RNA-linked nascent DNA pieces in phage T7-infected Escherichia coli. II. Primary structure of the RNA portion

Tetsunori Seki† and Tuneko Okazaki*

Institute of Molecular Biology, Faculty of Science, Nagoya University, Nagoya 464, Japan

Received 10 August 1979

ABSTRACT Short DNA chains were purified from phage T7 infected E. coli cells and 5' ends were labeled with ^{32}P. By an alkali-treatment, pNp's rich in pAp and pCp were liberated from the T7 short DNA chains. After digestion of the [5'-^{32}P]short DNA with the 3' to 5' exonuclease of T4 DNA polymerase, [5'-^{32}P]mono- to pentaribonucleotides tipped with a deoxyribonucleotide residue at their 3' ends were isolated. 5' terminal ribonucleotides were; exclusively AMP in the penta- and the tetraribonucleotides, mostly CMP in the triribonucleotide and mainly CMP and AMP in di- and monoribonucleotides. The 5' terminal dinucleotide of the penta- and the tetraribonucleotides was pApC. The nucleotide sequence of the tetraribonucleotide was mainly pApCpCpN and some pApCpApN, where N was mainly A and C. These results indicate that oligoribonucleotides shorter than trinucleotide may result from in vivo degradation of the tetra- and pentaribonucleotides. A possibility that the tetra- and pentaribonucleotides with a 5' triphosphate terminus are the intact primers for the discontinuous T7 DNA replication is discussed.

INTRODUCTION

DNA replication of T7 phage proceeds discontinuously via short chain intermediates (1, 2, 3). Detection of 5'-hydroxyl terminated DNA chains after alkali or RNase digestion of the T7 nascent short chains strongly suggested a covalent linkage of RNA to the 5' termini (3). The RNA-linked nascent T7 DNA accumulates in the absence of a functional T7 gene 6 exonuclease (3). The accumulation is particularly pronounced when both gene 6 exonuclease and host DNA polymerase I are deficient. These results support the proposition that discontinuous replication of T7 phage DNA is primed by RNA and that T7 gene 6 exonuclease and E. coli DNA polymerase I remove primer RNA. In support of this, both gene 6 exonuclease and DNA polymerase I degrade the RNA chains of RNA·DNA hybrid molecules in vitro (4, 5).

In the present study, we substantiate the T7 primer RNA by isolating oligoribonucleotides from the 5' ends of T7 short DNA chains. The short chains accumulated in T7 ts136 infected E. coli C-N3 cells lacking the two processing exonuclease provides a rich source of primer RNA(3). The 5' ends

© Information Retrieval Limited 1 Falconberg Court London W1V 5FG England

of the short chains are labeled with ^{32}P and the DNA portion removed by treatment with the 3' to 5' exonuclease of T4 DNA polymerase (6, 7, 8). [5'-^{32}P]oligoribonucleotides of chain lengths up to pentanucleotide are isolated. The 5' terminal dinucleotide of the tetra- and pentaribonucleotide is pApC and nucleotide sequence of the tetraribonucleotide is pApCpCpN or pApCpApN, where N is mainly A and C. Oligoribonucleotides shorter than trinucleotide may be intermediates of *in vivo* degradation of the tetra- and the pentaribonucleotides as judged by the base composition of the 5' terminal ribonucleotides of these chains. The possibility is discussed that intact primers for the discontinuous T7 DNA replication is the tetra- and the pentaribonucleotides with a 5' triphosphate terminus, pppApCp$^{C}_{A}$pN' and pppApC$_{p}^{C}_{A}$pN'pN", respectively.

MATERIALS AND METHODS

Bacterium, bacteriophage and culture medium

E. coli C-N3 [*his*, *polA*480 (formerly designated *polA*ex1)] and bacteriophage T7 *ts*136 (ts mutant of gene 6) have been described (3). Cells were grown in M9 medium (9) supplemented with 0.5% Casamino acids.

Radioactive compounds and enzymes

[Methyl-^{3}H]thymidine (50.8 Ci/mmol) was obtained from New England Nuclear Corp. [γ-^{32}P]ATP (118 Ci/mmol) was prepared by the procedure of Glynn and Chappell (10). T4 polynucleotide kinase, *E. coli* alkaline phosphatase, T4 DNA polymerase, pancreatic DNase I and snake venom phosphodiesterase have been described (6). Nuclease SW was purchased from Seikagaku Kogyo, Co. (26).

Isolation and terminal labeling of short DNA chains

E. coli C-N3 was grown with shaking at 30°C to 6 x 10^{8} cells/ml in 500 ml of M9 medium supplemented with 0.5% Casamino acids and infected with T7 *ts*136 phage at a multiplicity of 10. After 19 min at 30°C, the culture medium was shifted to 43°C and 2 min later pulse-labeled with 0.2 µM [^{3}H]thymidine (50.8 Ci/mmol) for 12 sec. The labeling was terminated with an equal volume of an ethanol/phenol mixture (11). Nascent short chains were purified essentially as described by Ogawa *et al.* (6). Phosphorylation of the 5' termini of the short chains with T4 polynucleotide kinase and [γ-^{32}P]ATP (118 Ci/mmol) were carried out using the nitrocellulose fraction of the nascent short chains (6). Resulting 5'-^{32}P labeled short chains were further purified as described previously (6).

Enzyme digestion of nucleic acids

Digestion of [5'-^{32}P]short DNA chains with the 3' to 5' exonuclease of T4 DNA polymerase was carried out in a 4.0-ml reaction mixture containing 67 mM Tris·HCl (pH 8.3), 6.7 mM $MgCl_2$, 11.4 mM 2-mercaptoethanol, 440 µg of E. coli tRNA, 40 pmol [5'-^{32}P]short chains and 24 units of the enzyme at 37°C for 2 hr (8). Ninety-three percent of ^{32}P radioactivity was rendered acid-soluble. Partial digestion of [5'-^{32}P]oligonucleotides with snake venom phosphodiesterase was carried out in 20-µl reaction mixture containing 21 mM TEAB** (pH 8.4), 1 mM $MgCl_2$, 7 to 10 µg of oligoribonucleotides of the corresponding chain length and 0.4 µg of the enzyme for 10 to 80 min at 37°C. [5'-^{32}P]oligonucleotides were digested completely with nuclease SW in a 100-µl reaction mixture containing 50 mM sodium carbonate buffer (pH 10.5), 1 mM $Mg(CH_3COO)_2$, 0.1 M NaCl, 22 µg of E. coli tRNA and 2 units of the enzyme for 2 hr at 37°C.

Identification of pNp liberated from short chains by alkaline hydrolysis

[5'-^{32}P]short chains or [5'-^{32}P]oligonucleotides were treated with 0.3 N NaOH for 18 hr at 37°C. The trinucleotide fraction was further digested with snake venom phosphodiesterase (6). When required, acid-insoluble material was removed by precipitation with 5% cold CCl_3COOH in the presence of sonicated salmon sperm DNA and resulting supernatant was used after neutralization with NaOH. Otherwise, the hydrolysates were neutralized with CCl_3COOH and applied to columns of DEAE-Sephadex A-25 (bed volume of 0.5 ml) with pNp otpical density markers. Elution was by a 100-ml linear gradient of 0 to 0.35 M NaCl in 5 mM Tris·HCl (pH 8.0)-7 M urea. Fractions containing ^{32}P and an optical density marker were pooled and desalted (6) and then chromatographed on PEI-cellulose** plates (Polygram Cel 300 PEI, Macherey-Nagel, Germany) either in 0.85 M potassium phosphate (pH 3.4) (12) or 2 M sodium formate (pH 3.0) (13).

Detection of cis-diol structure at the 3' termini of the [5'-^{32}P]oligonucleotide

Periodate oxidation and aniline treatments were carried out as described by Steinschneider & Fraenkel-Conrat (14). Oligonucleotides were analyzed before and after the treatments as described previously (6). Chromatography of [5'-^{32}P]oligonucleotides on a borate gel column** was carried out with oligoribonucleotide optical density markers and [^{3}H]oligodeoxynucleotide markers. The gel traps compounds having a cis-diol structure. Sample solution (300 µl) made 1 M in TEAB (pH 8.5), was applied to a column of borate gel (bed volume of 4 ml) equilibrated with 1 M TEAB (pH 8.5). The column was washed with one

bed volume of 1 M TEAB (pH 8.5) and then nucleotides retained to the gel were eluted with redistilled water. Borate gel was either a gift from Dr. K. Ueda (Kyoto University) or prepared according to Okayama et al. (15).

Other materials and methods

DNA·DNA hybridization was carried out as described (16). Two-dimensional chromatography of oligonucleotides was performed first by electrophoresis on a cellulose acetate strip (Fuji Photo Film Co.) in 7 M urea, 6% acetic acid, 1 mM EDTA, adjusted to pH 3.5 with pyridine and then by thin layer homochromatography on a 20 x 40 cm DEAE-cellulose plate (Polygram Cell 300 DEAE, Macherey-Nagel, Germany) at 65°C using "homomixture VI", pH 6.5 (17). Autoradiography was carried out using a prefogged Kodak RP Royal X-Omat film and intensifying screen (Kyokko HS) at -70°C for 2 weeks. Optical density markers of p(rN)p were prepared by alkaline hydrolysis of nuclease SW digest of E. coli tRNA. E. coli tRNA was supplied by Dr. T. Sekiya (National Cancer Research Institute, Tokyo). 5' phosphoryl-terminated oligoribonucleotides were prepared by partial digestion of E. coli tRNA with nuclease SW. [^{3}H] oligodeoxyribonucleotides were prepared by partial digestion of ^{3}H-labeled E. coli DNA with pancreatic DNase. When oligonucleotides of defined chain length were needed, the digest was fractionated by chromatography on a DEAE-Sephadex A-25 column in the presence of 7 M urea. Optical density markers of pApG, pApC and pApU were prepared by phosphorylation of the respective dinucleoside monophosphates (Sigma Co.) with T4 polynucleotide kinase and ATP. pApA was prepared by digestion of poly A (Sigma Co.) with nuclease SW followed by size fractionation on DEAE-Sephadex A-25 column. ^{32}P radioactivity in a solution was measured by Čerenkov counting (18) and the radioactivity on the PEI-cellulose thin layer plate was measured in a toluene-based scintillator fluid in a liquid scintillation counter.

RESULTS

Detection of T7 short DNA chains with ribonucleotides at the 5' ends

Phage T7 ts136 (temperature sensitive gene 6 mutant) infected E. coli polA480 cells were pulse labeled with [^{3}H]thymidine after a brief exposure to 43°C. Nucleic acids were extracted and short DNA chains were purified (6). The 5' ends of the short chains were, after dephosphorylation with bacterial alkaline phosphatase, labeled with ^{32}P using [γ-^{32}P]ATP and T4 polynucleotide kinase (6). 5'-^{32}P-labeled short chains were further purified by passing through a Sephadex G-100 column and by two successive Cs_2SO_4 equilibrium

density gradient centrifugation steps (6). The purified 5'-^{32}P-labeled short chains, which displayed in the gradient a single ^{32}P radioactive peak superimposable on the ^{3}H peak at the buoyant density of DNA (1.42 to 1.47 g/cm^3), were pooled for further analyses.

When the 5'-^{32}P-labeled short chains were treated with 0.3 N NaOH for 18 hr at 37°C, 32% of the ^{32}P radioactivity was hydrolyzed and produced pNp (96%) and unidentified material (4%). Number of ^{32}pNp released was roughly 100 per infected cell. The nucleotide composition of pNp was; pAp (44%), pCp (44%), pUp (6%), pGp (6%). The origin of the 5'-^{32}P-labeled short chains was analyzed by hybridization to T7 and E. coli DNA immobilized on membrane filters. Sixty-eight percent of the ^{32}P radioactivity which annealed to T7 DNA was alkali-labile, whereas almost all radioactivity which annealed to E. coli DNA was alkali-resistant (Table 1). ^{32}P-labeled T7 short chains were eluted from a membrane filter with 0.3 N KOH and incubated for 18 hr at 37°C. Resulting alkali-labile and acid-soluble ^{32}P radioactivities were identified as pNp's composed mainly of pAp and pCp (data not shown). The relative amounts of short chains are: T7, alkali-labile (30%); T7, alkali-resistant (14%), E. coli, alkali-labile (2%); and E. coli, alkali-resistant (54%). The alkali-labile fraction of ^{32}P-labeled T7 short chains is in good agreement with the proportion of T7 RNA-linked nascent DNA molecules measured by the spleen exonuclease assay (3). This result strongly suggests that the alkali-labile termini are derived from the nascent T7 DNA.

Table 1. Hybridization of [5'-^{32}P]short chains with T7 and E. coli DNA

DNA on membrane	[5'-^{32}P]label input	[5'-^{32}P]label hybridized	
		Alkali-labile	Alkali-resistant
	(cts/min)	(cts/min)	(cts/min)
T7 DNA	1809	196	91
E. coli DNA	1833	15	408

[5'-^{32}P]short chains, prepared from T7 ts136 infected E. coli C-N3 cells, were hybridized with 2 µg of T7 DNA or 20 µg of E. coli DNA immobilized on a membrane filter. Chains hybridized were eluted with 1.5 ml of 0.3 N KOH and incubated for 18 hr at 37°C. After addition of 50 µl of 2 mg/ml salmon sperm DNA and 110 µl of 11.6 N $HClO_4$, hydrolysates were centrifuged and the ^{32}P radioactivity in the supernatant (alkali-labile ^{32}P) and in the precipitate (alkali-resistant ^{32}P) were determined.

Isolation of oligoribonucleotide primer

DNA portion of [5'-^{32}P]short chains was extensively digested with the 3' to 5' exonuclease of T4 DNA polymerase under conditions in which most of the [5'-^{32}P]RNA-linked DNA molecules were degraded to [5'-^{32}P]RNA-deoxymononucleotides whereas [5'-^{32}P] DNA molecules were degraded to [5'-^{32}P]deoxydinucleotides (7, 8). The digest was chromatographed on a DEAE-Sephadex A-25 column in the presence of 7 M urea. As shown in Fig. 1, ^{32}P-labeled material eluted in six peaks at the positions corresponding to mono- to hexanucleotides. The bracketed fractions were pooled and their 5' and 3' nucleotides analyzed.

For the 5'-end analyses, a portion of each oligonucleotide was hydrolyzed in 0.3 N NaOH for 18 hr at 37°C and the hydrolysate fractionated by DEAE-Sephadex A-25 chromatography in the presence of 7 M urea. After the alkaline treatment, the trinucleotide fraction was further digested with snake venom phosphodiesterase to degrade possible contaminants, trideoxynucleotides, to mononucleotides and separate from pNp in subsequent DEAE-Sephadex A-25 chromatography. Almost all the ^{32}P in tri-, tetra-, penta- and hexa-

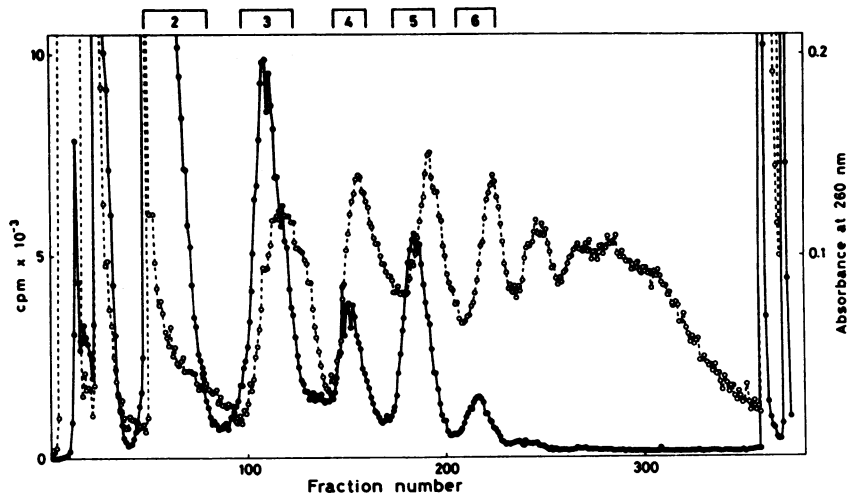

Fig. 1. DEAE-Sephadex A-25 column chromatography of [5'-^{32}P]short chains after digestion with the 3' to 5' exonuclease of T4 DNA polymerase. [5'-^{32}P] short chains were prepared and digested with the 3' to 5' exonuclease of T4 DNA polymerase as described in Materials and Methods. The digest was charged on a DEAE-Sephadex A-25 column (0.2 x 72 cm) and eluted with a linear gradient (400 ml) of 0.05-0.35 M NaCl in 5 mM Tris·HCl (pH 8.0)-7 M urea. A partial digest of E. coli tRNA with nuclease SW was co-chromatographed as size markers. ●, ^{32}P radioactivity; o, absorbance at 260 nm. Chain length of the each fraction is shown on the top of the figure.

nucleotides and 15% of that in dinucleotide eluted at the pNp position, showing that their 5' termini are ribonucleotides (data not shown). The total ^{32}pNp recovered from these five fractions was roughly equal to the ^{32}pNp released by alkali from [5'-^{32}P]short chains before digestion with T4 DNA polymerase. The distribution of alkali-labile termini among the various chain-length oligonucleotides is shown in Table 2. The nucleotide composition of [5'-^{32}P]pNp produced from each oligonucleotide fraction was determined using PEI-cellulose thin layer chromatography in two different solvent systems (Table 2). Only pAp was produced from the 5' termini of hexa- and pentanucleotides, and mainly pAp and pCp were produced from the shorter chains.

Nucleotides at the 3' ends of these oligonucleotides were analyzed by periodate and aniline treatments (14) and/or by borate gel column chromatography (15). Both of these methods detect the 2',3' cis-diol termini chracteristic of RNA. After periodate and aniline treatments of tri- to hexanucleotides, the net charge of the oligoribonucleotide internal references

Table 2. Structure and distribution of oligonucleotides with alkali-labile 5' termini produced by the digestion of [5'-^{32}P]short chains with the 3' to 5' exonuclease of T4 DNA polymerase

Chain length	Main oligonucleotide	5'-ribonucleotide termini				
		Distribution	Nucleotide composition			
			pAp	pCp	pUp	pGp
		%	%	%	%	%
2	prNpdN	32	34	47	17	2
3	$(prN)_2pdN$	32	20	69	7	4
4	$(prN)_3pdN$ + $(prN)_4$	12	57	34	8	< 1
	[$(prN)_3pdN$	7	15	76	9	< 1]
	[$(prN)_4$	5	98	2	< 1	< 1]
5	$(prN)_4pdN$	20	95	2	2	< 1
6	$(prN)_5pdN$	4	92	3	5	1
		Sum 100				

Di- to hexanucleotide fraction in Fig. 1 were pooled, desalted and the 5' and 3' terminal structures were analysed as described in the text. 5' terminal nucleotides of the two tetranucleotide fractions obtained by the periodate and aniline treatments were also analyzed [data shown in the brackets]. Alkaline hydrolysis of the oligonucleotides and the identification of [5'-^{32}P]pNp's were described in Materials and Methods.

was changed, whereas the profiles of ^{32}P radioactivity were unchanged except that of tetranucleotide, where 42% of the radioactivity moved with the marker. Therefore, all the tri-, penta- and hexanucleotides and 58% of tetranucleotide have deoxynucleotide(s) at their 3' termini. All [^{32}P]dinucleotides passed through a borate gel column, whereas the diribonucleotide marker added to the sample was retained (data not shown), indicating that dinucleotides consisted of prNpdN (15%) and $(pdN)_2$ (85%). Neither hexa- nor pentanucleotide was retained by the borate gel column, confirming the results obtained by the periodate and aniline treatments. The 5' terminal nucleotides were different in the two tetranucleotide fractions generated by the periodate and aniline treatments (Fig. 2). From the fraction which was sensitive to the treatments (molecules with ribonucleotide at 3' end) pAp was exclusively obtained, whereas from the resistant fraction (with deoxyribonucleotide at 3'

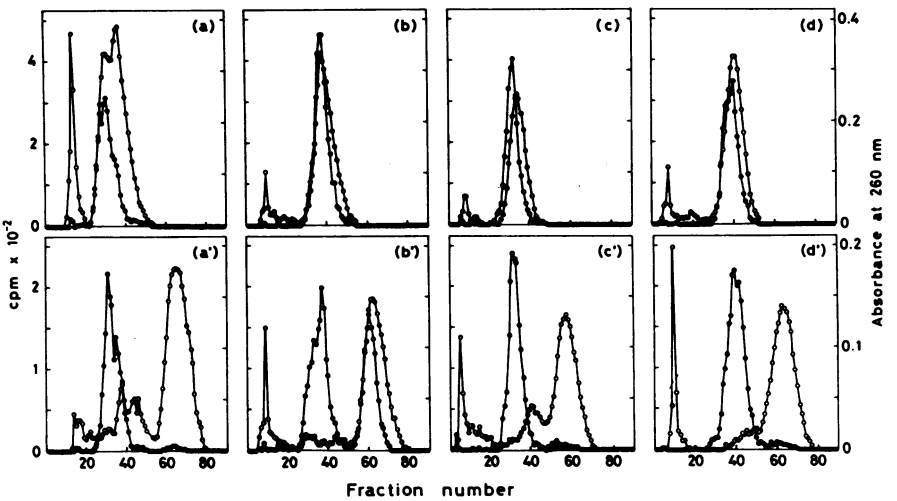

Fig. 2. DEAE-Sephadex A-25 column chromatography of the [5'-^{32}P]oligonucleotides before and after treatments with periodate and aniline. Tri- to hexanucleotide fractions in Fig. 1 were pooled and desalted. Portions were rechromatographed on a 1-ml DEAE-Sephadex A-25 column with the oligonucleotide markers of the corresponding chain length ((a) tri-; (b) tetra-; (c) penta-; (d) hexanucleotide). Fractions containing ^{32}P radioactivity and UV-absorbing material were pooled, desalted and treated with periodate and aniline. After the treatments, the samples were chromatographed on a 1-ml DEAE-Sephadex A-25 ((a') tri-; (b') tetra-; (c') penta-; (d') hexanucleotides). Elution was carried out with a linear gradient (100 ml) of 0.12-0.22 M NaCl ((a), (a')), 0.15-0.25 M NaCl ((b), (b')), 0.18-0.28 M NaCl ((c), (c')) or 0.20-0.30 M NaCl ((d), (d')) in 5 mM Tris·HCl (pH 8.0)-7 M urea. •, ^{32}P radioactivity; o, absorbance at 260 nm.

end) pCp was mainly obtained (Table 2).

The following analyses show that only one deoxynucleotide residue is at the 3' termini of the oligonucleotides. The pentanucleotide was partially digested with snake venom phosphodiesterase and the digest was fractionated by chromatography on a borate gel column. The ^{32}P radioactivity, that had not been retained by the gel before the venom digestion, separated into two fractions (Fig. 3A). Fraction I, which was not retained by the gel, was eluted with ^{3}H-labeled oligodeoxyribonucleotide marker, and fraction II which was retained by the gel, eluted with the oligoribonucleotide optical density marker. Each fraction was then chromatographed on a DEAE-Sephadex A-25 column in the presence of 7 M urea (Fig. 3B). From fraction I, one major ^{32}P radioactive peak was recovered at the position of a pentanucleotide optical

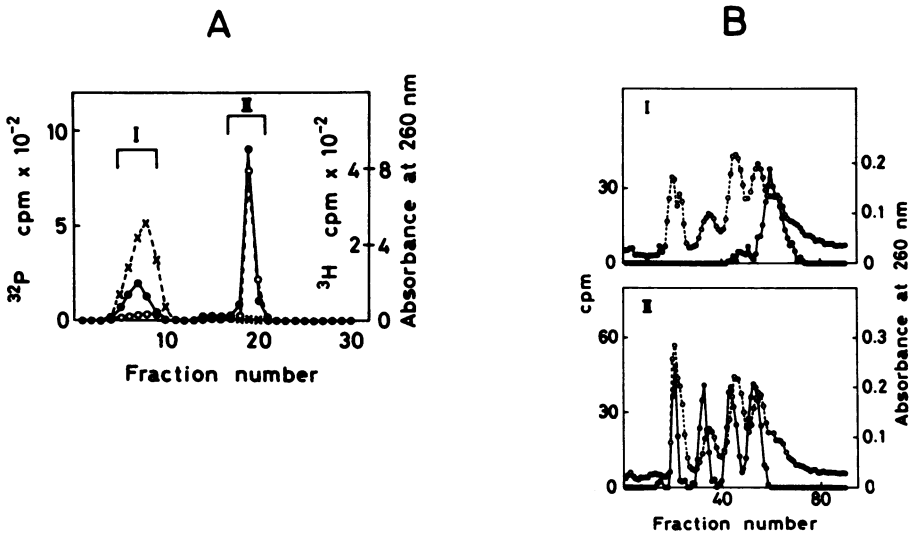

Fig. 3. Chromatographic separation of partial digestion products of the [5'-^{32}P]pentanucleotide with snake venom phosphodiesterase. A: The [5'-^{32}P] pentanucleotide fraction was partially digested with snake venom phosphodiesterase. The digest, after neutralization and desalting, was chromatographed on a 4-ml borate-gel column together with an oligoribonucleotide optical density marker and [^{3}H]pentadeoxynucleotide. Elution was carried out as described in Materials and Methods. •, ^{32}P radioactivity determined by Čerenkov radiation; x, ^{3}H radioactivity with 1/80 volume of each fraction in toluene-base scintillater; o, absorbance at 260 nm. B: Fractions I and II in Figure A were pooled separately and chromatographed on a DEAE-Sephadex A-25 column (bed volume of 0.5 ml) together with oligoribonucleotide optical density markers. Elution was carried out with a linear gradient (100 ml) of 0.00-0.35 M NaCl in 5 mM Tris·HCl (pH 8.0)-7 M urea. •, ^{32}P radioactivity by Čerenkov radiation; o, absorbance at 260 nm.

density marker, whereas from fraction II, four peaks of radioactivity were recovered at the positions of mono-, di-, tri- and tetranucleotides. Therefore the structure of the pentanucleotide must be $(prN)_4pdN$. Similar results on the hexanucleotide were shown in Fig. 4. Again, only one deoxyribonucleotide residue was found at its 3' end. Therefore, the hexanucleotide is $(prN)_5pdN$. The tri- and tetranucleotide with deoxyribonucleotide at 3' ends were also shown to be $(prN)_2pdN$ and $(prN)_3pdN$, respectively (data not shown).

These results prove that oligoribonucleotides of chain length of up to 5 nucleotides are covalently attached to the 5' ends of T7 short DNA chains and their 5' terminal ribonucleotides have a characteristic base composition.

Identification of 5' terminal dinucleotide of the penta-, hexa- and tetranucleotides

5' terminal dinucleotides of the penta- and the hexanucleotides ($(prN)_4pdN$ and $(prN)_5pdN$, respectively) were analyzed. The pentanucleotide was completely digested with nuclease SW and products were fractionated by DEAE-Sephadex A-25 chromatography in the presence of 7 M urea. Seventy-four percent and 18% of ^{32}P radioactivity were recovered as dinucleotides and trinucleotides, respectively. The dinucleotide fraction was then chromato-

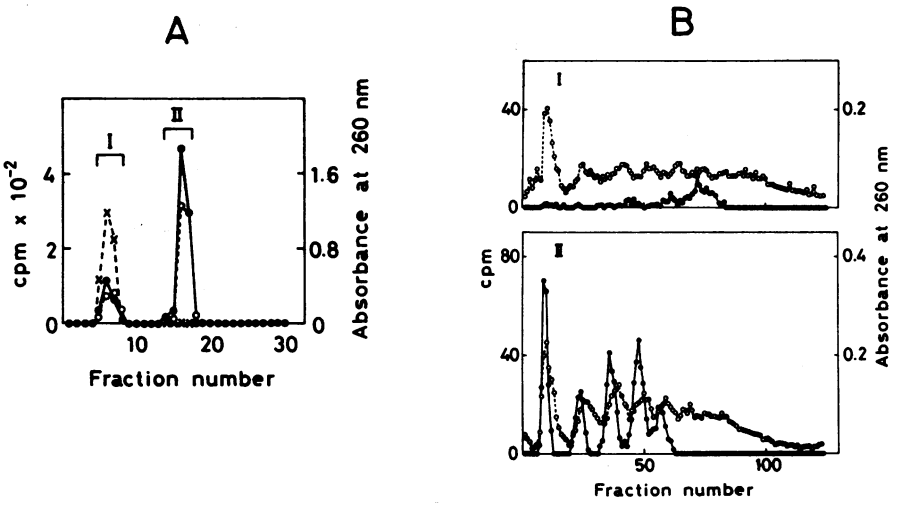

Fig. 4. Chromatographic separation of partial digestion products of the [5'-^{32}P]hexanucleotide with snake venom phosphodiesterase. The [5'-^{32}P]hexanucleotide fraction was partially digested with snake venom phosphodiesterase and the digest was fractionated by a borate-gel column and then by a DEAE-Sephadex A-25 column as mentioned in Figures 3A and B.

graphed on a Bio-Rad AG 1 x 2 column to resolve pApA and pApC from pApG and pApU. All the radioactivity was recovered in the elution position of pApA and pApC (Fig. 5A). The radioactive materials were pooled and further analyzed by a PEI-cellulose thin layer plate (Fig. 6A). The dinucleotide at the 5' termini of the pentanucleotide was pApC. For the analyses of the dinucleotide at the 5' termini of the hexanucleotide, a partial digestion with snake venom phosphodiesterase was used. The digest was chromatographed on DEAE-Sephadex A-25 in the presence of 7 M urea and mono- (68%), di- (18%) and trinucleotide (14%) were obtained. The dinucleotide fraction was further analyzed as described above and identified as pApC (Fig. 5B, 6B). The dinucleotide at the 5' termini of tetraribonucleotide ($(prN)_4$) (Table 2) was shown to be pApC (93%) and pApA (7%).

Nucleotide sequence of the pentanucleotide

[5'-^{32}P]pentanucleotide $(prN)_4pdN$ was partially digested with snake venom phosphodiesterase and the digest was analyzed by two dimensional chromatography. As shown in Fig. 7, an autoradiograph of the chromatography showed a single spot for the mononucleotide and for the dinucleotide, which corresponded to ^{32}pA and ^{32}pApC, respectively. Multiple spots were shown up for tri- and tetranucleotides, indicating that heterogeneity exists in inner sequence. As judged by the mobility, the major trinucleotide spot is pApCpC

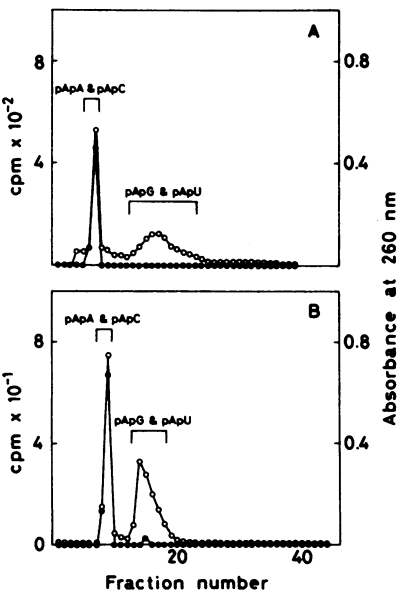

Fig. 5. Bio-Rad AG 1 x 2 column chromatography of the dinucleotide. The [5'-^{32}P] pentanucleotide was completely digested with nuclease SW. The [5'-^{32}P]hexanucleotide was partially digested with snake venom phosphodiesterase. Each digest was chromatographed on a DEAE-Sephadex A-25 column in the presence of 7 M urea and the dinucleotide fractions were obtained. Each dinucleotide fraction was then applied to a 0.3-ml Bio-Rad AG 1 x 2 column with pAprN optical density markers and eluted with a linear gradient (50 ml) of 0-0.05 M (A) or 0-0.10 M (B) NaCl in 0.01 N HCl. A, dinucleotide derived from the pentanucleotide; B, dinucleotide derived from the hexanucleotide; •, ^{32}P radioactivity by Čerenkov radiation; o, absorbance at 260 nm.

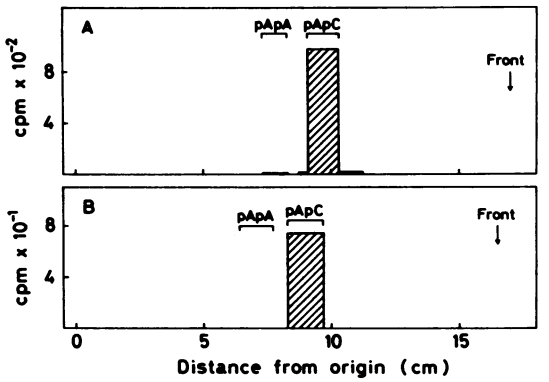

Fig. 6. Identification of the 5' nucleotides of the [5'-^{32}P]hexa- and the [5'-^{32}P]pentanucleotides by chromatography on PEI-cellulose thin layer plates. Radioactive fractions of Fig. 5 were neutralyzed and desalted. Concentrated samples were applied to a PEI-cellulose thin layer plate which was developed in 0.85 M potassium phosphate (pH 3.4). After drying, the plates were cut into pieces and the radioactivity determined in a toluene-base scintillation fluid. A, dinucleotide from the pentanucleotide; B, dinucleotide from the hexanucleotide.

and minor one is pApCpA. Two major spots of tetranucleotide may be pApCpCpC and pApCpCpA and two minor ones may be pApCpCpU and pApCpCpG. Multiple pentanucleotide spots were observed on the autoradiograph. Due to the sequence heterogeneity of the tri- and tetranucleotides, it is difficult to accurately deduce the nucleotide sequence of each pentanucleotide spot.

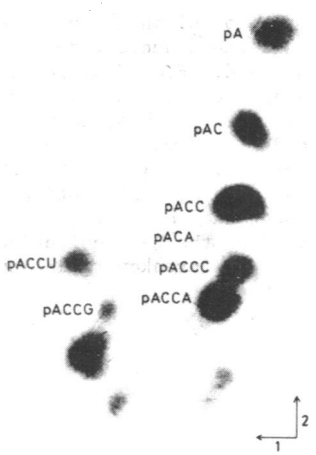

Fig. 7. Autoradiography of the two-dimensional chromatography of partial digestion products of the [5'-^{32}P]pentanucleotide with snake venom phosphodiesterase. The [5'-^{32}P]pentanucleotide (2600 cts/min) was partially digested with snake venom phosphodiesterase. The products were analyzed first by electrophoresis on a cellulose acetate strip and then by homochromatography on a DEAE-cellulose thin layer plate.

Deoxyribonucleotides at the RNA-DNA junction have been analyzed in an independent experiment using the method of specific radioactive labeling of the 5'-OH termini of DNA produced by an alkali-digestion of short DNA chains (19). All four deoxymononucleotides with enrichmant for dTMP (40%) have been found (A. Fujiyama, Y. Kohara and T. Okazaki; manuscript in preparation).

DISCUSSION

We have shown previously that primer RNA is linked to the 5' end of T7 nascent short DNA and that the T7 gene 6 exonuclease and E. coli DNA polymerase I remove the primer before the ligation of the short DNA chains (3). In the present study, the primer RNA has been isolated from short DNA chains which accumulate in T7 ts136 infected E. coli C-N3 cells at 43°C (3) and its primary structure has been elucidated. To detect the primer RNA, we labeled the 5' termini of the short chains with ^{32}P after a removal of the pre-existing phosphate groups with bacterial alkaline phosphatase (6). The proportion of the [5'-^{32}P]T7 short DNA chains which liberate [5'-^{32}P]pNp after alkali-digestion (Table 1) agrees well with the proportion of the nascent short DNA chains which have a 5'-hydroxyl end after alkali-treatment in the previous study (3). Therefore, most of the [5'-^{32}P]T7 short chains, including those with the alkali-labile termini, are probably newly synthesized chains as expected from their role in DNA synthesis. The [5'-^{32}P]pNp's liberated from the short chains are very rich in pAp and pCp, showing that the T7 primer RNA has characteristic sequences. The primer RNA was isolated from the terminally labeled short chains by digesting away the DNA moiety with the 3' to 5' exonuclease of T4 DNA polymerase (7, 8). Oligoribonucleotides in chain length up to pentanucleotide were obtained and most of them had one deoxyribonucleotide residue remaining at the 3' end (Fig. 1 & Table 2), as expected from the specificity of the exonuclease (7, 8). Analysis of the 5' terminal nucleotides of each chain length oligonucleotide has shown up the characteristic primary structure of the primer RNA. 5' terminal nucleotides of both $(prN)_5pdN$ and $(prN)_4pdN$ were exclusively AMP and those of $(prN)_3pdN$ were mostly CMP but some AMP (Table 2), whereas almost all dinucleotides at the 5' ends of both $(prN)_5pdN$ and $(prN)_4pdN$ were pApC (Figs. 5 & 6) and only a trace of pApA was found in $(prN)_4$. The two-dimensional chromatography of the products of partial digestion of $(prN)_4pdN$ with snake venom phosphodiesterase showed that the primary structure is predominantly pApCpCprNpdN and some pApCpAprNpdN, where rN is A and C rich in this order (Fig. 7). Judged

by the 5' terminal AMP residues, most tetraribonucleotides [$(prN)_4pdN$ and $(prN)_4$] are not derived by a degradation of the pentaribonucleotide [$(prN)_5pdN$]. Whereas nucleotide compositions of 5' ends of $(prN)_3pdN$ (CMP terminated portion), $(prN)_2pdN$ and prNpdN can be explained by a stepwise degradation of the tetraribonucleotides from the 5' end (Table 2). Our interpretation to these results is that the chain length of the intact RNA primers would mostly be tetra- and pentaribonucleotide and their sequence would be $pppApCp^C_ApN'$ and $pppApCp^C_ApN'pN''$, respectively, where N' is composed mainly of A and C and trace of U and G. The pentaribonucleotide primer might be made less frequently than the tetraribonucleotide primer, as the former was found in a small amount (Fig. 1 & Table 2) and its putative degradation product was not found in the tetraribonucleotides either, despite the previously shown fact that a tetraribonucleotide could be produced from the pentaribonucleotide primer by a function of T7 gene 6 exonuclease (4). Presence of a small amount of 5' AMP terminated $(prN)_3pdN$ (Table 2) might suggest either a presence of intact triribonucleotide primer or a degradation of the pentaribonucleotide by quick removal of its 5' dinucleotide, possibly with DNA polymerase I (4, 5). If the A terminated triribonucleotide would be originated in the latter pathway, then the first step of degradation of the pentaribonucleotide primer might preferentially be achieved by E. coli DNA polymerase I in vivo. The tetraribonucleotides [$(prN)_4$] may be derived from the DNA molecules with the tetraribonucleotide primer (or less likely from the pentaribonucleotide primer) as a result of complete digestion of DNA moiety with the exonuclease of T4 DNA polymerase, because pApC is the main dinucleotide sequence of their 5' ends. A possibility that they represent a contaminating free RNA seems less likely, because chain length of RNA would be longer than four nucleotides to maintain a non-covalent interaction with DNA after extensive purification steps and a long RNA, if ever, would not be degraded to chain length four by T4 DNA polymerase under the present conditions (6, 8).

Detection of the intact primer RNA with a triphosphate group at the 5' end is not possible by the methods used in this report but can be made by in vivo labeling of short chains with $H_3{}^{32}PO_4$. The accompaning paper by Ogawa and Okazaki (20) shows that the penta- and tetraribonucleotides are indeed intact primers synthesized with triphosphate termini. Therefore, the representative T7 primer RNA is tetra- or pentanucleotides with a unique pppApC sequence at the 5' end but with an A and C rich diverse sequence in the subsequent 3' portion. The diversity of primer sequence suggests that primers are synthesized in situ rather than being mobile initiators of DNA

synthesis. We are currently analyzing the distribution of transition loci from primer RNA to DNA synthesis on T7 genome and the base sequences around these loci. On a 339-nucleotide H-strand segment of T7 genome (a Hpa II fragment), five loci for the transition have been found (7; A. Fujiyama, Y. Kohara and T. Okazaki, manuscript in preparation). Thus potential sites for initiation of discontinuous replication seem to be distributing on T7 genome quite frequently. A short stretch of sequence which contains the complementary sequence to the T7 primer RNA has been found in common at the putative primer initiation sites on the template L strand, supporting the mechanism of *in situ* synthesis of the primer RNA (A. Fujiyama, Y. Kohara and T. Okazaki; manuscript in preparation). These data also suggest that the chain length of nascent DNA fragments, 1000 to 2000 nucleotides, is not simply determined by a frequency of the appearance of the potential primer initiation sites on the genome but determined by a probability by which one site is selected from multiple potential initiation sites along with the replication fork movement.

In vitro studies of T7 DNA replication by Scherzinger *et al.* and Richardson *et al.* showed that tetraribonucleotides, pppApCpCpA or pppApCpCpC, are synthesized by T7 gene 4 protein, an essential protein to T7 DNA synthesis, in the presence of a single stranded DNA template and that the oligoribonucleotides are extended with deoxyribonucleotides by T7 DNA polymerase (21, 22, 23). Unlike *dnaG* protein of *E. coli*, the *E. coli* primase, dNTP's are not utilized as substrate for polymerization by gene 4 protein, though they are hydrolyzed by the NTPase activity intrinsic of gene 4 protein (22, 23). Neither has the insertion of deoxynucleotide into the T7 primer been detected *in vivo* (Fig. 3 & 4). The excellent agreement between the *in vivo* and *in vitro* data persuasively argues that the gene 4 protein is the primase *in vivo* for T7 nascent short DNA synthesis.

The T4 primer RNA is mostly a pentanucleotide and the dinucleotide at the 5' ends is also pApC (7, 8). *In vitro*, similar oligonucleotides are synthesized with T4 gene 41 protein and X protein (24). The resemblance of the structure of the T7 primer RNA to T4 suggest some common mechanism of priming by the coliphages. From the 5' ends of *E. coli* short DNA, mono- to triribonucleotides were isolated but no information is yet available about the structure of intact primer (6, 7). Primers made in vitro by the *E. coli* *dnaG* protein vary in chain length up to nearly thirty nucleotide and contain a mixture ribo- and deoxyribonucleotides (25).

We thank Drs. K. Ueda and T. Sekiya for generous gifts of dihydroxyboryl Bio-gel P-60 and E. coli tRNA, respectively, and Dr. N. R. Cozzarelli for critical reading of the manuscript. This work was supported by grants from the Ministry of Education, Science and Culture, Japan.

† Present address: Cancer Institute (JFCR), Tokyo, Japan 170

* To whom correspondence should be addressed.

** Abbreviations used: TEAB, Triethylammonium bicarbonate; PEI-cellulose, polyethyleneimine cellulose; Borate gel, Bio-gel P-60 coupled with dihydroxyboryl group.

REFERENCES

1. Masamune, Y., Frenkel, G. D. and Richardson, C. C. (1971) J. Biol. Chem. 246, 6874-6879
2. Sternglanz, R., Wang, H. F. and Donegan, J. J. (1976) Biochemistry 15, 1838-1843
3. Shinozaki, K. and Okazaki, T. (1977) Molec. gen. Genet. 154, 263-267
4. Shinozaki, K. and Okazaki, T. (1978) Nucleic Acids Res. 5, 4245-4260
5. Roychoudhury, R. (1973) J. Biol. Chem. 248, 8465-8473
6. Ogawa, T., Hirose, S., Okazaki, T. and Okazaki, R. (1977) J. Mol. Biol. 112, 121-140
7. Okazaki, T., Kurosawa, Y., Ogawa, T., Seki, T., Shinozaki, K., Hirose, S., Fujiyama, A., Kohara, Y., Machida, Y., Tamanoi, F. and Hozumi, T. (1978) Cold Spring Harbor Symp. Quant. Biol. 43, 203-219
8. Kurosawa, Y. and Okazaki, T. (1979) J. Mol. Biol. in press
9. Studier, F. W. (1969) Virology 39, 562-574
10. Glynn, I. M. and Chappell, J. B. (1964) Biochem. J. 90, 147-149
11. Okazaki, R. (1974) in DNA Replication, Wickner, R. B., Ed., pp. 1-32, Marcel Dekker, Inc., New York
12. Cashel, M., Lazzarini, R. A. and Kalbacher, B. (1969) J. Chromatog. 40, 103-109
13. Randerath, K. and Randerath, E. (1967) in Methods in Enzymology, Grossman, L. and Moldave, K., Eds., vol. 12A, pp. 323-347, Academic Press, New York
14. Steinschneider, A. and Fraenkel-Conrat, H. (1966) Biochemistry 5, 2735-2743
15. Okayama, H., Ueda, K. and Hayaishi, O. (1978) Proc. Natl. Acad. Sci. USA 75, 1111-1115
16. Kainuma-Kuroda, R. and Okazaki, R. (1975) J. Mol. Biol. 94, 213-228
17. Jay, E., Bambara, R., Padmanabham, R. and Wu, R. (1974) Nucleic Acids Res. 1, 331-353
18. Clausen, T. (1968) Anal. Biochem. 22, 70-73
19. Okazaki, R., Hirose, S., Okazaki, T., Ogawa, T. and Kurosawa, Y. (1975) Biochem. Biophys. Res. Commun. 62, 1018-1024
20. Ogawa, T. and Okazaki, T. (1979) Nucleic Acids Res. 7, 1621-1634
21. Scherzinger, E., Lanka, E., Morelli, G., Seiffert, D., Yuki, A. (1977) Europ. J. Biochem. 72, 543-558
22. Hillenbrand, G., Morelli, G., Lanka, E. and Scherzinger, E. (1978) Cold Spring Harbor Symp. Quant. Biol. 43, 449-459
23. Richardson, C. C., Romano, L. J., Kolodner, R., LeClerc, J. E., Tamanoi, F., Engler, M. J., Dean, F. B. and Richardson, D. S. (1978) Cold Spring Harbor Symp. Quant. Biol. 43, 427-440

24. Liu, C. C. Burke, R. L., Hibner, U., Barry, J. and Alberts, B. (1978) Cold Spring Harbor Symp. Quant. Biol. 43, 467-487
25. Kornberg, A. (1978) in DNA Synthesis, Molineux, I. and Kohiyama, M. Eds., pp. 705-724, Plenum Press, New York
26. Mukai, J.-I. (1965) Biochem. Biophys. Res. Commun. 21, 562-567