Volume 9 Number 21 1981 Nucleic Acids Research

DNA-like duplexes with repetitions. III. Efficient template-guided chemical polymerization of d(TGGCCAAGCTp)

Z.A.Shabarova, N.G.Dolinnaya, V.L.Drutsa, N.P.Melnikova and A.A.Purmal

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

Received 13 August 1981

ABSTRACT

Self-association of a decanucleotide d(TGGCCAAGCTp) in an aqueous solution is shown by UV spectroscopy, CD and sedimentation analysis to yield a pseudopolymeric (concatemeric) duplex having a geometry similar to that of DNA B-type. It is demonstrated that in conditions when the concatemeric duplex is stable a water-soluble carbodiimide induces efficient polymerization of the 3'- or 5'-phosphorylated decanucleotide, and the resulting polymers $d(TGGCCAAGCTp)_{2-10}$ contain only natural phosphodiester bonds. In conditions optimal for template-guided polymerization of d(TGGCCAAGCTp) the overall yield of 20-100-member polynucleotides exceeds 90%. The obtained polymeric duplexes are cleaved by restriction endonuclease _Alu_ I, _Bsu_ RI, and _Hind_ III to corresponding decamers which were isolated and sequenced.

INTRODUCTION

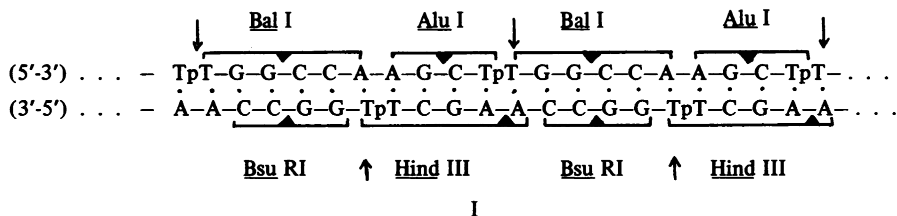
Creation of artificial DNA from synthetic oligonucleotide blocks brought together by complementary interactions on a template involves the enzyme DNA ligase. Development of techniques for chemical "ligase" reaction[1-4] would not only provide an alternative way for the synthesis of double-stranded polydeoxyribonucleotides, but would also open the way of constructing compounds with variously modified elements, including modified internucleotide bonds. Promising approaches are the activation of the terminal phosphate group directly in the duplex with condensing agents such as water-soluble carbodiimides, as well as the use of oligonucleotide derivatives with activated phosphate groups like amidates, pyrophosphates etc.

Expedient models for the development of the chemical ligation technique are the synthetic DNA duplexes with regularly repeated base-pair fragments (concatemeric duplexes) which we have described previously.[5] Such pseudopolymeric duplexes result from self-association of one or several synthetic oligonucleotides of certain primary structure, and have non-coincident interruptions in the complementary chains. These duplexes may contain sense repetitions, e.g. palindromes occurring in gene regulatory regions, sites recognized by restriction endonucleases etc.

A complex formed by an octanucleotide d(TGCACATG) was used to characterize a simplest concatemeric system with symmetrically overlapping blocks.[5] The present work describes a new concatemeric system of asymmetrically overlapping blocks which originates from a decanucleotide d(TGGCCAAGCTp) and contains sites recognized by several restriction endonucleases (duplex I;

© IRL Press Limited, 1 Falconberg Court, London W1V 5FG, U.K.

symbol "d" is omitted):



Optimal conditions were found for water-soluble carbodiimide-induced chemical polymerization of 3′- or 5′ -phosphorylated decanucleotide in the form of concatemeric duplex, and the products of chemical ligation of d(TGGCCAAGCTp) were shown to be cleaved by corresponding restriction endonucleases.

MATERIALS AND METHODS

General. Decanucleotide d(TGGCCAAGCTp) was synthesized in solution by the triester method as described.[6] Snake venom phosphodiesterase (VPDE) and E. coli alkaline phosphatase (BAP) were from Worthington Biochemical Corp.; T4 polynucleotide kinase, restriction endonucleases Bsu RI and Hind III – from VPO Biopreparat (Vilnius, USSR); CDI I, CDI II and MES[7] – from Merck; Aminosilochrom resin – from Soyuzkhimreactiv (Novosibirsk, USSR). Restriction endonuclease Alu I was kindly provided by Prof. H.-J.Fritz (Koln, FRG).

Optical measurements. Prior to optical assays the decanucleotide was additionally purified by paper chromatography in n-propanol: conc. NH_3:H_2O (55:10:35 v/v). Calculated amounts of oligo- or polynucleotides were dissolved in one of the following buffers: 4mM phosphate buffer, pH 7.0, 0.2 M NaCl, 75 mM $MgCl_2$ (buffer A); 50 mM MES, pH 6.1, 20 mM $MgCl_2$ (buffer B); 5 mM Tris-HC1, pH 7.4, 1 mM EDTA, 0.1 mM dithioerythritol (buffer C). Nucleotide concentration was determined spectrophotometrically. The molar extinction coefficient, ϵ_{260} (per monomer), of d(TGGCCAAGCTp) determined upon exhaustive hydrolysis with a mixture of VPDE and BAP, was 8980 M^{-1} cm^{-1}. Molar extinction coefficients of d(pTGGCCAAGCT) and oligonucleotide polymerization products were taken equal to ϵ_{260} of d(TGGCCAAGCTp). photometer; CD spectra were recorded with a Roussel-Jouan II dichrograph. Dependences of UV absorption and CD on temperature were studied in thermostatted quartz cuvettes (Hellma, FRG) with path lengths of 20, 10, 2, 1, and 0.1 mm. Temperature was monitored with a copper-constantan thermocouple. Hypochromicity (h) resulting from complex formation was calculated by equation:

$$h = \frac{A_{45} - A_{-5}}{A_{45}} \quad 100\%,$$

where A_{45} and A_{-5} are the absorbances of the decanucleotide solution at 45 and -5°C (λ = 256nm).

Sedimentation analysis. Sedimentograms of d(TGGCCAAGCTp) and polymerization products were obtained in a Beckman Model E analytical ultracentrifuge at 5, 20, and 35°C. Concentration of oligo- or polynucleotides per monomer, C_o, was 0.5-1.0 · 10^{-4}M. Solvents used were buffer A, buffer B, and 0.1 M NaOH in 0.9 M NaCl.

Chemical template-guided oligodeoxyribonucleotide polymerization. Condensing agents used were water-soluble carbodiimides I and II (CDI I and CDI II). Polymerization of d(TGGCCAAGCTp) induced by CDI I was carried out in conditions listed in Table 1. Carbodiimide was added to solutions of the decanucleotide in buffer B (or buffer B diluted by 20% with isopropanol) cooled to 0°C. Reaction mixtures were incubated at 0 or -10°C (see Table 1). After 8 days from the onset of polymerization, reaction mixtures II and IV were supplemented with fresh portions of CDI I to a total concentration of 0.04 M.

CDI II-induced polymerization of d(TGGCCAAGCTp) was performed in buffer B at 0°C with $C_o = 10^{-3}$ and CDI II concentration of 0.02 M. The same conditions were used to chemically polymerize d(TGGCCAAGCTTGGCCAAGCTp) and d(pTGGCCAAGCT). The reaction mixture containing the 5'-phosphorylated decanucleotide was 8 days after the start of reaction treated with additional CDI II to make it 0.04 M in CDI II.

Reaction mixtures were analyzed by microcolumn chromatography on Aminosilochrom resin (separation conditions see in legend to Fig. 3) at certain times after carbodiimide addition. Yields of polymerization products were determined as the relation of corresponding peak area to the total peak area of the initial oligonucleotide and all polymerization products. Products of decanucleotide d(TGGCCAAGCTp) polymerization were isolated on a preparative scale by chromatography of the reaction mixtures on an Aminosilochrom 0.2x10 cm column in a linear gradient of Na-phosphate, pH 7.0, containing 7 M urea. Isolated polymeric fractions were desalted on DEAE-cellulose 0.2x2 cm column in HCO_3^- form. Nucleotide material was eluted with 1 M ammonium bicarbonate.

The nucleotide sequence of $d(TGGCCAAGCTp)_2$ was determined by a modified method of Maxam and Gilbert.[8]

The primary structure of decanucleotides produced by cleaving the polynucleotide duplexes with restriction endonucleases was analyzed by nucleotide mapping.[9]

Enzymic methods. Hydrolysis of oligonucleotides (0.5 OD unit, 260 nm) with BAP (1 act. unit) was carried out for 3 hr at 37°C in 40 mM NH_4HCO_3, pH 8.5, 20 mM $MgCl_2$, 0.02 mg/ml serum albumin, in a total volume of 50 μl. A mixture of BAP (0.1–1.0 act. unit) and VPDE (0.02–0.2 act. unit) was used to hydrolize oligonucleotides (0.01–0.5 OD unit, 260 nm) for 16 hr at 37°C or for 3 hr at 50°C in 40 mM NH_4HCO_3, pH 8.5, 20 mM $MgCl_2$, 0.2 mg/ml serum albumin (v = 50 μl). Polynucleotides (0.04 OD unit, 260 nm) were treated with restriction endonucleases Bsu RI (20 act. units), Alu I (26 act. units), or Hind III (20 act. units) in 25 mM Tris-HCl, pH 7.5 containing 10mM $MgCl_2$, 50 mM NaCl and 10 mM 2-mercaptoethanol (v = 10μl).

Phosphorylation of d(TGGCCAAGCT) (1 OD unit, 260 nm) with T4 polynucleotide kinase (10 act. units) was performed for 15 min at 37°C in 50 mM Tris HCl, pH 9.0, containing 10 mM $MgCl_2$, 5 mM 2-mercaptoethanol, 2 mM spermine and 1 mM ATP (v = 100μl).

RESULTS AND DISCUSSION

1. Physico-chemical properties of the concatemeric duplex formed by d(TGGCCAAGCTp)

UV spectroscopy, CD, and sedimentation analysis were applied to the determination of thermal stability, geometry and length of the concatemeric duplex formed by the decanucleotide d(TGGCCAAGCTp).

Fig. 1 depicts the temperature dependences of UV absorption of the decanucleotide dissolved in buffer A at various concentrations. The shape of melting curves, considerable hypochromicity, and concentration dependence of melting temperatures (T_m) indicate the formation of an intermolecular complex. Addition of isopropanol (20% v/v) entails a practically complete dissociation of this complex already at 0°C. The thermal stability of the concatemeric duplex is obviously enhanced with increasing nucleotide concentration C_o (Fig. 1). Hypochromicity resulting from self-complexation of d(TGGCCAAGCTp) is virtually independent of nucleotide concentration and constitutes 21-22%.

The CD spectrum of the concatemeric duplex formed by d(TGGCCAAGCTp) is conservative in the range 230-290 nm with a zero transition point at about 260 nm (Fig. 2), which is characteristic of DNA B-type.[10]

The CD spectrum of the complex does not appreciably change with varying nucleotide concentration from 10^{-4}M to 10^{-2}M, thus testifying to the identical geometry of helices produced at different nucleotide concentrations. Increase of temperature to the region of T_m evokes a change in the shape of the positive CD band (Fig. 2) indicating dissociation of the complex. As in the case of

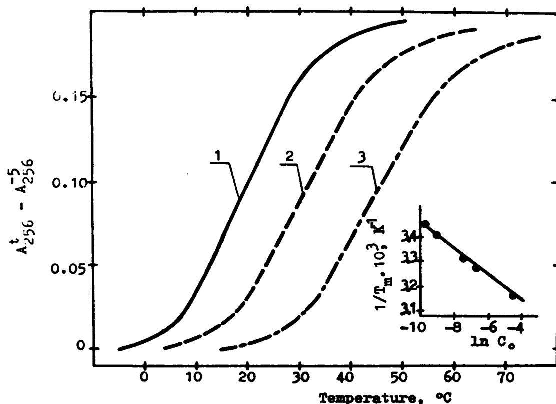


Fig. 1 Melting profiles of d(TGGCCAAGCTp) in buffer A; C_o= 10^{-4}M (1), 10^{-3}M (2) and 10^{-2} M (3). Insert: plot of $1/T_m$ against ln C_o.

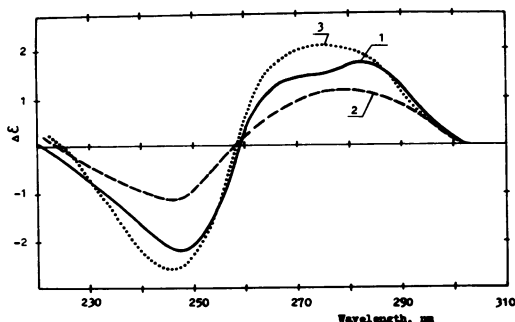


Fig. 2 CD spectra of d(TGGCCAAGCTp) at 5°C (1) and 60°C (2); $C_0=10^{-2}$M, buffer A. The CD spectrum of phage λ DNA containing 50% G · C pairs (3) is given for comparison.[11]

the concatemeric duplex formed by d(TGCACATG)[5], the absolute values of the positive and negative CD bands of the decanucleotide are lower than that of DNA with the same G · C content (Fig. 2) Provided native DNA and concatemeric duplexes are isogeometric, the CD spectra of the latter may be highly informative as the source of input data for computation of the first neighbour frequencies in various DNAs. Indeed, owing to periodic repetition of nucleotide sequence in the concatemeric duplex the frequencies of nearest neighbour base pairs differ from random values characteristic of most natural DNAs. In this respect the concatemeric duplexes resemble satellite DNA[11] and may serve as DNA models with known first neighbour frequencies.

The length of duplexes produced was estimated by determining the sedimentation coefficients of the decanucleotide in conditions favouring intermolecular complex formation. Nucleotide concentration C_0 was $1.0 \cdot 10^{-4}$ M or $0.5 \cdot 10^{-4}$ M in either buffer A or buffer B, temperature 4°C; in all systems used the sedimentation coefficients were nearly identical and equal to 2.7± 0.2 S (20°C). Such sedimentation coefficients are characteristic of DNA fragments comprising 32-35 base pairs[12] and point therefore to association of 6-8 decanucleotides. The sedimentation data indicate an appreciable homogeneity of the complexes in molecular mass, i.e., the absence of a wide set of duplexes of variable length. In conditions denaturing the complementary complex at 35°C, d(TGGCCAAGCTp) sediments with a coefficient of 1.0± 0.1 S (20°C) which corresponds to the single-stranded decanucleotide.[13]

Thus, several independent physico-chemical methods provide evidence that self-association of the decanucleotide d(TGGCCAAGCTp) in aqueous solutions yields thermostable DNA-like duplexes I of appreciable length.

Further detailed studies on the peculiarities of the geometry, thermodynamics and kinetics of formation of concatemeric duplexes are of self-contained theoretical interest, since such complexes with palindromic repetitions directly mimic satellite DNAs and eukaryotic genome regions with high degree of repetition.

2. Template-guided chemical polymerization of decanucleotides in the concatemeric duplex

The most known way of chemical ligation is the template-guided polymerization of mono- or

oligonucleotides induced by water-soluble carbodiimide.[3, 14, 15] The mechanism of this reaction was studied mainly on triple helical complexes of homogeneous oligomers and polymers [e.g. oligo (A) · 2 poly(U)]. However, because of the irregular mutual orientation of oligonucleotides in this system the predominant condensation product is the symmetric pyrophosphate of the oligonucleotide.[14, 15] Carbodiimide linking of homogeneous and heterogeneous oligonucleotides in some double-helical complexes [4, 16, 17] has demonstrated the expedience of this approach. To assess the applicability of this technique to the synthesis of sufficiently long polydeoxyribonucleotides of defined sequence the present work made use of the above-characterized concatemeric duplex I formed by d(TGGCCAAGCTp). Duplex I has two substantial advantages over the duplex of d(pTGCACATG) studied previously.[5] Firstly, the decanucleotide yields a more thermostable duplex; secondly, d(TGGCCAAGCTp) bears the 3'-terminal phosphomonoester group, i.e., the activated 3'-phosphate is accepted by a much more reactive 5'-hydroxyl of the adjacent substrate molecule (see e.g. [16]). Conditions of chemical ligation were optimized using CDI I as a condensing agent.

The conditions of chemical ligation of the decanucleotide are given in Table 1.

A preliminary study was made of the stability of the concatemeric duplex under conditions of template-guided polymerization. T_m of the duplex in reaction mixture II was found to be 26°C (6°C lower than that in buffer A). Hypochromicity of complex formation does not change therewith. A 5-fold increase in CDI I concentration (reaction mixtures I and III) entails pronounced carbodiimide modification of the heterocyclic bases in the decanucleotide during melting. Introduction of isopropanol into the solution of d(TGGCCAAGCTp) (reaction mixture V) virtually destroys the complementary complex already at 0°C (this mixture was used as a control).

Reaction mixtures were analyzed by microcolumn chromatography on Aminosilochrom resin. Accumulation of the products of template-guided d(TGGCCAAGCTp) polymerization in reaction mixtures I-IV was observable already in several hours. Chemical ligation was the most efficient in

TABLE 1

Conditions of CDI I-induced template-guided polymerization of d(TGGCCAAGCTp) in buffer B

Reaction mixture No.	C_o, M	CDI I concn., M	Incubation temp., °C
I	10^{-3}	0.10	0
II	10^{-3}	0.02	0
III	10^{-2}	0.10	0
IV	10^{-3}	0.02	-10
V*	10^{-3}	0.02	0

* Incubation in buffer B diluted by 20% with isopropanol.

reaction mixture II (chromatographic pattern can be found in [18]). The overall yield of polymeric products reached 70% in 12 days of incubation, products with four, five, and more links prevailing in the reaction mixture. It should be noted that after 6-8 days of incubation at 0°C the chemical ligation slowed down owing to carbodiimide hydration. Reaction in mixture II was after 8 days boosted by adding fresh CDI I to a cumulative concentration of 0.04 M, which raised the total yield of polymeric products from 50 to 80%. In reaction mixtures I and III, polymerization of d(TGGCCAAGCTp) was superimposed on notable carbodiimide modification of thymine and guanine bases in the decanucleotide. Modification products differ markedly from the original decanucleotide in chromatographic mobility and are readily detectable by microcolumn chromatography on Aminosilochrom resin.

Freezing of the reaction mixture (IV) retarded somewhat the template-guided polymerization. In discord with the data of Uesugi and Ts'o[4], lowering of incubation temperature to -10°C did not practically affect the length of chemical ligation products. Only negligible dimerization of the decanucleotide occurs in reaction mixture V where the complementary complex is not formed, which points to the template-dependent nature of polymerization in reaction mixtures I- IV.

The comparison between the action of CDI I and CDI II in optimized conditions demonstrated the higher efficiency of CDI II. Fig. 3 presents the chromatographic separation profiles of reaction mixtures containing different carbodiimides after 48 hr from the onset of reaction. Peaks 2, 3 and 4 correspond to the products of d(TGGCCAAGCTp) polymerization: di-, tri-, and tetramers. The higher efficiency of CDI II is obvious.

The chromatographic patterns of d(TGGCCAAGCTp) polymerization mixture after various periods of time from CDI II addition are depicted in Fig. 4. Gradual accumulation of the products with increasingly higher molecular weights is observable with a concomitant decrease in the relative amounts of di-and trimers of the decanucleotide. After 90 hr the overall yield of decanucleotide polymerization products approximates 90%. Formation of internucleotide bonds beyond the con-

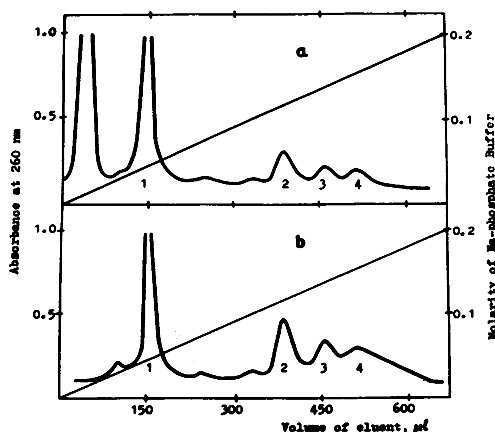


Fig. 3 Aminosilochrom chromatography (gradient of sodium phosphate, pH 7.0, in 7 M urea) of reaction mixtures containing 10^{-3} M d(TGGCCAAGCTp) and 0.02 M of either CDI I (a) or CDI II (b) in buffer B, after 2 days of incubation at 0°C. Peak 1, d(TGGCCAAGCTp); the peak of CDI I preceding the decanucleotide can be seen in Fig. 3a (CDI II has no UV absorption).

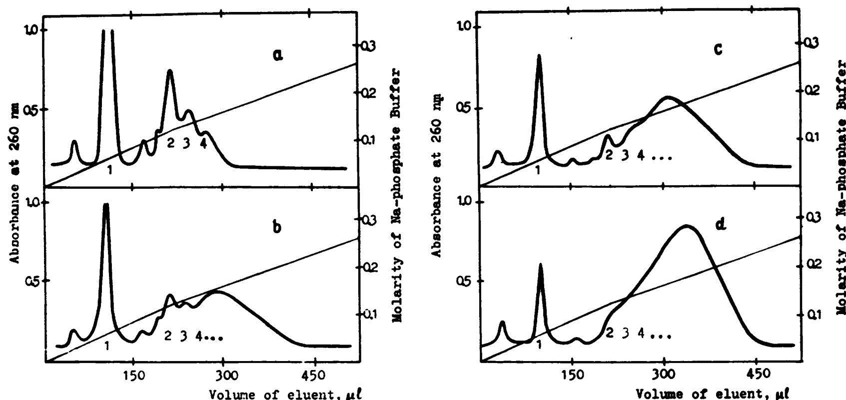


Fig 4. Chromatography (see Fig. 3 caption) of reaction mixtures containing 10^{-3} M d(TGGCCAAGCTp) and 0.02 M CDI II in buffer B after 1(a), 2(b), 3(c), and 4(d) days of incubation at 0°C. Peaks: 1, d(TGGCCAAGCTp); 2, 3, 4 etc., polymerization products.

catemeric duplex can be ruled out since in this case decanucleotide condensation would majorly yield a symmetric pyrophosphate. Lack of modification of thymine and guanine bases by carbodiimide (Fig. 4) also argues in favour of practically all decanucleotide molecules being integrated in the complementary complex.[19]

To elucidate the nature of the internucleotide bond produced in chemical ligation of d(TGGCCAAGCTp), a substance corresponding to the decanucleotide dimer (e.g. peak 2 on Fig. 4) was isolated and analyzed. The body of evidence presented below unequivocally testifies to this compound being an icosanucleotide with natural 3'-5' phosphodiester bonds.

(1) Treatment of the decanucleotide dimer with a combination of VPDE and BAP at 50°C for 3 hr produced a mixture of nucleosides. This directly indicates that the chemical ligation product contains only natural 3'-5' phosphodiester bonds. Treatment of the decanucleotide dimer with the mixture of the above enzymes at 37°C produces a compound corresponding in chromatographic mobility to a 16-member oligonucleotide (Fig. 5). This result was interpreted as follows.

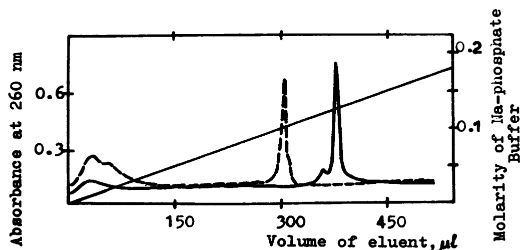
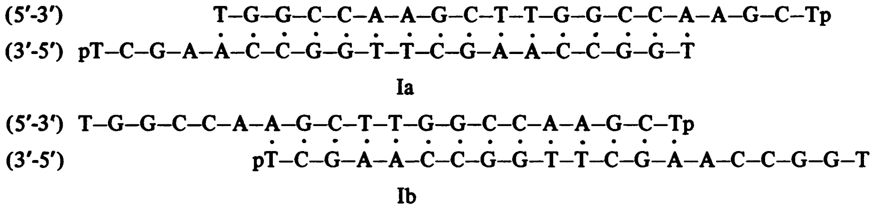


Fig. 5 Chromatography (see Fig. 3 caption) of $d(TGGCCAAGCTp)_2$ before (solid line) and after (broken line) treatment with a mixture of VPDE and BAP at 37°C

In a neutral buffer solution the icosanucleotide exists as an equilibrium mixture of thermostable duplexes Ia and Ib (symbol "d" is omitted)



and their longest concatemeric associates formed at the expense of the cohesive single-stranded ends. At 37°C the VPDE and BAP blend digests only the protruding 3'-ends of duplex Ia (duplex Ib is not affected by VPDE), producing thereby the 16-member self-complementary hexadecanucleotide d(TGGCCAAGCTTGGCCA).

(2) The complementary complex formed by self-association of the decanucleotide dimer is marked by high thermal stability. T_m of $d(TGGCCAAGCTp)_2$ in buffer C is 65°C even at C_0 of 10^{-4} M. The 3'-3' pyrophosphate of the decanucleotide, on the contrary, is incapable of forming thermostable complementary complexes.

(3) The isolated dimer of the decanucleotide was used as substrate for chemical ligation. As can be seen from Fig. 6, $d(TGGCCAAGCTp)_2$ is readily polymerized by CDI II to yield products comparable in length with polymers formed upon chemical ligation of the decanucleotide.

(4) Sequencing of the decanucleotide dimer according to Maxam and Gilbert is presented in Fig. 7. The clearly readable 14-member 5'-terminal sequence containing a (5')...-dC-dT-dT-dG...(3') fragment proves the formation of a natural internucleotide bond in the template-guided chemical polymerization of d(TGGCCAAGCTp).

The sequence of the 3'-terminal region cannot be read unambiguously because of pronounced

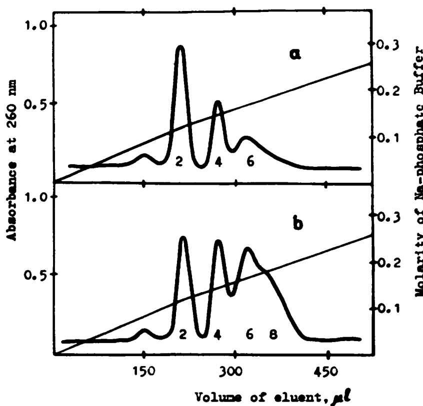


Fig. 6 Chromatography (see Fig. 3 caption) of reaction mixtures containing 10^{-3} M $d(TGGCCAAGCTp)_2$ and 0.02 M CDI II in buffer B after 1(a) and 2 (b) days of incubation at 0°C. Peaks: 2, $d(TGGCCAAGCTp)_2$; 4, 6, and 8 are tetra-, hexa-, and octamers of d(TGGCCAAGCTp).

self-association of long (about 15 members) products of icosanucleotide degradation during electrophoresis; such an effect of "band compression" has been reported.[20, 21]

Obviously, the products of d(TGGCCAAGCTp) chemical ligation with higher molecular weight as well contain natural phosphodiester bonds. This is also confirmed by the total cleavage of polynucleotide duplexes by restriction endonucleases (see section III).

The length of decanucleotide polymerization products was estimated by sedimentation analysis. The sedimentation coefficient of the total polymeric fraction under conditions disrupting complementary interactions (0.1 M NaOH in 0.9 M NaCl) is 2.7± 0.2 S(20°C) which corresponds to single-stranded polynucleotides of 60-70 monomeric units on average.[12] In conditions favouring the formation of complementary complexes (buffer A) the sedimentation coefficient of the same specimen increases to 5.4± 0.2 S (20°C). The latter value corresponds to DNA duplexes comprising 240-260 nucleotide pairs[12], i.e. products of self-association of 6-8 above polynucleotide molecules. These data suggest that chemical ligation of the decanucleotide is accompanied by

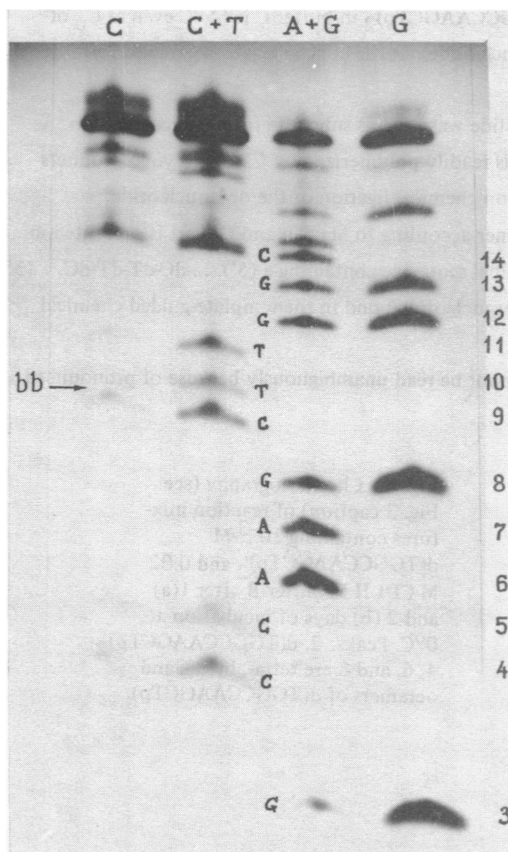


Fig. 7 Sequence analysis of $d(TGGCCAAGCTp)_2$ by a modified method of Maxam and Gilbert.[8] The products of selective chemical degradation of the icosanucleotide were electrophoresed in 20% polyacrylamide gel, 50 mM Tris-borate, pH 8.3, 0.1 mM EDTA, 7 M urea. Position of the bromophenol blue marker is indicated by arrow. Numbers on the gel photograph denote the consecutive numbers of corresponding nucleotides counting from the 5'-terminus of the icosanucleotide. The leading mono- and dinucleotide were allowed to migrate out of the gel to provide increased resolution of larger fragments.

elongation of the concatemeric duplex. It must be noted that, despite the different rates of chemical template-guided d(TGGCCAAGCTp) polymerization obtained with CDI I and CDI II, the length of resultant polymers differ only slightly.

To compare the reactivity of 3'- and 5'- phosphorylated oligonucleotides, d(pTGGCCAAGCT) was obtained by enzymic phosphorylation of the 5'-terminus of the dephosphorylated decanucleotide. As should have been expected, the template-guided CDI II-induced polymerization of this decanucleotide proceeds substantially more slowly than in the case of the 3'-phosphorylated one. However, the yield of polymeric products increases with time, and the reaction is intensified by addition of fresh carbodiimide after 8 days (data not shown). It must be noted that though the overall yield of d(pTGGCCAAGCT) polymerization products reaches 70% the bulk of these contain 3-4 internucleotide linkages, i.e. the degree of polymerization of d(pTGGCCAAGCT) is lower than with d(TGGCCAAGCTp).

Accumulation of chemical ligation products in all systems studied is displayed in Fig. 8.

UV spectra of the products of decanucleotide polymerization do not differ from those of d(TGGCCAAGCTp), but T_m of the duplexes formed by decanucleotide polymers is extremely high and in buffer C exceeds 80°C. This peculiarity of the "linked" concatemeric duplexes may have the following explanation. Duplexes formed by short oligomers are marked by pronounced dependence of CD and T_m on base-pair sequence.[22, 23]
This effect is gradually attenuated with chain elongation and is, for instance, completely unobser-

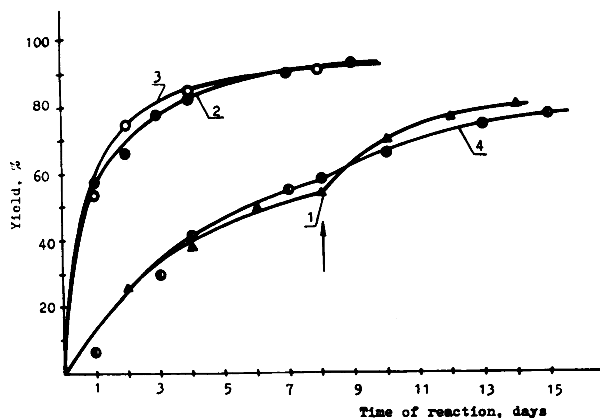


Fig. 8 Kinetics of accumulation of polymerization products. 1, CDI I-induced polymerization of d(TGGCCAAGCTp) (mixture II in Table 1); the arrow denotes the addition of fresh CDI I. 2, CDI II-induced polymerization of d(TGGCCAAGCTp), conditions see in the legend to Fig. 4. 3, CDI II-induced polymerization of d(TGGCCAAGCTTGGCCAAGCTp), conditions see in the legend to Fig. 6. 4, CDI II-induced polymerization of d(pTGGCCAAGCT), conditions see in Materials and Methods.

vable in DNA with quasi-random base distribution. Polymerization of oligonucleotides, in concatemeric formation, on the contrary, enhances this effect owing to multiple repetitions of a certain nucleotide sequence. If the polymer, as in the present case, reiterates a "favourable" sequence, the thermal stability of duplexes formed by such polymers would be substantially higher than that of DNA with the same G · C content but with a quasi-random nucleotide sequence. Another characteristic feature of concatemers is the rapid renaturing of melted duplexes (e.g., no annealing is required for the formation of regular helical structures in concatemeric systems). Together with enhanced thermal stability, the rapid renaturing may become a source of errors and complications in the work with concatemers (e.g. in PAG electrophoretic sequencing).

If one directs attention to the concatemeric structures in native DNAs[24, 25] and considers the biological significance of the mentioned anomalous properties of concatemeric duplexes, it becomes obvious that regions containing recurrent short oligonucleotide sequences must differ significantly in thermal stability and dynamics of denaturation-renaturation. This would provide for pronounced variability in the conformational and kinetic facilities of nucleic acid-protein interactions in such regions, which may be involved in vivo in regulation of genome function.

3. Cleavage of the products of template-guided chemical ligation of d(TGGCCAAGCTp) by restriction endonucleases

As already mentioned, the nucleotide sequence d(TGGCCAAGCTp) is such that the duplex formed upon its self-association and carbodiimide-induced polymerization contains sites recognized by restriction endonucleases Bsu RI, Bal I, Alu I, Hind III and their isoschizomers. Further, owing to the concatemeric nature of the duplex each site is regularly repeated at intervals of 10 nucleotide pairs. Thus, complete cleavage of such polymeric duplex by a restriction endonuclease at all possible recognition sites would yield a decanucleotide as a major product and shorter oligonucleotides corresponding to the terminal fragments of polymeric chains. To confirm the structure and purity of polymers obtained in the present work by chemical ligation, and to elucidate the peculiarities of cleavage of closely positioned (tandem) sites, we have treated a duplex formed by a mixture of polymers d(TGGCCAAGCTp)$_{4-10}$ with restriction endonucleases Alu I, Bsu RI, and Hind III. Reaction mixtures were analysed by chromatography on Aminosilochrom resin.

Fig. 9 depicts the analyses of the final restriction digests of duplexes composed of d(TGGCCAA-GCTp)$_{4-10}$. As expected, the major products of exhaustive digestion are in all cases the corresponding decanucleotides. The structure of decanucleotides d(pCCAAGCTTGG), d(pCTTGGCCAAG), and d(pAGCTTGGCCA) isolated from the Bsu RI, Alu I, and Hind III digests, respectively, was confirmed by terminal nucleotide analysis and nucleotide mapping.[9] A practically complete cleavage of the initial duplex is attained in the chosen conditions (18°C, 24h), with Bsu RI and Alu I (Figs. 9b, 9c), whereas Hind III cleaves only about two thirds of the polymeric duplex to the corresponding decanucleotide (Fig. 9d). This appears to be due not to the substrate specificity of the enzyme or particular structure of the concatemeric duplex, but rather to the mere lability of the

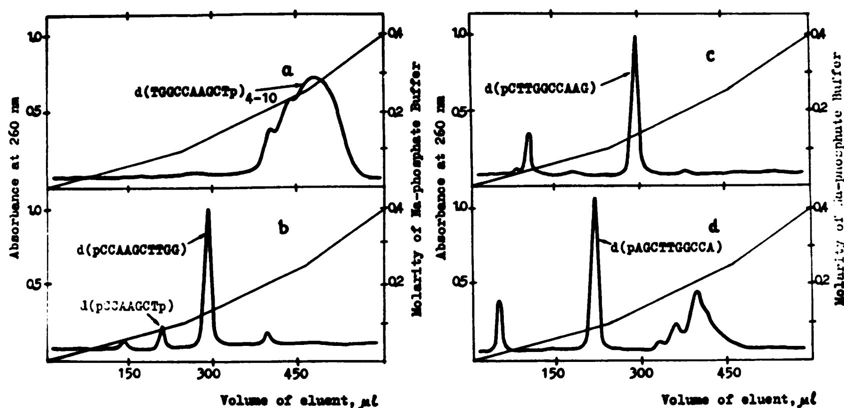


Fig. 9. Chromatography (see Fig. 3 caption) of a mixture of $d(TGGCCAAGCTp)_{4-10}$ before (a) and after cleavage with restriction endonucleases Bsu RI (b), Alu I (c), and Hind III (d).

enzyme preparation which quickly loses Hind III restriction activity in the assay buffer. The suggestion is further supported by the fact that the residual undigested polynucleotides in the Hind III digest are then completely split under the same conditions by a fresh portion of Hind III.

Polymeric duplexes obtained by chemical ligation are completely cleaved by restriction endonucleases, the sites of which cover without gaps the entire nucleotide sequence of the duplexes; this not only unambiguously confirms their primary structure but also testifies to the virtual absence of carbodiimide modification of heterocyclic bases. Furthermore, the quantitative relation of the decanucleotides and shorter oligonucleotide fragments in the total restriction endonuclease digests provides a direct estimate of the mean length of polymers subjected to hydrolysis. Thus, the ratio of the areas of UV absorption peaks corresponding to d(pCCAAGCTp) and d(pCCAAGCTTGG) in the Bsu RI digest (1:10, Fig. 9b) indicates that the initial duplex comprised d(TGGCCAAGCTp) polymers with an average length of 60-80 nucleotides. This estimate is in good correlation with the polynucleotide mean length value obtained by sedimentation analysis.

The complete cleavage of the concatemeric duplex at all possible sites signifies that at least the enzymes we have used (Alu I, Bsu RI, and Hind III) recognize all their tandem sites in extended duplexes, and readily cleave each such site. It is quite probable that the interval of 10 nucleotide pairs created in this concatemeric system is critical or close to critical for normal site cleavage by restriction endonucleases. This is also evidenced, for instance, by experiments with satellite DNA built of repeated heptanucleotide sequences with Eco RI* sites, which could not be split in all tandem sites.[26] On the other hand, satellite DNA comprising recurrent 46-base-pair fragments was practically completely cleaved by Alu I to 46-member oligonucleotides.[24]

In conclusion it should be noted that the chemical ligation technique followed by restrictive cleavage of the resulting polymers is an excellent means of obtaining high-purity linkers and adap-

tors for genetic engineering. In our case this method applied to a single chemically synthesized decanucleotide yielded three different self-complementary decanucleotides carrying sites of restriction enzymes.

Thus, we have tested, optimized, and for the first time applied the chemical template-guided polymerization (chemical ligation) technique to the preparation of DNA-like synthetic polymers carrying certain genetic information. The major advantage of this approach is that it allows polymers with unnatural, modified internucleotide bonds to be obtained. It is noteworthy that both 3′- and 5′-phosphorylated oligonucleotides can be used as chemical ligation substrates though the 3′-phosphates are the more reactive.

The concatemeric duplexes obtained by the methods tested in this work are also important as such. Indeed, synthetic concatemeric duplexes can serve as expedient models of palindromes in DNA or whole DNAs built on a concatemeric principle (e.g. satellite DNAs). The ability of concatemers to form stable long duplexes with controlled number of interruptions ("nicks") in the chains makes them promising substrates for the studies on the fine mechanisms of nucleic acid-protein interactions in such major events as processing, transcription initiation, restriction and methylation of DNA etc. Fundamentally new possibilities in studying the functioning of numerous DNA-binding proteins and enzymes are opened by the use of synthetic concatemeric duplexes with unnatural (phosphoramide, pyrophosphate, or other) internucleotide bonds, as well as of those with chemically active (photoactivated, alkylating, phosphorylating etc.) groups introduced into chain interruptions or heterocyclic basis for affinity modification of proteins within the nucleoprotein complex. Creation of such DNA-like structures is a matter of nearest future.

ACKNOWLEDGEMENTS

The authors wish to thank Prof. H.-J.Fritz (Cologne, FRG) and Dr. A.Yanulaitis (Vilnius, USSR) for the generous gifts of Alu I, Bsu RI, and Hind III; donation of CDI II by Prof. P.Gilham (Indiana, USA) is gratefully acknowledged; thanks are due to Dr. V.Kagramanova (Moscow, USSR) for $d(TGGCCAAGCTp)_2$ sequencing.

REFERENCES

1. Naylor,R. and Gilham,P.T. (1966) Biochemistry 5,2722-2728
2. Shabarova,Z.A. and Prokofiev,M.A. (1970) FEBS Lett. 11, 237-240.
3. Badashkeeva,A.G., Kabasheva,G.N., Knorre,D.G., Shamovsky,G.G. and Shubina,T.N. (1972) Dokl. AN SSSR 206, 870-873.
4. Uesugi,S. and Ts'o,P.O.P. (1974) Biochemistry 13, 3142-3151.
5. Shabarova,Z.A., Dolinnaya,N.G., Turkin,S.I. and Gromova,E.S. (1980) Nucleic Acids Res. 8, 2413-2429.
6. Shabarova,Z.A., Volkov,E.M., Orezkaya,T.S., Turkin,S.I., Dolinnaya,N.G., Kagramanova,V.K. and Prokofiev,M.A. (1981) Dokl. AN SSSR 258, 914-917.
7. Abbreviations: CDI I, N-cyclohexyl-N′-[β-(N-methylmorpholinio)ethyl] carbodiimide p-toluene sulfonate; CDI II, 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride; MES, 2-morpholinoethanesulfonate.

8. Korobko,V.G. and Grachev,S.A. (1977) Bioorgan. khimiya 3, 1420-1422.
9. Sanger,F. (1973) in Virus Research (Fox,C.F. and Robinson,W.S., eds), pp. 573-599, Academic Press, London-New York.
10. Zimmer,C. and Luck,G. (1984) Biochim. et biophys. acta 361, 11-32.
11. Marck,C. and Guschlbauer,W. (1978) Nucleic Acids Res. 5, 2013-2031.
12. Studier,F.W. (1965) J. Mol. Biol. 11, 373-390.
13. Martin,F.H., Uhlenbeck,O.C. and Doty,P. (1971) J. Mol. Biol. 57, 201-215.
14. Badashkeeva,A.G., Gorbunov,N.P., Shamovsky,G.G. and Shubina,T.N. (1974) Izvestiya sibirskogo otdeleniya AN SSSR, No. 2, 96-102.
15. Popov,S.G. and Shamovsky,G.G. (1976) Molecul. biologiya 10, 576-583.
16. Popov,S.G. and Shamovsky,G.G. (1976) Izvestiya sibirskogo otdeleniya AN SSSR, No. 7, 14-19.
17. Dolinnaya,N.G. and Shabarova,Z.A. (1980) Bioorgan. khimiya 6, 209-215.
18. Shabarova,Z.A. (1980) Nucleic Acids Symposium Series 7, 259-279.
19. Augustini-Tocco,G. and Brown,G.L. (1965) Nature 206, 683-685.
20. Mills,D.R. and Kramer,F.R. (1979) Proc. Natl. Acad. Sci. USA 76, 2232-2235.
21. Ambartsumyan,N.S. and Maso,A.M. (1980) FEBS Lett. 114, 265-268.
22. Borer,P.H., Uhlenbeck,O.C., Dengler,B. and Tinoco,I. (1973) J. Mol. Biol. 80, 759-771.
23. Borer,P.H., Dengler,B. and Tinoco,I. (1974) J. Mol. Biol. 86, 843-853.
24. Pöschl,E. and Streeck,R.E. (1980) J. Mol. Biol. 143, 147-153.
25. Sealy,L., Hartley,J, Denelson,J. and Chalkley,R. (1981) J. Mol. Biol. 145, 291-318.
26. Kirsch, D.R. and Cohen,E.H. (1980) Biochem. Biophys. Res. Commun. 96, 1603-1609.