DNA-like duplexes with repetitions. I. Properties of concatemer duplexes formed by d(T-G-C-A-C-A-T-G)

Z.A.Shabarova, N.G.Dolinnaya, S.I.Turkin and E.S.Gromova

Department of Chemistry and A.N.Belozersky Laboratory of Molecular Biology and Bioorganic Chemistry, Moscow State University, Moscow 117234, USSR

Received 16 April 1980

#### ABSTRACT

A new class of synthetic DNA duplexes containing repeating oligonucleotide sequences,double-helical concatemers, is characterized. The UV-absorption and circular dichroism of a concatemer formed in self-association of d(T-G-C-A-C-A-T-G) have been studied. The thermodynamical parameters of complex formation are the following:  $\triangle H_1^o = -9.210.3$  kcal/mol,  $\triangle S_2^o =$  $= -27\pm1$  e.u. The data obtained show that pseudopolymeric duplexes having structures that are similar to DNA-B-type helices are formed in solutions of d(T-G-C-A-C-A-T-G). Polymerization of 35P-d(pT-G-C-A-C-A-T-G) induced by water-soluble carbodiimide has been carried out under the conditions of concatemer stability. The yield of the dimer, a 16-member oligonucleotide, was 13%.

#### INTRODUCTION

The present communication describes a new class of synthetic DNA-like duplexes containing repeating oligonucleotide sequences (concatemer duplexes). The duplexes are built from one or several synthetic single-stranded oligodeoxyribonucleotides whose primary structures are such that the initial duplex has cohesive ends that make concatemers grow. The three simplest examples given below illustrate formation of the duplexes from one or two oligodeoxyribonucleotides (symbol d is omitted).

1) From one oligonucleotide, e.g. octanucleotide:  $T-G-C-A-C-A-T-G$  -  $T-G-C-A-C-A-T-G$ \* \* \* 0  $G-T-A-C-A-C-C-I$  $(5' - 3')$  - -  $\frac{1}{T}$ -G-C-A-C-A-T-G<sup>i</sup> T-G-C-A-C-A-T-G<sup>'</sup>  $(3'-5')$  - -  $A-C-G-T$ <sub>+</sub> $G-T-A-C-A-C-G-T$ <sub>+</sub> $G-T-A-C$ I

© IRL Press Umited, <sup>1</sup> Falconberg Court, London W1V 5FG, U.K.

2) From two oligonucleotides,

a) e.g. two decanucleotides:

 $T-G-C-A-T-T-T-T-T-T$  +  $T-G-C-A-A-A-A-A-A-A-A-A$  $T-G-C-A-T-T-T-T-T-T-T$  $A-A-A-A-A-A-A-AC-G-T$  $(5:-3")$  -- $\frac{1}{T}-G-C-A-T-T-T-T-T-T-T-G-C-A-T-T-T-T-T-T-T$  $\sqrt{(3'-5')}$  ---A-C-G-T<sub>1</sub>A-A-A-A-A-A-A-C-G-T<sub>1</sub>A-A-A-A IIa

```
b) e.g. deca- and dodecanucleotides:
```
 $T-G-C-A-T-T-T-T-T-T-T$  +  $G-A-A-T-T-C-A-A-A-A-A-A-A$ |T-G-C-A-T-T-T-T-T-T<sup>+</sup>G-A-A-T-T-C-A-A-A-A-A-A<sup>|</sup> 1A-A-A-A-A-A-C-T-T-A-A-G1T-T-T-T -T-T-A-C-G-T1 <sup>†</sup>G-A-A-T-T-C-A-A-A-A-A-A<sup>+</sup>T-G-C-A- $-C-G-T+A-A-A-A-A-A-C-T-T-A-A-G_T$  $r_{T-G-C-A-\frac{1}{X}}+r_{G-G}$  $(5'-3')$  --- A-C-G-T<sub>1</sub>X-A-C-G-T<sub>1</sub>X-A-C-**IIb** 

where X is a self-complementary sequence, T-T-T-T-T-T G-A-A-T--T-C-A-A-A-A-A-A, containing a single interruption between T and G.

Complex I is formed of an oligonucleotide comprising two self-complementary sequences of an arbitrary length (four, six, eight, etc. nucleotide units):

d(T-G-C-A-C-A-T-G), d(A-A-G-C-T-T-G-G-C-C) et al.

Formation of complex II requires the starting oligonucleotides to contain complementary sequences of equal length attached to self-complementary sequences of identical (IIa) or different (IIb) length and primary structure. The self-complementary sequences must be at the same  $(5'$ - or  $3'$ -)end of the chain. Complexes comprising three or more oligonucleotides may be buit likewise.

The resultant DNA-like duplexes with repetitions offer much promise for the study of nucleic acid-protein interactions, in particular for the study of the mechanism of the action of enzymes that recognize duplexes of a certain primary structure (restriction endonucleases, processing enzymes, etc.). They may also prove useful in preparation of substrates for affinity chromatography of these enzymes, for activity assays and for genetic engineering. Duplexes that are built from a limited number of synthetic oligonucleotides are excellent model systems for the study of various template-guided reactions including "chemical ligation" (template-guided chemical joining)<sup>1,2</sup>.

There are no reported data on the properties of concatemer oligonucleotide complexes, except for a short note on possible formation of triple helices containing overlapping sequences from oligo(A) and oligo(U)  $3\overline{)}$ . Bahl et al. <sup>4</sup> dealt with linkers containing cohesive ends and Dobrynin et al.  $^5$  with concatemer complexes in their works on the synthesis of oligodeoxyribonucleotide duplexes containing sites recognized by restriction endonucleases. Oligonucleotide complexes with broken sugar-phosphate chains (both concatemer-type and complexes formed by complementary oligonucleotides of different length) have been studied insufficiently. Such complexes were first prepared in Khorana's laboratory in the course of gene synthe sis <sup>6</sup>. Recently, the physico-chemical properties of this type of systems have been studied  $7$ . However, quite a number of points, e.g. self-assembly, enzymatic and chemical joining, etc., remain to be elucidated.

We began the study of the physico-chemical properties of the concatemers and the development of the "chemical ligation" method using complexes of type I which are comparatively easy of access. This work deals with concatemers made of synthetic oligonucleotide d(T-G-C-A-C-A-T-G):

 $(5:-3")$  --1-G-C-A-C-A-T-G<sup>i</sup>T-G-C-A-C-A-T-G<sup>i</sup>  $(3'-5')$  --A-C-G-T<sub>1</sub>G-T-A-C-A-C-G-T<sub>1</sub>G-T-A-C

The octanucleotide sequence rules out the possibility of formation of hairpin structures and duplexes lacking cohesive ends.

## MATERIALS AND METHODS

General. Octanucleotide d(T-G-C-A-C-A-T-G) was synthesized by the triester method as previously described  $8$ . Snake venom phosphodiesterase and bacterial alkaline phosphatase were purchased from Worthington Biochemical Corp.; CIEC 9 and MES, from Merck. The sample of  $32P-d(pT-G-C-A-C-A-T-G)$  was prepared by Dr. P.M. Rubtsov (Institute of Molecular Biology, Academy of Sciences of the USSR) by phosphorylation of d(T-G-C-A-C-A-T-G) with  $\gamma$  -<sup>32</sup>P-ATP and T4 polynucleotide kinase  $^6$ .

To prepare the concatemer complex, d(T-G-C-A-C-A-T-G) was purified by paper chromatography in n-propanol- conc.  $NH_3-H_2O$ (35 : 10 : 55  $v/v$ ), dissolved in 0.004 M phosphate buffer (pH  $7.3$ ), 0.2 M NaCl, 0.075 M MgCl<sub>2</sub> (nucleotide concentration per monomer,  $C_0$ , was 10<sup>-4</sup> to 10<sup>-2</sup> M), and stored at -5<sup>o</sup>C. The molar extinction coefficient,  $\mathcal{E}_{260}$  of d(T-G-C-A-C-A-T-G) determined after hydrolyzing octanucleotide with snake venom phosphodiesterase was 9350  $\text{M}^{-1}$ . cm<sup>-1</sup>.

Optical measurements. UV absorption spectra were recorded in a Cary-15 spectrophotometer. CD spectra were obtained in a Roussel-Jouan II dichrograph. To determine temperature-dependent changes in the absorbance and CD, thermostatted quartz cells of 10, 1 and 0.1 mm path lengths were used. Temperature was measured with a copper-constantan thermocouple. The cell compartments were blown through with dry nitrogen. Hypochromicity, h, due to complex formation was calculated by equation:

$$
h = \frac{A_{45} \cdot A_{-2} \cdot A_{-2}}{A_{45} \cdot A_{45}}
$$

where  $A_{450}$  and  $A_{220}$  are the absorbances of the octanucleotide solutions at  $45^{\circ}$  and -2°C (  $\lambda$  258 nm).

Chemical polymerization of  $32P-d(pT-G-C-A-C-A-T-G)$  as a concatemer constituent. A solution of terminally labelled with  $32<sub>P</sub>$  octanucleotide (10<sup>-3</sup>M per monomer) in 0.05 M MES (pH 6.0), 0.02 M MgCl<sub>2</sub>, was made 0.1 M in CMEC at room temperature and allowed to stand at 0°C. The condensation products were resolved by polyacrylamide gel electrophoresis as follows. 10 to 20)11 of a solution of marker dyes (0.025% xylene cyanol, 0.025% bromphenol blue, 80% formamide) was added to 2 to 3  $\mu$ l

of the reaction mixture. The samples were then heated on a water bath at 90° for 15 to 30 sec, loaded on 40 x 20 x 0,15 cm 20% acrylamide, 0.66% N,N'-methylenebisacrylamide, 7 K urea, 50 mM tris-borate (pH 8.3), 0.1 ml EDA slab gel and electrophoresed at a constant voltage of 20 V/cm. Prephoresis was carried out at constant voltage of 20 V/cm for two hr. After electrophoresis, the gel was exposed against an X-ray film. The products were eluted as recommended in Ref. 10.

### RESULTS AND DISCUSSION

#### 1. Physico-chemical properties

UV absorption. Complementary interactions of d(T-G-C-A-C-  $-A-F-G$ ) were studied in 0.004 M phosphate buffer (pH  $7.3$ ), 0.2M NaCl, 0.075 M MgCl<sub>2</sub>. UV absorbance vs. temperature curves for octanucleotide solutions of  $C_0 = 10^{-4}$ ,  $10^{-3}$  and  $10^{-2}$ M are depicted in Fig. 1. The shape of the melting curves, the considerable hypochromicity and concentration dependence of melting temperature  $(T_m)$  suggest formation of an intermolecular complex



Fig. 1. Melting profiles of d(T-G-C-A-C-A-T-G) at various concentration:  $10^{-4}$ M (I),  $10^{-3}$ M (2) and  $10^{-2}$ M (3). Insert: plot of  $1/T_m$  against logarithm of nucleotide concentration (per monomer).

in octanucleotide solutions at reduced temperatures. The magnitude of the hypochromicity due to the complex formation was equal to  $16.8$ , 19.1 and 20.4% for  $10^{-4}$ ,  $10^{-3}$ , and  $10^{-2}$ M solutions, respectively. Decrease of the oligonucleotide concentration to  $10^{-4}$ M thus results in reduction of the hypochromicity of the complex. This is likely to be due to incomplete complexation at  $C_{\rho}$  below 10<sup>-3</sup>M rather than to concatemer chain length variations with nucleotide concentration.

The optical data are insufficient to determine the length of the double helix formed. This was estimated from indirect evidence, i.e. from analysis of the products of enzymatic joining of octanucleotides (see below). The enzymatic reaction was found to give polymer products containing several hundred mononucleotide units. of the double helix formed. This was estimated from indirect<br>evidence, i.e. from analysis of the products of enzymatic<br>joining of octanucleotides (see below). The enzymatic reaction<br>was found to give polymer products conta

 $1/T_m$  vs. lnC<sub>o</sub> plot being a straight line (Fig. 1). Using the "all-or-none" model which describes the build-up of oligomer complexes to a reasonably good approximation  $11-13$ , the concatemer formation may be written in a simplified form as follows: K

$$
\mathbf{A} \quad \stackrel{\mathbf{A}}{\longrightarrow} \quad \mathbf{A}_{\mathbf{A}},
$$

where n is the number of molecules involved. The oligomer concentration (per monomer) will be denoted as  $C_0$ , and the equilibrium concentration of the free oligomer and complex comprising n oligonucleotide molecules  $C_A$  and  $C_{A_n}$ , respectively, so that

$$
C_o = mc_A + mc_{A_n},
$$

where m is the number of monomer units in the initial oligomer  $(m = 8$  in our case). Using notation mmC<sub>A<sub>n</sub></sub>/C<sub>o</sub> = f for the fraction of monomer units participating in the complex formation, the equation for the equilibrium constant may be written as

$$
K = \frac{C_{A_{n}}}{(C_{A})^{n}} = \frac{f \cdot m^{n-1}}{n(1-f)^{n} C_{0}^{n-1}}
$$
 (1)

With f of 1/2,  $=\frac{(2m)^{n}}{n\cdot C_0^{n-1}}$  equation  $-RTlnK = \Delta H^{\circ} - T \Delta S^{\circ}$ then gives

$$
(n - 1)(lnC_0 - ln2m) + ln n = \frac{\Delta H^o}{RT_m} - \frac{\Delta S^o}{R}
$$
 (2)

The  $1/T_m$  vs.  $lnC_o$  straight line slope thus gives  $\Delta H^o/n-1$ which is equal to  $-37.0^{\text{+}}1.2$  kcal/mol in our case. With n equal to unity, no complex is formed. The enthalpy of the reaction between two oligonucleotide molecules (n = 2) is  $-37.0^{+}1.2$ kcal/mol (four complementary base pairs are formed in the process). When n is 3, which corresponds to the formation of eight nucleotide pairs,  $\triangle H^o$  is -74.0<sup>+</sup>2.4 kcal/mol, etc. The entropy of complex formation is equal to -108<sup> $\pm$ </sup>4 e.u. (n = 2), -216<sup> $\pm$ </sup>8 e.u.  $(n = 3)$ , etc. Assuming that the thermodynamic parameters of the formation of duplex base pairs,  $\Delta H^o$ <sub>1</sub> and  $\Delta S^o$ <sub>1</sub>, are the same, the mean values of  $\triangle H^o$ <sub>1</sub> and  $\triangle S^o$ <sub>1</sub> are -9.2<sup>±</sup>0.3 kcal/mol and  $-27\pm1$  e.u., respectively. These values are close tothe base pair formation parameters in oligomeric complexes with the same G.C content though free of internal interruptions <sup>12</sup>. The enthalpy,  $\Delta H^o$ <sub>1</sub>, determined by direct calorimetric measurements of heat of denaturation of DNA containing 34% of G.C pairs, is equal to  $+9.65$  kcal/mol  $^{14}$ .

Complex formation following the "all-or-none" model is known to occur when the chain growth stability constant, S, is sufficiently high,  $S > 10$  (ref. 15). In our case, S decreases rapidly as the concatemer chain length increases, i.e. S is equal to 82 (21°C) at  $n=2$ , 20 (21°C) at  $n=4$ , etc. (S was calculated by formula:  $^{16}$  K =  $\beta$  S<sup>N</sup>, where N is the number of nucleotide pairs in the duplex,  $\beta$  is the nucleation parameter set equal to  $10^{-4}$  $M^{-1}$ , ref. 17). It thus follows that the "all-ornone" model can hardly be applied to describe the formation of sufficiently long concatemers. The process may be treated in terms of oligonucleotide interactions with a polymer template. To calculate the enthalpy of concatemer formation, the "all-or-none" model developed to handle oligomer-polymer system may be applied. With the latter approach, the polymer template chain length is unimportant. Following Damle <sup>18</sup>, we calculated the  $N \triangle H^o$ , value, where N is the number of monomer

units in the oligomer bound with the polymer template, from the slope of the  $1/T_m$  vs.  $lnC_m$  straight line ( $C_m$  is the free oligomer concentration at  $T_m$ ). This was found to be equal to -37 kcal/mol, i.e. near the  $\triangle H^{\circ}$  (n=2) value obtained above for the formation of four nucleotide pairs. In terms of this model, the apparent oligomer chain length is thus equal to 4 monomer units. Porschke and Eggers  $19$  suggested a thermodynamic model, which we have used for a more adequate description of concatemer formation. In terms of this model, the complex formation of octanucleotide d(T-G-C-A-C-A-T.-G) results in a set of complexes ranging from dimer to very long polymers:



```
Scheme
```
Let us assume that  $C_{A_1}$ ,  $C_{A_2}$ ,  $C_{A_3}$  ... are equilibrium concentrations of a free oligomer and the complexes formed by two, three etc. molecules of the octamer,  $C_{\bf a}$  is the concentration of the oligonucleotide chains, and  $C_m$  is the total concentration of all the complexes present in the solution. Then

$$
C_{a} = C_{A_{1}} + 2C_{A_{2}} + 3C_{A_{3}} + 4C_{A_{4}} + ...
$$
\n
$$
C_{m} = C_{A_{1}} + C_{A_{2}} + C_{A_{3}} + C_{A_{4}} + ...
$$
\n(3)

The parameter of the UV absorption of the concatemer system  
\n
$$
\mathcal{E} = \mathcal{E}_{\text{C}} \frac{c_{A_1}}{c_0} + \mathcal{E}_{\text{h}} \frac{c_{A_2} + 2c_{A_3} + 3c_{A_4} + \dots}{c_o} + \tag{5}
$$

+ 
$$
\mathcal{E}_c(c_{A_2} + c_{A_3} + c_{A_4} + \cdots)
$$

where  $\epsilon_{\rm c}$  and  $\epsilon_{\rm h}$  are the molar extinction coefficients of free

octamer and octamer in a double-stranded complexes, respectively.

It is obvious from equations  $(3)$ ,  $(4)$  and  $(5)$  that

and

$$
\frac{c_m}{c_a} = \frac{\mathcal{E}_h - \mathcal{E}}{\mathcal{E}_h - \mathcal{E}_C}
$$

 $\mathcal{E} \cdot C_A = \mathcal{E} \cdot \partial C_m + \mathcal{E} \cdot \partial C_a - C_m$ 

With 
$$
\frac{C_m}{C_a} = 1/2
$$
 (6)  
 $\mathcal{E} = \frac{\mathcal{E}_C + \mathcal{E}_h}{2}$ , which corresponds to the value of the UV

absorption at the melting point of the concatemer.

According to Porschke and Eggers  $^{19}$ , if condition (6) is fulfilled,  $K = \frac{2}{C}$ , where K is the equilibrium constant of a the process at  $\texttt{T}_{\texttt{m}}$ . From the slope of the lnK vs. 1/ $\texttt{T}_{\texttt{m}}$  straight line we find the enthalpy of the formation of four nucleotide pairs of the complex (this is the number of pairs that is formed on consecutive addition of the molecules of the octanucleotide), which is equal to  $37.0\frac{1}{1.2}$  kcal/mole. The value of  $\triangle H^{\circ}$ is exactly the same as that obtained with the help of less sophisticated models (see above).

CD studies. The CD spectra of double-helical polydeoxyribonucleotides are believed to depend on the kind of mononucleotide units and first neighbour frequencies, provided the species compared have geometrically identical helical structu $res$ <sup>20,21</sup>. Arnott<sup>22</sup> has proposed that it is the frequency of nucleotide triplet sequences that determines the CD characteristics of DNA duplexes. Some of authors  $23$  attribute the differences in the CD spectra of various DNA's to the differences in their double helix geometries. Collation of the literature data  $23-25$  shows that the CD spectra of double-helical polynucleotides containing repeating di(tri)nucleotide sequences and bearing mixed purines and pyrimidines on each chain provide a better fit to the CD spectra of DNA (characterized by approximately the same G\*C content) than the CD spectra of polymers with all purines on one chain and all pyrimidines on the other.

Apparently, the best DNA models are synthetic nucleotide sequences that resemble most the statistical DNA sequences,  $e_{\alpha}E_{\alpha}$ concatemer complexes of a heterogeneous composition. Actually, Fig. 2 shows that the conservative CD spectrum of the concatemer complex in the range of 230-290 nm with positive maximum near 270 nm resembles the CD spectrum of DNA-B form with 50%  $G^{\bullet}$ C content  $21$ . The differences in the CD spectra of the concatemer and quasi random sequence DNA (Fig. 2) are, probably, due to the fact that the synthetic complex has a regular structure and its CD spectrum should therefore bear a stronger resemblance to the spectra of satellite DNA's whose frequencies of hydrogen-bonded dimer pairs differ strongly from the random values  $21$ . The CD data are in accord with the suggestion that the concatemer complex has DNA-B-type helix. Such a structure appears to predominate in synthetic polydeoxyribo<sup>22</sup> and oligodeoxyribonucleotide 26 duplexes in solution.



Fig. 2. The CD spectra of d(T-G-C-A-C-A-T-G) at different temperatures and concentrations, a: -4° (1) and 35° (2),  $\bar{C}_{0}$  10<sup>-4</sup> M; b:- 4° (1) and 37° (2),  $C_{0}$  - 10<sup>-3</sup>M; G:  $\frac{1}{2}$  (1) and 45<sup>o</sup> (2), C<sub>0</sub> - 10<sup>-2</sup>M. For comparison, the CD spectrum of DNA from phage A containing 50% of C.G<br>pairs (dotted line) is given <sup>21</sup>.

The CD pattern remains practically the same within a 100-fold variation of the octamer concentration which therefore has no effect on the structure of the double-helical complex. Elevation of temperature causes a decrease in the absolute intensity of the negative CD band of d(T-G-C-A-C-A-T-G). The temperature dependence of  $\Delta \xi_{250}$  (Fig. 3), which is similar to that of  $A_{258}$  (Fig. 1), reflects the cooperative process of the melting of the complex. Temperature increase affects the shape of the positive CD band (Fig. 2) in the same manner as denaturation of the DNA double helix affects the respective feature in its spectrum  $20$ . Similarly to native DNA  $27,28$  the temperature dependence of  $AC_{280}$  (Fig. 3) does not afford data on the helix-coil transition region. The dissimilarity in the temperature dependences of  $AC_{280}$  between the concatemer and DNA  $^{28,29}$  near  $T_m$  may be due to difference between the helixcoil transition temperatures in these systems and also differences in the final "melted" products.

The concatemer thermodynamic and optical characteristics suggest that association of  $d(T-G-G-A-C-A-T-G)$  leads to pseudopolymer DNA-like helices whose geometry resembles that of DNA-B form.



Temperature de-<br>pendence of Pig. 3. Temperature d<br>pendence of<br> $\alpha$ C<sub>280</sub> and  $\alpha$ C<sub>250</sub> for<br>d(T-G-C-A-C-A-T-G) at<br>different oligonucleo different oligonucleo- $\lambda$ 280 HM  $10^{-4}$  M (1),  $10^{-3}$ M (2).<br> $\lambda$ 280 HM and  $10^{-2}$ M (3).

# 2. Chemical polymerization of <sup>32</sup>P-d(pT-G-C-A-C-A-T-G) as concatemer constituent

The peculiar features of the aecondary structure of the concatemer synthesized by us make them handy models for studying template-guided reactions. As the first example, we chose chemical self-condensation of concatemeric  $32P-d(pT-G-C-A-C-A-$ -T-G) induced by water-soluble carbodiimide (ClEC). CIEC was used previously for template-guided joining of homogeneous oligonucleotides  $30-32$ , Samples of  $32P-d(pT-G-C-A-C-A-T-G)$  were prepared by phosphorylation of  $d(T-G-G-A-C-A-T-G)$  with  $y - \frac{32}{P}$ -ATP and polynucleotide kinase  $^6$  and isolated by gel filtration on Sephadex G-50. Polyacrylamide gel electrophoresis of the preparations thus made showed that they were practically free of  $r = \frac{32}{P+AP}$ , the yield amounting to 80%. The phosphorylation product according to the electrophoresis data contained about 20% of nucleotide admixtures. Since the latter did not interfere with further analysis of template-guided condensation products, no additional purification of the octanucleotide was made. Condensation of  $^{32}$ P-d(pT-G-C-A-C-A-T-G) was carried out in 0.05 M MES (pH 6.0), 0.02 M MgCl<sub>2</sub>, 0.1 M CMEC,  $C_0 - 10^{-3}$ M.  $T_m$  of the concatemer under these conditions decreases to about 13°C. With this in mind, the reaction mixture was kept at O°C. The products of octanucleotide condensation were detected in the reaction mixture 36 hr after addition of CIEC. The autoradiogram of the slab gel after electrophoresis of the mixture is shown in Fig. 4a. Together with the sample to be analyzed, the gel was loaded with  $32P-d(pT-G-C-A-C-A-T-G)$  treated with  $T4$ polynucleotide ligase under the concatemer formation conditions, Fig.4b and c (the latter reaction was carried out by Dr. P.M. Rubtsov). The bands of the radioactive substances from CIECinduced condensation of the  $3^2$ P-d(pT-G-C-A-C-A-T-G) (Fig. 4a) and from ligase-linked octanucleotide (Fig. 4c) move close together. That proves identity of the chemical and enzymatic condensation products. In a reference experiment (in the absence of CMEC or polynucleotide ligase), the autoradiogram of polyacrylamide gel showed no indications of the presence of higher molecular weight products than the initial octanucleotide (data



not shown). Unlike the enzymatic reaction, which gives almost quantitative yields of polymeric products, "chemical ligation" produces mostly a dimer, a 16-membered oligonucleotide. A 13% yield of the latter was obtained after 6 days. The low efficiency of chemical joining is, probably, the result of insufficient concatemer stability under the reaction conditions and low reactivity of the 3'-OH group (according to Popov and Shamovsky  $32$ , the vield of the products of template-guided condensations increases greately if the primary alcohol 5'-hydroxyl group acts as acceptor of an activated phosphate group).

After the product of chemical condensation, the dimer, was eluted from gel, it was subjected to a more detailed analysis. The isolated hexadecanucleotide was treated with bacterial alkaline phosphatase and analyzed by electrophoresis in polyacrylamide gel together with a reference ample of the reaction mixture not treated with the enzyme. As seen from Fig. 5, the two samples showed different mobilities. The data obtained are indicative of, firstly, covalent joining of octamers (as follows from label retention after treatment with the enzyme), and, se-



Pig. 5. Polyacrylamide gel electrophoresis of the reactiQ mixture containing <sup>32</sup>Pd(pT-G-C-A-C-A-T-G) ,0. M1 CMEC,  $0.02$  M MgCl<sub>2</sub> and 0.05 lMES Cpf 6.0) (a) and of the major product of chemical joining, hexadecanucleotide, treated with baoterial alkaline phosphatase (b). Length markers were obtained by degradation of DNA fragment according to Maxam and Gilbert  $10$ :  $A > G$  (C) and  $C + T$ ( d ). Chain lengths of the DNA fragments are indicated with numbers (e.g. 10 corresponds to decanucleotide). The position of marker dyes, xylene cyanol and bromphenol blue, are shown.

condly, the presence of a terminal phosphate group in the hexadecanucleotide synthesized (i.e. the "chemical ligation" product contains a natural  $3'-5'$  phosphodiester bond).

The chain lengths of the oligonucleotides synthesized and their mobilities in polyacrylamide gel were estimated by comparison with reference oligonucleotides. A mixture of reference oligonucleotides differing by one mononucleotide residue was obtained by chemical degradation of a native DNA fragment according to Maxam and Gilbert  $^{10}$ . A DNA fragment was treated in two separate reactions, with the sequence being cleaved at A> G (Fig. 5c) and  $C + T$  (Fig. 5d). The mobilities of both the chemical ligation (Pig. 5) and enzymatic ligation products in polyacrylamide gel were different from those of the reference 16-, 24-, etc. membered oligonucleotides that should be formed in consecutive addition of octamers. The ligase joining products migrate at an interval of four monomer units (the band positions in Pig. 4 were compared with those of chain length markers). A similar tendency is observed for the products of the chemical ligation (Fig. 5). This phenomena may be due to intermolecular aggregation of the octanucleotide condensation products. Although gel was 7 M in urea and was not cooled, the possibility that the secondary structure was formed could not be ruled out  $33^*$ . Each ligation product in the concatemer system, e.g. hexadecanucleotide, may be thought to form a number of different intermolecular aggregates showing different mobilities in electrophoresis. The specific feature of these complementary aggregates is the presence of identical cohesive ends and consecutive elongation of the double-helical fragment by four nucleotide pairs, similarly to the structures shown in the scheme.

The physico-chemical properties of the concatemer described and the data on the chemical and enzymatic linking of octanucleotides given above indicate that the use of these systems for template-guided reactions is very promising. Approaches involving synthetic concatemer complexes are being developed in this laboratory.

#### **ACKNOWLEDGEMENTS**

The authors wish to express their thanks to Dr. P.M. Rubtsov and Dr. LG. Skryabin, who provided some of the enzymes used in this work and carried out enzymatic ligation reactions and to Dr. A.I. Dolinny for his help in the calculation of thermodynamic parameters.

Correspondence should be sent to Prof. Z.A. Shabarova

#### REFERENCES

- 1 Naylor, R. and Gilham, P.T. (1966) Biochemistry 5, 2722-2728<br>2 Shabarova, Z.A. and Prokofiev, M.A. (1970) FEBS Lett. 11.
- Shabarova, Z.A. and Prokofiev, M.A. (1970) FEBS Lett. 11, 237-240
- 3 Gennis, R.B. and Cantor, C.R. (1970) Biochemistry 9, 4714-4722
- 4 Bahl, C.P., Wu, R., Brousseau, R., Sood, A.K., Hsiung, H.M.<br>and Narang, S.A. (1978) Biochem.Biophys.Res.Comm. 81, 695-703
- 5 Dobrynin, V.No, Korobko, V.G., Severtsova, I.V., Bystrov, N.S., Boldyreva, E.F., Chernov, B.K. and Kolosov, M.N. (1978) Bioorgan.Khim. 4, 1600-1610
- 6 Sekiya, T., Besmer, P., Takeya, T. and Khorana, H.G. (1976) J.Biol.Chem. 251, 934-641
- 7 Dolinnaya N.G., Gromova, E.S., Mikhailov, S.N. and Shabarova, Z.A. t1978) Bioorgan.Khim. 4, 535-549.
- 8 Turkin, S.I., Yamschikov, V.F., Potapov, V.K., Shabarova, Z.A. and Prokofiev, M.A. (1979) Dokl.AN SSSR 245, 614-617
- 9 Abbreviations: CMEC, N-cyclohexyl-N'- $\beta$  -(N-methylmorpholi- nio)ethyl] carbodiimide p-toluene sulfonate; MES, 2-morpholinoethanesulfonate
- 10 Maxam, A.M. and Gilbert, W. (1977) Proc.Natl.Acad.Sci USA  $74$ , 560-564 74, 590-564 <sup>11</sup> Porschke, D., Uhlenbeck, 0.C. and Martin, F.1H. (1973) Biopo-
- lymers 12, 1313-1335
- 12 Pohl, F.11. (1974) Eur,J.Biochem. 42, 495-504
- 13 Borer, P.N., Dengler, B. and Tinoco , I., Jr. (1974) J.Mol. Biol. 86, 843-853
- 14 Privalov, P.L. (1969) Molecul.Biol. 3, 690-695
- 15 Porschke, D. (1971) Biopolymers 10, 1989-2013
- 16 Riesner, D. and Romer, Ro (1973) in Physico-chemical Properties of Nucleic Acids (Duchesne, Jo, ed.), vol. 2, pp. 237-318, Academic Press, London- New York
- <sup>17</sup> Hoggett, J.G. and Maass, G. (1971) Ber. Bunsenges. Phys. Chem. 75, 45-54
- 18 Dam3le, VoN. (1970) Biopolymers 9, 353-372
- 19 Porschke, D. and Eggers, F. (1972) Eur.J.Biochem. 26, 490- 498
- 20 Allen, F.S., Gray, D.M., Roberts, G.P. and Tinoco, I., Jr. (19725 Biopolymers 11, 853-879
- 21 Marck, C. and Guschlbauer, W. (1978) Nucleic Acids Res. 5, 2013-2031
- 22 Arnott, S. (1975) Nucleic Acids Res. 2, 1493-1502<br>23 Gray, D.M., Morgan, A.R., Ratliff, R.L. (1978) Nu
- 23 Gray, D.M., Morgan, A.R., Ratliff, R.L. (1978) Nucleic Acids Res. 5, 3679-3695
- 24 Wells, R.D., Tarson, J.E., Grant, R.C., Shortle, B.E. and Cantor, C.R. (1970) J.Mol.Biol. 54, 465-497
- 25 Zimmer, C. and Luck, G. (1974) Biochim.Biophys.Acta 361,  $11 - 32$
- 26 Patel, D.J. and Tonelli,  $A.E.$  (1974) Biopolymers 13, 1943-<br>1964. 1964.
- 27 Pale6.k, E. and Pri6, I. (1972) Biochem.Biophys.Res.Co=n., 47, 1262-1269
- 28 Vsatyi, A.F. and Shlyakhtenko, L.S. (1973) Biopolymers 12, 45-51
- 29 Gennis, R.B. and Cantor, C.R. (1972) J.Mol.Biol. 65, 381-<br>399 399
- 30 Badashkeeva A.G., Gorbunov, N.P., Shamovsky, G.G. and Shubina, T.N. 61974) Izvestia Sibirakogo otdelenia AN SSSR No. 2, 96-102
- 31 Uesugi, S. and Ts'o,P.O.P. (1974) Biochemistry 13, 3142- 3152
- 32 Popov, S.G. and Shamovsky, G.G. (1976) Izvestia Sibirskogo otdelenia AN SSSR, No. 7, 14-19
- 33 Mills, D.R. and Kramer, F.R. (1979) Proc. Natl.Acad.Sci USA 7& 2232-2235