code for a eukaryotic message.

INTRODUCTION

The chemical and enzymatic synthesis of oligodeoxyribonucleotides of defined sequence and potential biological significance proceeds apace. Thus, the pioneering work of Khorana's group on the synthesis of the structural gene for yeast alanine transfer RNA (2) has been followed by the construction of the precursor of Escherichia coli tyrosine suppressor transfer RNA (3). Synthesis of lactose operator oligonucleotide fragments have also been reported (4,5). These DNAs may be useful in various biochemical studies, for instance in sequencing DNA by serving as primers for DNA polymerases (6) and as tools for gene isolation (7). More recently, the minimal structural requirements of the EcoRI restriction and modification system have been determined by means of a self-complementary octanucleotide (8).

The question of whether it is possible to use chemically synthesized DNA to program the in vitro or in vivo synthesis of a specific protein is intriguing. One possible scheme involves the construction of bihelical DNA of defined sequence,
its insertion into an appropriate locus of a naturally occurring genome, and a study of its possible transcription and translation. In this paper, we report the synthesis of two complementary 45-mers (Fig. 1), a duplex designed to code for a peptide related to the S-peptide of Ribonuclease A (S-peptide<sub>2-14</sub>). This peptide can replace S-peptide proper in regenerating enzymatic activity to S-protein after subtilisin cleavage (9).

The strategy followed closely the approaches outlined by Khorana (10). Oligomers (11) (Fig. 1) 10-15 nucleotides in length, synthesized by the phosphorothioate procedure (12) and possessing overlapping complementarity as shown in Fig. 1, were joined enzymatically with phage T<sub>4</sub> polynucleotide ligase (13) and separated from unreacted fragments and the "Splint".

JOINING OF OLIGOMERS BY T<sub>4</sub> POLYNUCLEOTIDE LIGASE.

General Procedures. The 5'-termini of Fragments (Fig. 1) were labeled with <sup>32</sup>P in the polynucleotide kinase reaction after dephosphorylation with calf intestinal alkaline phosphatase (14). Isolation was carried out on Sephadex G-75 or G-100.

![Diagram of DNA segment coding for S-peptide<sub>2-14</sub>](https://example.com/f1.png)

Fig. 1. Structure of double-stranded DNA segment coding for S-peptide<sub>2-14</sub> (9). The sites of enzymatic joining are shown by arrows. Also shown is the hypothetical messenger RNA and the amino acid sequence directed by this DNA sequence.

2008
both superfine, in jacketed columns at 70°. The kinetics of joining were followed with the Norit assay (14). The joined Fragments were separated on the gel column with 0.1 M triethylammonium bicarbonate pH 7.6 (pH 8.3 after heating). The gel was heated in the above buffer to 100° and poured into the column at this temperature.

Joining Fragment G to 2. The separation of the 19-mer product from the Splint and the unjoined Fragments is shown in Fig. 2. The first peak contained the desired 19-mer. The second peak (not radioactive) contained the Splint, Fragment 1. The third peak was unused Fragment 2 (not radioactive) and peak 4 (radioactive) contained Fragment G, not joined. The separation of all unused components in this reaction mixture was facilitated by dephosphorylating Fragment 2 before the joining reaction.

**Fig. 2.** Purification of Segment (G + 2). The reaction mixture (0.9 ml) contained 66 mM Tris-HCl buffer (pH 7.6), 6.6 mM MgCl₂, 6.6 mM dithiothreitol, 111 μM ATP, and 55.5 μM each of the Fragments 1 (5'-ethylthioate blocked) (15), 2 (5'-phosphate removed) (16), and G (5'-32P of S.A. 1 x 10⁴ cpm/nmol). The mixture was heated at 70° for 5 min. and cooled, over a period of 1 hr., to room temperature. 50 Units of ligase were added, and the reaction was incubated overnight in ice. At this time, 60% of the Fragments had joined. The oligonucleotides were separated on a Sephadex G-75 superfine column (0.9 x 147 cm) at 70°. (———) A₂₆₀' (- - - -) CPM Cerenkov radiation.
It can be seen that the Splint overlaps Fragment G through seven bases, while Fragment 2 has only six bases complementary to the Splint. In order to obtain maximum extent (60-70%) as well as maximum rate of joining, the reaction had to be kept at 0\(^{\circ}\). The presence of a 6-base self-complementary sequence in Fragment 1 did not appear to have any deleterious effect on the reaction.

**Synthesis of Segment (3 + 1).** The formation of a 25-mer by linkage of Fragment 3 to Fragment 1 was carried out under the same conditions as those used for Segment (G + 2). In both reactions, the stability of the resulting duplex seemed to be similar. Fig. 3 shows the separation by gel chromatography of this reaction mixture. The first peak contained the joined Segment (3 + 1); it is followed by a peak consisting of the Splint and unjoined Fragments.

**Synthesis of Segment (5 + 4E).** This reaction was found to proceed very slowly and low yields were obtained when the reaction was carried out with the conditions used for synthesis of Segments (G + 2) or (3 + 1). Examination of the various reaction parameters revealed that, by raising both the temperature and magnesium concentration, it was possible to obtain a 35% yield. Separation of the reaction components (Fig. 4) yielded two peaks of radioactivity. The second peak, which contained the unjoined Fragments and the Splint, was evaporated to dryness and rejoined, using the same conditions as before. A yield of 37% was obtained for the second reaction, thus raising the total yield to an acceptable level.

**Joining Fragment T to 6.** The joining of Fragment T to Fragment 6E did not proceed well (10%). A study of this reaction showed that most of the phosphatase-resistant product was actually an adenylated DNA intermediate of the ligase reaction (20, Fig. 5) and not the desired Fragment. Intermediates of this type were not observed in any other joining carried out in the course of this work although their presence was always
Fig. 3. Gel-filtration of Segment (3 + 1). Reaction mixture and joining procedure were the same as in Fig. 2. Equal concentrations (55.5 μM) of Fragment 2 (5'-ethylthioate blocked), Fragment 3 (14) (5'-32P of S.A. 1 x 10^4 cpm/nmol), and Fragment 1 (5'-phosphate) were used. Separation was on a Sephadex G-75 superfine column (0.9 x 93 cm) at 70°. The final yield was 65%. (---) A_{260}' (-----) CPM Cerenkov.

Fig. 4. Separation of Segment (5 + 4) by gel filtration. A reaction mixture was made up as in Fig. 2, except that the MgCl2 concentration was 13.2 mM. The mixture contained 55.5 μM each of Fragment 6£ (5'-ethylthioate blocked) (17), Fragment 4£ (5'-phosphate) (18), and Fragment 5 (5'-32P of S.A. 1 x 10^4 cpm/nmol) (19), and 50 units of ligase were added; after 24 hr. at 10°, 35% of joining had occurred. The reaction was incubated another 24 hr. at 10° but no further joining was noted. Separation was on Sephadex G-75 superfine (0.9 x 93 cm) at 70°. (-----) A_{260}' (-----) CPM Cerenkov.
Fig. 5. Identification of ligase-mediated DNA-intermediate. An aliquot of 5 μl (3,500 CPM) from a 24 hr. synthesis of Segment (T + 6f) was incubated with 5 μg bacterial alkaline phosphatase in 100 μl of 0.1 M Tris-HCl buffer (pH 8.0) for 30 min. at 60°. The reaction was added to a 3 x 56 cm strip of DEAE paper (Whatman DE-81) with 0.1 μmol of Fragment T as a marker. The strip was developed in the descending manner with 0.5 M triethyl ammonium bicarbonate (TEAB) buffer (pH 7.6) for 15 hr. The strip was cut into 2 cm sections and counted in a toluene-based cocktail. The first phosphatase-resistant peak was Segment (T + 6f). The second peak (radioactivity not sensitive to phosphatase) at the position of Fragment T was eluted with 1 M TEAB, heated to 100° for 15 min. in 1 N-HCl. The sample was neutralized and the 32p-label was now found to be phosphatase sensitive (20).

The amount of unjoined adenylated DNA intermediate could be lowered by decreasing the ATP concentration (21); however, it was also necessary to use a large excess of T₄ ligase to attain an acceptable yield. Thus, the reaction was carried out with 10 μM of both Fragments T and 6f, and 20 μM of 5'-EtSp-Fragment 4E (15) as Splint and in an ATP concentration of 20 μM. The 5 ml reaction contained the same concentrations of other ingredients as those indicated in Fig. 2. After incubation with 500 units of T₄ ligase for two days at 10°, over 60% of the proper strands had joined. Separation (not shown)
was carried out on Sephadex G-75 superfine, as described for Fig. 3.

Completion of the plus (or nonsense) strand. The joining of Segment (5 + 4f) to (G + 2) in the presence of Fragment 3 as Splint was extremely slow. After 24 hrs., only 14% had joined. The reaction was heated to 70° and reannealed and 25 units of ligase added again. This was repeated on each day until the third day. By this time, 30% joining had taken place, no additional joining could be obtained. The separation of oligonucleotides is shown in Fig. 6. The first peak was the completed 45-mer. The second peak was excess Segment (5 + 4f). Peak 3 contained unjoined (G + 2) and peak 4 the Splint, Fragment 3.

Fig. 6. Purification of the plus strand. The reaction (0.6 ml) contained: 66 mM Tris HCl buffer (pH 7.6), 6.6 mM MgCl₂, 6.6 mM dithiothreitol, 33.3 μM ATP, and 16.7 μM each of Segment (G + 2) (5'-32P of S.A. 5 x 10⁴ cpm/mmol), Segment (5 + 4f) (5'-phosphate), and Fragment 3 (5'-ethylthioate blocked). The reaction was heated and annealed as in Fig. 2. 50 Units of T₄ ligase were added and the mixture incubated at 0°. At 24, 48, and 72 hr. the reaction was heated, reannealed, and 50 units of ligase added each time. By 72 hr., the reaction had stopped at 30% joined and the components were separated on a Sephadex G-100 column (0.9 x 70 cm). (-----) A₂₆₀' (--- - - -) CPM Cerenkov.
Synthesis of the minus (or sense) strand. An attempt was made to join Segment (T + 6S) to (3 + 1) in the presence of Fragment 5. Although numerous trials were made under a variety of conditions, this joining always failed. In order to further stabilize the minus strand for joining, a Splint consisting of the entire Segment (S + 4S) was used in place of Fragment 5. Incubation of this reaction at 10°C for two days converted 60% of the Segments (T + 6S) and (3 + 1) to the completed 45-mer minus strand. The separation of oligonucleotides is shown in Fig. 7.

Fig. 7. Separation of completed minus strand. A reaction mixture with 1.2 ml was made up as in Fig. 6. The following Segments were added: 8.3 μM (T + 6S) (5'-32P of S.A. 5 x 10^4 cpm/nmol), 8.3 μM (1 + 3) (5'-phosphate) and 7.5 μM (4S + 5) (5'-phosphate removed). The mixture was heated and annealed as in Fig. 2. Ligase (100 units) was added and the reaction was incubated at 10°C for 66 hr. At this time it was found that 65% of the strands had joined. Another 25 units of ligase was added and the reaction incubated another 24 hr.: no further joining had taken place. The strand was separated on a Sephadex G-100 gel column (0.9 x 93 cm). (- - - -) A260, (---) CPM Cerenkov.

2014
CHARACTERIZATION

The structure of each 45-mer strand arises almost by necessity from the established structures of its components: thus, in each joining reaction, the three oligomer components used were themselves homogeneous and of the proper sequence. In order to assure that each of the proper Fragments or Segments participated in the joining event and that the newly formed phosphodiester bond connected the proper termini, the following criteria had to be met: 1. The 5'-32p label of the fragment to be joined was changed during ligase joining from a phosphatase sensitive linkage (phosphomonoester) to a linkage resistant to this enzyme. 2. The size of the newly constructed polymer was larger than either component, as apparent from the gel filtration elution pattern (Fig. 8). 3. After separation, the labeled product gave products of enzymatic hydrolysis predicted from the known mode of action of specific nucleases (Table I). Thus venom phosphodiesterase

<table>
<thead>
<tr>
<th>Product</th>
<th>% Distribution of [$^{32}$P]</th>
<th>Expected</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>d-pT d-pA d-pC d-pG</td>
<td></td>
</tr>
<tr>
<td>G + 2</td>
<td>0 0 0 $&gt;$95</td>
<td>d-pG</td>
</tr>
<tr>
<td>3 + 1</td>
<td>$&gt;$95 0 0 0</td>
<td>d-pG</td>
</tr>
<tr>
<td>5 + 4</td>
<td>$&gt;$95 0 0 0</td>
<td>d-pG</td>
</tr>
<tr>
<td>T + 6</td>
<td>$&gt;$95 0 0 0</td>
<td>d-pG</td>
</tr>
<tr>
<td>plus strand</td>
<td>0 $&gt;$95 0 0</td>
<td>d-pG</td>
</tr>
<tr>
<td>minus strand</td>
<td>95 0 0 0</td>
<td>d-pG</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>d-Tp d-Ap d-Cp d-Gp</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>G + 2</td>
<td>$&gt;$95 0 0 0</td>
<td>d-Tp</td>
</tr>
<tr>
<td>3 + 1</td>
<td>0 0 $&gt;$95 0</td>
<td>d-Tp</td>
</tr>
<tr>
<td>5 + 4</td>
<td>0 0 $&gt;$95 0</td>
<td>d-Tp</td>
</tr>
<tr>
<td>T + 6</td>
<td>0 0 $&gt;$95 0</td>
<td>d-Tp</td>
</tr>
<tr>
<td>plus strand</td>
<td>0 $&gt;$95 0 0</td>
<td>d-Tp</td>
</tr>
<tr>
<td>minus strand</td>
<td>$&gt;$95 0 0 0</td>
<td>d-Tp</td>
</tr>
</tbody>
</table>

Table I. Characterization of joined strands by digestion with phosphodiesterases. Upper panel: venom (24); lower panel: spleen (14)
Fig. 8. Elution volume for Sephadex G-100 vs. log of number of nucleic acid residues. The two lines connect data points from columns of different length.

**FUTURE DIRECTIONS**

It is to be noted that the quantities of full-length strands available are extremely small. Several schemes have been proposed for the multiplication of such rare and valuable species (23,24): we are currently engaged in extending each 45-mer at its 5'-terminus with a self-complementary ("restriction") sequence; in the bihelical duplex these extensions would constitute single-stranded projections. It is intended then to integrate the resulting molecule into a genome bearing identical (hence, complementary) single-stranded termini --
the product of restriction endonuclease fission — in order to assure a supply of "minigene" by in vivo replication (25,26, 27).

There is, of course, no assurance that the particular sequence chosen is identical with that occurring naturally in the mammalian genome. The universality of the genetic code (28,29) precludes outright translation errors, but the actual use of any particular triplet in a specific system is, as yet, problematical. In this connection, it is gratifying to note that the triplets selected several years ago, then by necessity with a view only toward chemical efficiency (i.e., to minimize the number of different Fragments to be chemically synthesized) and to the avoidance of adverse complementary (i.e., self-annealing, or otherwise undesired annealing, causing interference to proper ligase sealing), have since been demonstrated to code for the designated amino acids, at least in prokaryotic systems (Table II). The problem of optimal strategy has recently been discussed (45).

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Codon Assigned</th>
<th>Ref.</th>
<th>Amino Acid</th>
<th>Codon Assigned</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Met</td>
<td>AUG</td>
<td>31,32,33, 34,35,36</td>
<td>Phe</td>
<td>UUU</td>
<td>34,35,38, 39,41</td>
</tr>
<tr>
<td>Glu</td>
<td>GAA</td>
<td>31,33,34, 37,38,39</td>
<td>Arg</td>
<td>AGA</td>
<td>39,40</td>
</tr>
<tr>
<td>Thr</td>
<td>ACU</td>
<td>33,34,38, 39,40</td>
<td>Gln</td>
<td>CAG</td>
<td>34,38</td>
</tr>
<tr>
<td>Ala</td>
<td>GCG</td>
<td>34,39,40</td>
<td>His</td>
<td>CAU</td>
<td>34,39,43, 44</td>
</tr>
<tr>
<td>Ala</td>
<td>GCU</td>
<td>33,34,41</td>
<td>Asp</td>
<td>GAU</td>
<td>34,41</td>
</tr>
<tr>
<td>Lys</td>
<td>AAA</td>
<td>34,38,42</td>
<td>ochre</td>
<td>UAA</td>
<td>35,36</td>
</tr>
</tbody>
</table>

Table II. In vivo occurrence of codons used in the present work as deduced from actual sequencing and frameshift mutations, primarily in E. coli and its phages. References are selective only.
As to the use of the polynucleotides herein described, at the very least they can be envisaged to serve as primer-templates in sequence studies of a defined site in mammalian DNA (6): occasional miscoding, even at the terminal base, should not prevent proper annealing of these relatively long molecules; their length should insure that such annealing occur uniquely at the desired site.

REFERENCES
1 Present Address: Department of Biological Chemistry, Harvard Medical School, Boston, Massachusetts 02115, USA
7 Doel, M. T. and Smith, M. (1973) FEBS Letters 34, 99-102
11 The chemically synthesized oligomers are called "Fragments", and numbered as shown in Fig. 1; two enzymatically joined Fragments are called a "Segment", its component Fragments being indicated in parentheses. Oligomers complementary to portions of Fragments or Segments to be joined are termed "Splints". The completed 45-mers are called strands. The strand complementary to the hypothetical messenger is called the minus (or sense) strand; the other strand is, correspondingly, the plus (or nonsense) strand. S.A. denotes specific activity.
24 Olson, K., Gabriel, T., Michalewsky, J. and Harvey, C. L. (1975) Nucleic Acids Res. 2, 43-60
33 Fiers, W., Contreras, R., de Wachter, R., Haegeman, G.,
Nucleic Acids Research


2020