Volume 14 Number 18 1986 Nucleic Acids Research

Selection of template initiation sites and the lengths of RNA primers synthesized by DNA primase are strongly affected by its organization in a multiprotein DNA polymerase alpha complex

Jamboor K.Vishwanatha, Masamitsu Yamaguchi[1,2], Melvin L.DePamphilis[1] and Earl F.Baril

Cell Biology Group, Worcester Foundation for Experimental Biology, Shrewsbury, MA 01545, [1]Department of Biological Chemistry, Harvard Medical School, Boston, MA 02115, USA and [2]Laboratory of Cell Biology, Aichi Cancer Center, Research Institute, Chikusa-ku, Nagoya, Japan

Received 28 May 1986; Revised 19 August 1986; Accepted 26 August 1986

ABSTRACT

Synthesis of (p)ppRNA-DNA chains by purified HeLa cell DNA primase-DNA polymerase α (pol α-primase) was compared with those synthesized by a multiprotein form of DNA polymerase α (pol α_2) using unique single-stranded DNA templates containing the origin of replication for simian virus 40 (SV40) DNA. The nucleotide locations of 33 initiation sites were identified by mapping G*pppN-RNA-DNA chains and identifying their 5'-terminal ribonucleotide. Pol α_2 strongly preferred initiation sites that began with ATP rather than GTP, thus frequently preferring different initiation sites than pol α-primase, depending on the template examined. The initiation sites selected *in vitro*, however, did not correspond to the sites used during SV40 DNA replication *in vivo*. Pol α_2 had the greatest effect on RNA primer size, typically synthesizing primers 1-5 nucleotides long, while pol α-primase synthesized primers 6-8 nucleotides long. These differences were observed even at individual initiation sites. Thus, the multiprotein form of DNA primase-DNA polymerase α affects selection of initiation sites, the frequency at which the sites are chosen, and length of RNA primers.

INTRODUCTION

The initiation of DNA synthesis in prokaryotes and eukaryotes involves the synthesis of RNA oligomers (primers) that are elongated by the replicative DNA polymerase (1, 2). A major fraction of nascent DNA chains (Okazaki fragments) that are isolated from viral infected (3-6) and noninfected (7,8) eukaryotic cells contain transient oligoribonucleotides at the 5' terminus that are covalently linked to the nascent DNA. The RNA oligomers were coined initiator or iRNA by Eliasson and Reichard (9), and details of their structure and synthesis were elucidated from studies of the replication of the DNA of the closely related polyoma virus and simian virus 40 (SV40). Accordingly, it was shown that the RNAs initiate with ATP and GTP, are of unique size (8-10 nucleotides), without a unique internal sequence (3,6,9-11) and with a near random composition at the RNA-DNA junction (4,5,11). Analysis of Okazaki fragments in replicating SV40 DNA *in vivo* indicated that the primary initiation site was 3'-purine-dT-5' in the DNA template and the secondary initiation site

was 3'-purine-dC-5' and the 5' end of the iRNA was complementary to either dT or dC (11). Moreover, the length of the RNA primers was not uniform but varied from 3 to 12 bases depending on the initiation site that was used.

A DNA primase activity, similar to the dna G gene product in Escherichia coli (1), has now been shown to be tightly associated with major or minor species of the replicative DNA polymerase from a wide variety of eukaryotic cells (12-19). The DNA primase-polymerase complex in the presence of single-stranded DNA template and complementary rNTPs and dNTPs synthesizes RNA oligomer primed DNA in vitro in a manner similar to the synthesis of Okazaki fragments in vivo (19,20).

A multiprotein (640 kDa) form of DNA polymerase α (pol α_2) was previously purified from HeLa cells and characterized (21,22). The pol α_2 complex was shown to have in addition to polymerase α activity, associated DNA primase (22), exonuclease activities, and the primer recognition proteins C1,C2 (21,23). The C1,C2 proteins from HeLa cells allow the homologous polymerase α to function with single-stranded DNA templates and RNA or DNA primers of 2 to 6 nucleotides in length (24). In contrast, the DNA polymerase α catalytic subunit resolved from pol α_2 required RNA or DNA primers of 6 to 10 nucleotides in length for maximal rates of in vitro synthesis with single-stranded DNA templates (23).

In this study, we have investigated the role of the additional proteins of DNA polymerase α_2 in the synthesis of RNA primers. By comparing the selection of initiation sites and sizes of the RNA primers synthesized by pol α_2 and pol α-primase on unique single-stranded regions of SV40 DNA, we discovered that the lengths of RNA primers was most strongly affected by association of primase in a multiprotein complex (pol α_2). We attribute this to the association of the primer recognition proteins C1,C2 with pol α_2 that were previously shown to promote the recognition of RNA or DNA primers with DNA templates as short as dinucleotides. The rate of DNA synthesis in those studies was found to increase proportionally with increased primer length from 2 to 6 nucleotides (24). Thus, we believe that the C1,C2 proteins may function as proteins involved in switching from transcription to replication in DNA synthesis in vivo.

MATERIALS AND METHODS

Materials.

Nucleotides, GpppNs and bacteriophage T4 DNA polymerase were purchased from P-L Biochemicals and polyethyleneimine (PEI)-cellulose was from

Brinkman Instruments, Inc. Radioactive nucleotides were from New England Nuclear Corporation. Biogel P-60 was obtained from Bio-Rad Laboratories. DNA polymerase I Klenow fragment and restriction enzymes were purchased from New England Biolabs; nuclease P1 from Boehringer-Mannheim Chemicals and vaccinia virus guanylyltransferase, calf thymus terminal deoxynucleotidyl transferase and RNase T2 from Bethesda Research Laboratories. Linear single-strand SV40 DNA inserts from mSV01 and mSV02 clones (10) were isolated by making use of the self-complementary sequence in the polylinker region of M13mp7 DNA as described previously (20).

RNA-DNA synthesis reactions.

Coupled RNA-DNA synthesis reactions were carried out at 30°C for 20 minutes in 100 µl containing 50 mM HEPES pH 7.8, 15% glycerol, 13 mM magnesium acetate, 10 mM sodium acetate, 10 µg bovine serum albumin, 1 mM dithiothreitol, 5 µg of mSV insert, 0.1 mM (4 µCi) each of [^{3}H]dNTPs, 0.2 mM each of GTP, CTP and UTP, 4 mM ATP and 8 units of polymerase α. Reactions were terminated by the addition of 15 mM EDTA and 0.5% SDS and then incubated for 30 minutes at 37°C with 100 µg/ml proteinase K. After phenol extraction, the products of the reaction were purified by gel filtration through a 5 ml Biogel P-60 column equilibrated in 10 mM Tris-HCl (pH 8.0), 1 mM EDTA, and 20 mM NaCl. DNA in the flow-through was concentrated by ethanol precipitation. Nascent DNA chains were then elongated to the ends of the template by the addition of 30 units/ml of *E. coli* Pol I Klenow fragment in a 100 µl reaction consisting of DNA, *E. coli* Pol I, 50 mM Tris-HCl (pH 7.4), 10 mM $MgCl_2$, 1 mM dithiothreitol and 1 mM of each dNTP. The reaction was incubated at room temperature for 2 hours and was terminated by the addition of 6 mM EDTA. After phenol extraction, the products of the reaction were purified by gel filtration through a 5 ml Biogel P-60 column and DNA in the flow-through was concentrated by ethanol precipitation.

Capping reaction.

RNA-DNA chains containing a di- or triphosphate at their 5' ends were labeled using [α-^{32}P]GTP and vaccinia virus guanylyltransferase as described before (20).

Analysis of the size of RNA primers.

Nascent DNA chains, end-labelled at their 5'-ends by capping, were denatured for 2 minutes at 100°C in 50 mM Tris-HCl (pH 8.0), 6 mM $MgCl_2$, 25 mM KCl and 0.5 mg/ml carrier tRNA, and then rapidly cooled in ice-water. Digestion of DNA was started by addition of 800 units/ml of bacteriophage T4 DNA polymerase - 3',5' exonuclease and carried out for 2

hours at 37°C. Products of the reaction were fractionated by electrophoresis in 22% polyacrylamide gels (0.5 mm x 33 cm x 42 cm) containing 8 M urea, 100 mM Tris-borate (pH 8.3) and 1 mM EDTA (10). Electrophoresis was carried out for 5 hours at 1500 V using bromophenol blue and xylene cyanol as tracking dyes. Gels were then exposed to Kodak XAR-5 film with a Cronex-plus intensifying screen at -70°C.

To determine the migration of the cap structure in the above gel, an aliquot of the denatured, capped nascent DNA was incubated with RNase T_2 in a 10 µl reaction containing 50 mM sodium acetate (pH 5.3), 2 mM EDTA and 20 units of RNase T_2. The reaction was incubated for 2 hours at 37°C and the products of the reaction were electrophoresed on 22% polyacrylamide gels described above.

Nucleotide locations of initiation sites.

The template locations of 5'-end labeled, capped nascent DNA were mapped at single nucleotide resolution by electrophoresis in an 8% polyacrylamide/8 M urea gel. The sequence of an appropriate mSV insert was carried out in parallel on the same gel. The 3'-ends of the mSV DNA were labelled by incorporation of [α-^{32}P]-cordycepin-5'-triphosphate using terminal transferase in a tailing reaction. The [3'-^{32}P]DNA fragment was sequenced by the chemical degradation method (24). Polyacrylamide gels (0.25 mm x 33 cm x 42 cm) containing 8% acrylamide with a 20:1 ratio of acrylamide to N,N1-methylenebisacrylamide were prepared in 8 M urea, 100 mM Tris-borate (pH 8.3) and 1 mM EDTA (25). Electrophoresis was carried out at 2,000 V for 2-3 hours. Gels were then transferred to used X-ray film and exposed to Kodak X-Omat AR film with a Cronex-plus intensifying screen at -70°C.

Recovery of DNA from polyacrylamide gels.

Radiolabeled DNA was recovered from 8% polyacrylamide/8 M urea gels by slicing the radioactive region out of the gel and extracting the individual slices in 0.5 M ammonium acetate, 10 mM magnesium acetate and 0.1% SDS at 60°C for 16 hours. After centrifugation, DNA in the supernatant was precipitated with ethanol using tRNA as carrier. Ethanol precipitation was repeated twice to remove traces of SDS.

Identification of 5'-terminal GpppN in nascent RNA-DNA chains.

Nascent DNA labeled at the 5'-end by capping was digested with nuclease P1 to identify the terminal GpppN. In a 10 µl reaction, labeled DNA was incubated with 30 mM Na- acetate (pH 5.3) and 2 µg of nuclease P1 for 2 hours at 37°. Products of the reaction were then subjected to ascending chromatography on a PEI-cellulose plate in 1.6 M LiCl for 8-9 hours along

Table I. Summary of the Catalytic Activities Associated With DNA Polymerase α_2 and the DNA Primase-Polymerase α Complex.

Activity	DNA polymerase α_2	DNA primase-polymerase α
	unit/mg protein	
DNA polymerase α with:		
a. activated DNA	105	1736
b. denatured DNA	71	0
DNA primase	234	6225
3'—>5' exonuclease	26	0
RNase H under:		
a. conditions for RNase H assay	109	107
b. conditions for DNA primase assay	23	71
ATPase:		
a. DNA dependent	0	0
b. DNA independent	0	0

DNA polymerase α_2 was purified from HeLa cell extract according to a published procedure (22). HeLa cell DNA primase-polymerase α was resolved from DNA polymerase α_2 by hydrophobic affinity chromatography on phenylSepharose and purified further by phosphocellulose chromatography and glycerol gradient centrifugation (Vishwanatha and Baril, submitted for publication). The procedures for assay of DNA polymerase α, primase, exonuclease, and ATPase activities were as previously described (22). RNase H activity was assayed with poly(dT).[^{3}H] poly(A)(12μM, 800 cpm/pmol) as substrate according to the procedure of DiFrancesco and Lehman (31). Units of enzyme activity are defined as follows: DNA polymerase α, 1nmol of total dNTP incorporated/ h at 30°C, ATPase, 1nmol of charcoal nonadsorbable [$^{32}PO_4$] produced/ h at 30°C and RNase H, 1nmol of [^{3}H] poly(A) hydrolyzed in 15 min at 35°C.

with unlabeled cap standards. After autoradiography of the PEI-cellulose plate, radioactive spots were excised and their radioactivity was quantitated by scintillation counting. Positions of the unlabeled cap standards were visualized by UV-light.

RESULTS

Enzyme activities associated with purified DNA polymerase α2 and DNA polymerase α-primase.

The catalytic activities associated with the multiprotein DNA polymerase α_2 complex and the DNA polymerase α-primase complex that is resolved from it are summarized in Table I. As previously reported the DNA polymerase

α_2 complex has associated DNA polymerase α, primase, and a single-strand specific exodeoxyribonuclease activity (22). It also has the associated primer recognition accessory proteins C1,C2 (21-23,27). The accessory proteins permit polymerase α to function with DNA templates having a low primer to template ratio such as denatured DNA (Table I)(27). The purified DNA polymerase α_2 also has a tightly associated RNase H activity. It was previously shown that additional polypeptides are present during SDS polyacrylamide gel electrophoretic analysis of the homogeneous DNA polymerase α_2 multiprotein complex to which catalytic or accessory functions could not be assigned (22). It is possible that the RNase H activity is attributable to one or more of these polypeptides. This is currently being investigated.

The DNA primase-polymerase α complex that is resolved from polymerase α_2 and then further purified (Vishwanatha and Baril, submitted for publication) lacks exodeoxyribonuclease activity (Table I). It also does not function with DNA templates containing a low primer to template ratio due to the separation of the C1,C2 primer recognition proteins during the course of purification of the DNA primase-polymerase α complex (22). The purified DNA primase-polymerase α complex, like DNA polymerase α_2, has a tightly associated RNase H activity (Table I and S. Coughlin and E. Baril, unpublished results). The ratio of DNA primase to polymerase α activities in the DNA polymerase α_2 complex and during the course of resolution and purification of the DNA primase-polymerase α remain similar (Table I and Vishwanatha and Baril, submitted for publication). The recovery of RNase H activity relative to primase and polymerase activities decreases during the purification of the DNA primase-polymerase α complex. This is due to a loss or inactivation of RNase H activity in the resolution from polymerase α_2 rather than separation of a free RNase H activity (S. Coughlin and E. Baril, unpublished results). Interestingly, the RNase H associated with DNA polymerase α_2 complex is much less active under conditions for synthesis of RNA primers (Table I) and for DNA synthesis from the RNA primers (data not shown).

In view of the complex nature of DNA polymerase α_2 it was of interest to determine the result of this molecular complexity on the recognition of RNA primer starts and the length of RNA primers synthesized *in vitro* on SV40 single-stranded DNA templates.

Mapping the locations of the 5' ends of RNA primers synthesized on single-stranded linear mSV01 and mSV02 DNA templates.

The location of the initiation sites for DNA primase on linear mSV01 and

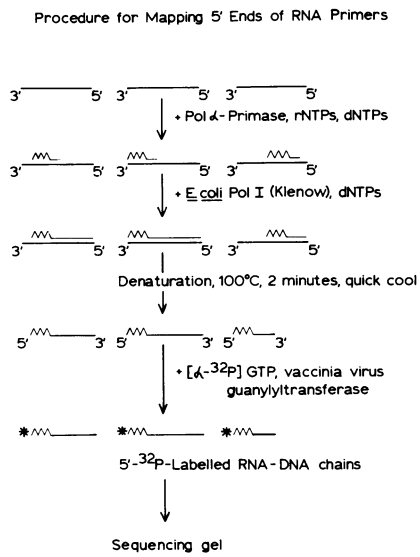


Fig. 1. Outline of procedure used for mapping the 5'-ends of RNA primers on mSV01 and mSV02 insert DNAs.
Symbols are 3'----- 5' DNA template; ΛΛΛ RNA primer; ΛΛΛ_____ nascent DNA; asterisks indicates 5' - α^{32}P label. The single-stranded linear mSV01 (311 bases) insert DNA containing the SV40 ori region or the complimentary mSV02 insert DNA (Fig. 3) was incubated with pol α - primase or with DNA pol α_2 in the presence of all four NTPs and dNTPs under the assay conditions for RNA-DNA synthesis. The nascent DNA chains were extended to the end of the template by incubation with E. coli DNA polymerase I (Klenow). The product was isolated and the 5' termini of the isolated (p)ppRNA-DNA chains were radiolabeled in the capping reaction by incubation with [α-^{32}P]GTP and vaccinia virus guanylyltransferase (20,26). The size of the 5'-^{32}P-RNA-DNA chains were determined by electrophoresis in 8% polyacrylamide gels containing 8M urea. The nucleotide location of the 5'ends of RNA primers was determined by comparing the migration of these bands with those released by sequencing ^{32}P-labeled DNA chains of the same polarity as the nascent DNA. When the single-stranded mSV01 DNA insert was used as the template the nascent DNA was extended to the end of the template with Klenow fragment past a single HindIII site that is located 239 bases downstream from the 3' end of this template (Fig. 3, position 5184). One aliquot of the incubated sample was digested with HindIII to release nascent DNA chains with identical 3' termini and the remaining aliquot was used as the control. The aliquots were both heat-denatured and fractionated in an 8% polyacrylamide-8M urea gel. Details are given in Materials and Methods.

mSV02 DNA templates was performed as described by Yamaguchi et al., (20) and outlined in Fig. 1. The 5' termini of the isolated (p)ppRNA-DNA chains were radiolabeled in the capping reaction by incubation with [α-^{32}P]GTP and vaccinia virus guanylyltransferase (20,26). The advantage of this technique is that guanylyltransferase specifically adds GMP to RNA chains with

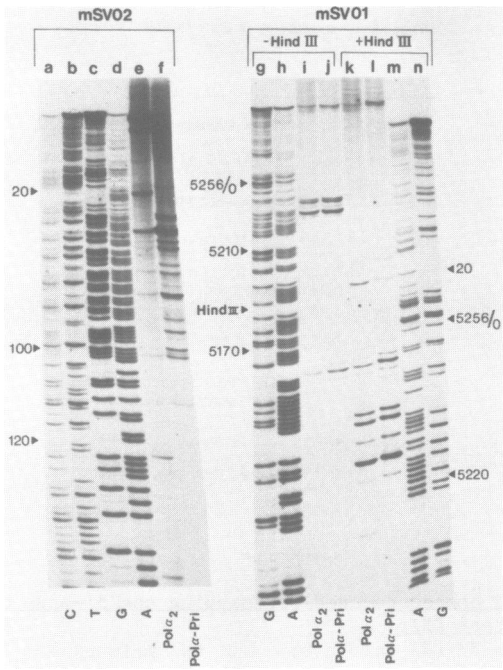


Fig. 2. Mapping the locations of 5'-ends of RNA primers synthesized on mSV01 and mSV02 templates by DNA polymerase.
(p)pp RNA-DNA chains synthesized on either mSV01 or mSV02 template were extended to the end of the template with Klenow fragment and were then radiolabeled in the capping reaction (Fig. 1; Materials and Methods). The nucleotide sequence of DNA chains with the same polarity as nascent DNA was determined concurrently using the Maxam-Gilbert sequencing procedure on [3'-^{32}P]DNA. Lanes a,b,c and d are cleavage at C, cleavage at C and T, cleavage at G and cleavage at A and G respectively. G*pppN-RNA-DNA chains made by pol α_2 on mSV02 are in lane e and those made by pol α - primase are in lane f. Lanes g and h are cleavages at G and A and G respectively. Lane i and j are G*pppN-RNA-DNA chains made by pol α_2 and pol α - primase respectively, on mSV01. To resolve the 5'-ends of RNA primers made on mSV01, the material shown in lanes i and j were digested with HindIII and the digested material is shown in lanes k and l. Lanes m and n show the nucleotide sequence of HindIII digested [3'-^{32}P]DNA of the same polarity. Lane m shows cleavage at A and G and lane n shows cleavage at G. Numbers on the vertical axis are nucleotide locations in SV40 wt 800 (11). The single HindIII site is also indicated.

diphosphate or triphosphate at the 5' terminus so only RNA chains with these 5' termini are radiolabeled and the amount of radioactivity in the nascent chain is independent of its length or composition.

The overall patterns and intensity of the gel bands of nascent RNA-DNA chains synthesized by pol α_2 or pol α-primase on mSV01 DNA were very

similar; although some differences were observed in the choice of initiation sites and the frequency at which a particular site was used on mSV02 DNA (Fig. 2 lanes e and f). Since each 5'-terminus of the RNA-DNA chains is ^{32}P-labeled in the capping reaction, the amount of radioactive label at a given position in the gel represents the relative frequency of initiation at that position (20). With mSV01 insert DNA as the template, incubation with pol α_2 or pol α-primase resulted in about the same pattern of migration of nascent RNA-DNA chains in the gel (Fig. 2, lanes i and j). Slight differences in the frequency of initiation at a few positions were apparent. A total of 14 initiation sites on the mSV01 insert DNA were selected both by pol α_2 and pol α - primase (Fig. 3). The relative frequency of initiations from each of these sites was about the same for each form of DNA primase - polymerase α. The preferred initiation sites exist predominately within clusters around positions 5224 and 5156. On the mSV02 template the specific initiation sites were more dispersed and fewer initiation sites were used in common by both forms of primase. Out of 19 sites for initiation, only 7 were used by both enzymes while 7 were unique for pol α - primase. The location of the in vitro initiation events for the synthesis of RNA-DNA chains on mSV01 and mSV02 DNA templates do not correspond to the location of 5'-terminal ribonucleotides of RNA primers observed in SV40 DNA replication in vivo (10).

The 5'-terminal nucleotide composition of nascent RNA-DNA chains.

The 5' terminal nucleotide of nascent RNA-DNA synthesized in vivo is about 70% pppA and 30% pppG (9,10). To confirm the mapping data described in Fig. 3, the capped nascent RNA-DNA chains were digested with nuclease P1, the 5'G*pppN products were separated by thin layer chromatography on PEI-cellulose and the radioactivity in the excised spot determined (Fig. 4). The 5' terminal nucleotide of RNA-DNA chains synthesized by pol α_2 on mSV02 DNA were slightly higher in A over G while with mSV01 DNA the 5' terminal nucleotide of synthesized RNA-DNA chains were predominately A (over 80%) over G (Fig. 4 lanes a and c). With pol α-primase the 5' terminal nucleotide of RNA-DNA chains synthesized on mSV02 DNA were biased in favor of G (Fig. 4 lane b). In these experiments 5'-terminal U or C on RNA-DNA chains synthesized by pol α_2 or pol α-primase with either DNA template were not observed even after exposure of the autoradiographs for 72 hours. These data indicate that pol α_2 shows a preference for initiation with A irrespective of the template while the initiation by pol α-primase shows a greater template dependence. These data are in agreement with the results

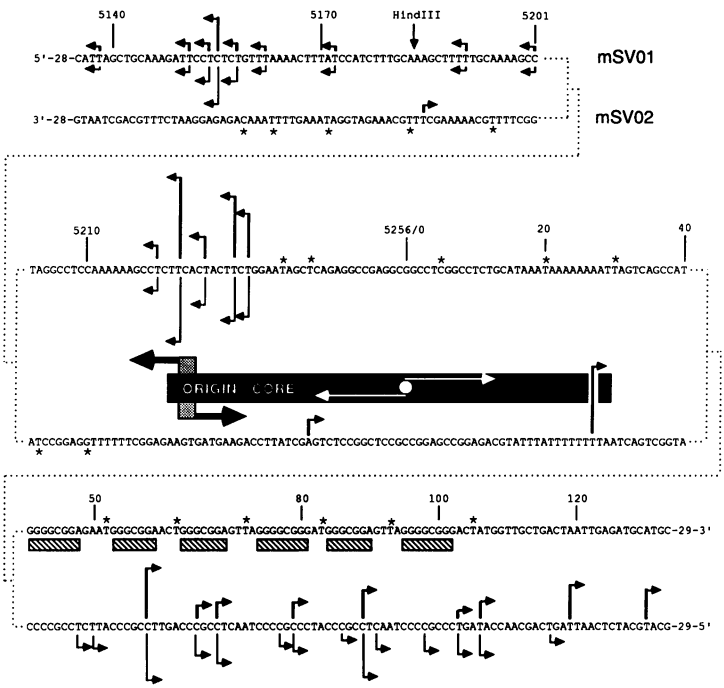
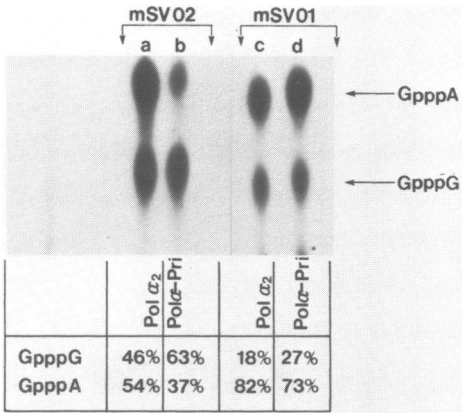


Fig. 3. Nucleotide locations of 5'-ends of RNA primers synthesized on SV40 DNA templates in vitro and in vivo.
In vitro results. The nucleotide sequence for the region analyzed (mSV01 and mSV02) is shown and the nucleotides are numbered with respect to the SV40 wt800 sequence (10). Vertical bars above the typed sequence indicate the initiation sites of pol α_2 and those below the typed sequence indicate the initiation sites used by pol α - primase. Arrows indicate the direction of synthesis. The relative height of each bar represents the relative frequency of initiation events at each site.
In vivo results. Asterisks over nucleotides indicate the positions of 5'-terminal nucleotides of RNA primers observed during SV40 DNA replication in vivo (11). The minimum sequence required for the origin of replication (origin core) is indicated by the shaded box. The transition point from discontinuous to continuous DNA synthesis in vivo (i.e. the origin for bidirectional replication) is indicated by the stippled vertical area and solid arrows. A 27 bp palindrome is indicated by ⟵○⟶ . Hatchmarked rectangles identify consensus sequences 5'-$(G)_{3-4}$ CGGA on mSV01 DNA. The single HindIII site indicated by an arrow is 239 bases (position 5184) from the 3'-end of the template.

from mapping of 5' nucleotide initiations on these templates (Fig. 2, lanes i,j, and Fig. 3). On mSV01 DNA template, both pol α-primase and pol α_2 appeared to use the same initiation sites and with about the same relative frequency of initiation events while on mSV02 DNA there were fewer initiation sites used in common by the two enzymes (Fig. 2, lanes e,f, and Fig. 3,).



	Pol α_2	Polα-Pri	Pol α_2	Polα-Pri
GpppG	46%	63%	18%	27%
GpppA	54%	37%	82%	73%

Fig. 4. 5'-Terminal nucleotide composition of nascent RNA-DNA chains
Nascent RNA-DNA chains synthesized with pol α_2 and pol α - primase were radiolabeled at their 5'-ends by capping with [α-^{32}P]GTP and vaccinia virus guanylyltransferase. These were digested with nuclease P1 and chromatographed on PEI (Material and Methods). Lanes a and c are products of pol α_2 on mSV02 and mSV01, respectively. Lanes b and d are products of pol α_2 and pol α - primase on mSV02 and mSV01 respectively. Positions of the unlabeled capped standards are indicated by arrows.
Table. Radioactive spots seen after autoradiography were cut out and the radioactivity was determined by scintillation counting. Percentage of total radioactivity corresponding to each spot is indicated in the table.

Average size of RNA primers synthesized on mSV01 and mSV02 DNA templates. The effect of the molecular form of DNA primase and the template on the length of RNA primers was determined. The products synthesized by pol α - primase or pol α_2 on mSV01 and mSV02 DNA templates were radiolabeled in the capping reaction and divided into three aliquots. One aliquot was digested by T4 polymerase associated 3' —> 5' exonuclease (T4 exo) under conditions that remove all but a single dNMP from the 3' end of the RNA chains (20). A second aliquot was digested with RNase T2 and the remaining aliquot was used as the untreated (uncut) control. The control and treated products were fractionated by electrophoresis in a polyacrylamide-urea gel (Fig. 5).

The untreated (uncut) radiolabeled polynucleotides synthesized by pol α_2 and pol α -primase on mSV01 (lanes a and b) and mSV02 DNA (lanes g and h) ranged in length from 25 bases to the size of the template. The radiolabel in the polynucleotide was released following digestion with RNase T2. Synthesis was influenced both by the molecular form of the DNA primase and the template sequence (Fig. 5; Table II). After digestion of the DNA of

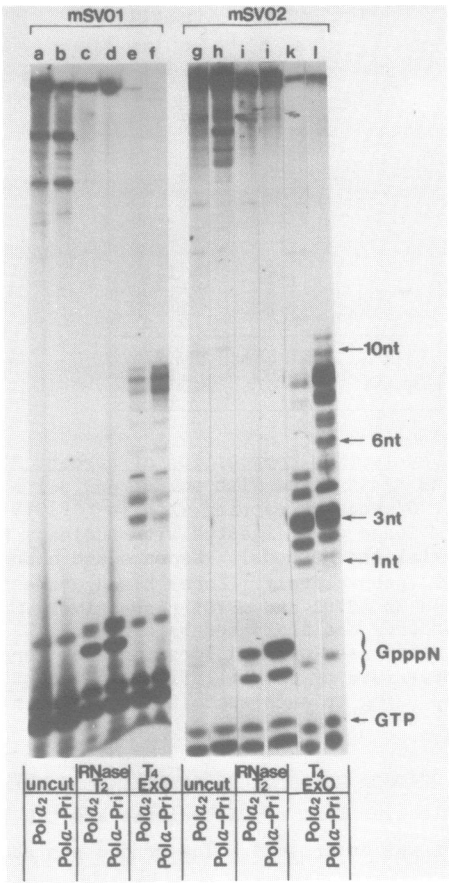


Fig. 5. Size of RNA primers synthesized on mSV01 and mSV02 templates by HeLa cell DNA pol α_2 and pol α - primase.
(p)pp RNA-DNA chains were synthesized and their 5'-ends were labeled with [α-^{32}P]GTP using vaccinia guanylyltransferase as described under 'Materials and Methods'. The G*pppN-labeled products were either left untreated (lanes a,b,g,h), treated with RNase T_2 (lanes c,d,i,j) or treated with T_4-DNA polymerase associated exonuclease (lanes e,f,k,l) and then fractionated by electrophoresis on a 22% polyacrylamide-8M urea gel (24). Lanes a,c and e represent products of pol α_2 on mSV01 template. Lanes b, d and f represent products of pol α - primase on mSV01 template. Lanes g, i and k represent the products of pol α_2 on mSV02 template and lanes h, j and l represent products of pol α - primase on mSV02 template. Numbers on the vertical axis indicate the size in number of nucleotides of the primers after accounting for the cap size and the last dNMP left after T_4-DNA polymerase associated exonuclease digestion.

TABLE II. Length distribution of RNA primers synthesized on mSV01 and mSV02 DNA templates.

Length (bases)	Distribution (%)[a]			
	DNA polymerase α_2		DNA primase - polymerase α	
	mSV01	mSV02	mSV01	mSV02
1-5	62	89	37	49
6-10	38	11	63	51

[a]The proportion of RNA primers of each length was calculated by excising the appropriate sections of each gel lane (Fig. 5) and measuring the radioactivity by liquid scintillation counting. The lengths of the RNA primers were corrected for the presence of a cap structure and 3'-dNMP (20).

the RNA-DNA chains by T4 DNA polymerase associated exonuclease, radiolabeled oligonucleotides appeared in the region identified by the markers $p(Ap)_3$, $p(Ap)_6$ and $p(Ap)_{10}$ (Fig. 5, lanes e,f,k and l). A striking difference was seen in the length of oligoribonucleotides that were synthesized by pol α_2 (lanes e and k) compared to pol α - primase (lanes f and l). This difference was most pronounced for the mSV02 template in which the length of 89% of the RNA primers synthesized by pol α_2 were in the range of 1 to 5 nucleotides (Table II) while the length of only 49% of the primers synthesized by the pol α - primase were in this range. With the mSV01 DNA template 62% of the RNA primers synthesized by pol α_2 had a length in the range of 1 to 5 nucleotides while 63% of the RNA primers synthesized by pol α - primase were in the range of 6 to 10 nucleotides in length.

Since DNA polymerase α_2 and DNA primase-polymerase α have RNase H activity that cofractionates with the enzymes it was necessary to first determine if the difference in primer lengths synthesized by the enzymes is attributable to the RNase H. To ascertain this, preformed RNA oligomers $(A)_6$ and $(A)_{10}$ were 5'[^{32}P] end labeled in separate reactions using T_4 polynucleotide kinase and [γ-^{32}P]ATP. The labeled product was isolated and hybridized to poly $(dT)_{1000}$ at a molar nucleotide ratio of 10:1. The poly (dT).[^{32}P]RNA hybrids were isolated and used as substrates in reactions under optimal conditions for the assay of

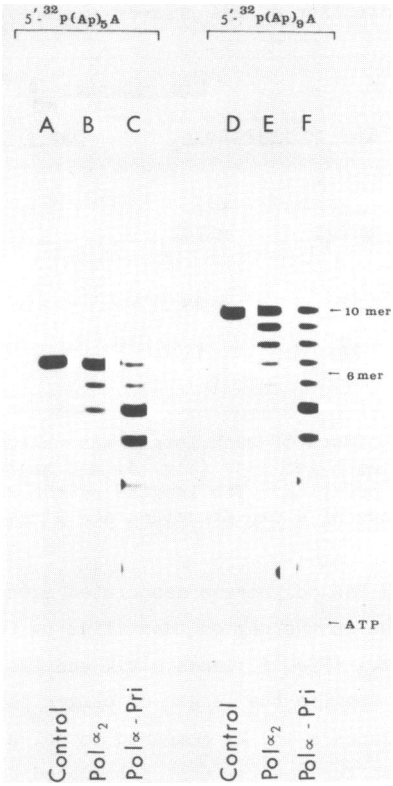


Fig 6. Analysis of RNase H activity with preformed [^{32}P]RNA primers. Hela cell DNA polymerase α_2 or DNA primase-polymerase α were incubated with either poly(dT)$_{1000}$.5'-^{32}P$_2$(Ap)$_5$A or poly(dT)$_{1000}$.5'-^{32}P (Ap)$_9$A hybrids at a 10:1 molar nucleotide ratio under optimal conditions for the assay of RNase H. After incubation at 35°C for 10 min the reaction was terminated by the addition of EDTA and SDS to final concentrations of 10mM and 0.5% respectively. The products were isolated by phenol extraction, precipitation with ethanol and separated by electrophoresis on a 22% polyacrylamide gel (0.3cm x 33 cm x 42 cm) containing 8m urea in a 100mM Tris borate pH 8.3, 1mM EDTA buffer. Electrophoresis was for 5 hours at 1500 V using bromophenol blue and xylene cyanol as tracking dyes. The gel was exposed to Kodak XAR-5 film using a Cronex-plus intensifying screen. Lanes A and D, controls of the 5'-^{32}P(Ap)$_5$A and 5'-^{32}P(Ap)$_5$A oligomers respectively, of the polydT.A hybrids incubated in the absence of enzyme. Lanes B and C, products from the hydrolysis of poly(dT).5'-^{32}P (Ap)$_5$A by 0.24 units of RNase H associated with DNA polymerase α_2 and DNA primase-polymerase α respectively. Lanes E and F, products from the hydrolysis of poly(dT).5'-^{32}P(Ap)$_9$A by 0.24 units of RNase H associated with DNA polymerase α_2 and DNA primase-polymerase α respectively. The positions of the 5'-^{32}P (Ap)$_5$A, 5'-^{32}P(Ap)$_9$A and [γ-^{32}P]ATP markers that were run in parallel on the same gel are indicated by arrows.

the polymerase α_2 and DNA primase-polymerase α associated RNase H (S. Coughlin and E. Baril, unpublished). The products from the hydrolysis were separated by electrophoresis in 22% polyacrylamide gels in the presence of 8M-urea. As shown in Fig. 6, hydrolysis of the poly (dT) hybridized [5'-^{32}P](Ap)$_5$A and [5'-^{32}P](Ap)$_9$A did occur during the incubation with polymerase α_2 and DNA primase-polymerase α under optimal assay conditions for the RNase H. However, the hydrolysis was much more extensive during the incubation with DNA primase-polymerase α than with polymerase α_2 (compare lanes B,C and E,F in Fig. 6). The hydrolysis by the polymerase α_2 associated RNase H was markedly reduced (over 80%) under the conditions for synthesis of RNA primers and DNA synthesis reported here (Table 1 and data not shown). The RNase H activity associated with DNA primase-polymerase α was also reduced under these conditions but to a lesser extent. On the basis of these results it does not appear that the shorter RNA primer lengths synthesized by DNA polymerase α_2 relative to DNA primase-polymerase α is attributable to the RNase H activity associated with the enzymes.

Size of individual RNA primers.

Our analysis of the total population of RNA primers synthesized by pol α_2 and pol α - primase on mSV01 and mSV02 DNAs indicated that size of the primer synthesized varied from 1 to 9 depending on the DNA template and whether pol α - primase or the pol α_2 was used in the assay. As discussed above this variation does not appear to be due to the RNase H activity associated with these enzymes (Fig. 6). These data suggested that the initiation site on the DNA and/or the form of the DNA primase influence the length of RNA primers that are synthesized. Since pol α - primase and pol α_2 used similar initiation sites on mSV01 insert DNA (Fig. 2), we used this DNA template for synthesis of RNA-DNA chains by each enzyme. Five nascent capped RNA-DNA chains were separated by gel electrophoresis (Fig. 7a), excised and eluted from the gel. The DNA segment of the eluted RNA-DNA chains was digested by T4 DNA polymerase associated exonuclease and the size of the RNA primers was determined by gel electrophoresis (Fig. 7b). In at least one instance (Fig. 7 lanes 2 and 7) the length of RNA primers synthesized by pol α_2 at the same initiation site on this template was clearly shorter (4 to 6 bases) than the length of RNA primers synthesized by pol α - primase (5 to 9 bases) at this site. In most cases there are 2 to 6 major products produced per site and the products synthesized by pol α_2 (3 to 7 bases) were shorter than those synthesized by pol α - primase (3 to 9

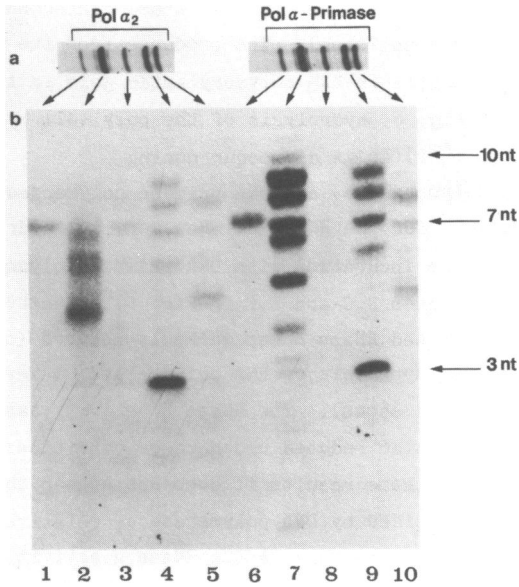


Fig. 7. Sizes of RNA primers from individual initiation sites on mSV01 template.
(p)pp RNA-DNA chains were synthesized, labeled and fractionated by electrophoresis on a sequencing gel as described in Figure 2. The five major products shown in Fig. 5, lanes k (for pol α_2) and l (for pol α - primase) were eluted separately from the gel, digested with T_4-DNA polymerase associated exonuclease and fractionated by electrophoresis in a 22% polyacrylamide-8M urea gel. Panel a shows the five major products each from pol α_2 and pol α- primase and Panel b shows the T_4-DNA polymerase associated exonuclease digestion products of the five major products. Arrows connect the major bands of the sequencing gel with their corresponding T_4-DNA polymerase associated exonuclease products. Numbers on the vertical axis indicate the size in nucleotides of the primers.

bases). In one case, a unique product of 7 bases in length was synthesized by both enzymes at the same initiation site (Fig. 7 lanes 1 and 6). Since the 5' end of each RNA-DNA chain originated from a single unique base in the template, variation in primer length must depend on the sequence of the initiation site. However, since the length of RNA primers synthesized by pol α_2 at the same site were frequently shorter than those synthesized by pol α - primase, the molecular form of DNA primase must also influence the length of primer.

DISCUSSION

The results presented here reveal that the organization of DNA primase into a multiprotein, DNA polymerase α complex (pol α_2) affects both the

length of RNA primers synthesized, and the template sites at which they are initiated. Both enzymes were capable of synthesizing RNA primers from 1 to 11 ribonucleotides in length. These preferences were apparent even at individual template initiation sites. The sizes of RNA primers measured in these experiments were consistent with the sizes of RNA primers synthesized by DNA primers from other organisms that were measured by the same technique. For example, crude Drosophila embryo extracts synthesized primers of 6 to 8 ribonucleotides on natural single-stranded DNA templates (26) and purified DNA primase-DNA polymerase α from CV-1 cells also synthesized primers of 6 to 8 ribonucleotides (20).

The bias to shorter RNA primers synthesized by pol α_2 is a result of enzyme complexity rather than an in vitro artifact. It does not appear to be the result of the RNase H activity associated with DNA polymerase α_2 since the RNase H activity associated with DNA primase-polymerase α was shown to be more active under optimal conditions for RNase H (Fig. 6) and under the conditions for synthesis of RNA primers used here (Table I). It is also not due to misincorporation of dNMPs into the RNA primer since pol α-primase was purified from the pol α_2 complex, and the templates and assay conditions were identical. We propose that the shorter RNA primers synthesized by pol α_2 may be attributable to the association of the C1 and C2 primer recognition proteins (21, 22). These proteins were shown to specifically interact with the homologous DNA polymerase α from HeLa (21) and CV-1 cells (27) and to promote recognition of 3'-hydroxyl termini on RNA and DNA primed single-stranded DNA templates (23). In the case of HeLa cells, association of C1,C2 with DNA polymerase α promoted recognition of primers as short as dinucleotides and the rate of in vitro DNA synthesis increased proportionally with increased primer length from 2 to 6 nucleotides (23). In contrast, the DNA polymerase α catalytic subunit required primers of 6 to 10 nucleotides in length for maximal in vitro synthesis with single-stranded DNA templates (23). The requirements for primer length previously observed for HeLa cell DNA polymerase α2 and the DNA polymerase α catalytic subunit are remarkably similar to RNA primer lengths syntheized by pol α_2 and pol α-primase, respectively, with natural single-stranded DNA templates. Thus, C1 and C2 primer recognition proteins may function in the switch mechanism from transcription to replication phase in synthesis of nascent DNA chains by the DNA polymerase α complex.

In vivo, the typical length of RNA primers synthesized on the same DNA templates used in this study and characterized by the same techniques was 9

to 11 residues and the range was 2 to 12 residues (11). The longer average length of RNA primers synthesized *in vivo* may reflect the presence of proteins uniquely associated with DNA templates in replicating chromosomes.

A second, but less dramatic, way in which pol α_2 differs from pol α -primase is in their selection of DNA template initiation sites. Both enzymes utilized the same templates to approximately the same extent, but pol α_2 strongly preferred sites that initiated synthesis with ATP, even when the template promoted initiation at sites that used GTP (e.g. mSV02). Thus, the more complex form of DNA primase-DNA polymerase α was less influenced by template sequence and more frequently initiated synthesis with ATP, consistent with the initiation of RNA primed-DNA synthesis *in vivo* (11). Nevertheless, both enzymes initiated synthesis at essentially the same pyrimidine-rich sequences previously observed with purified CV-1 cell DNA primase-DNA polynmerase α (20). These *in vitro* initiation sites are not the same as those observed on the same DNA templates during SV40 DNA replication *in vivo* (Fig. 3; 10), suggesting that additional proteins associated either with the enzyme or template strongly affect selection of initiation sites. All three forms of DNA primase-DNA polymerase α tested on mSV01 strongly preferred to initiate synthesis at the origin for bidirectional replication used *in vivo*, but the significance of this observation is not clear since these initiation sites are not observed *in vivo* (Fig. 3). One possible explanation is that the 27 bp palindrome located in *ori*-core (Fig. 3) promotes *in vitro* initiation at these sites. Deletions in this palindrome reduce initiation by purified DNA primase *in vitro* (28) and the cruciform configuration does not appear to exist in replicating SV40 chromosomes *in vivo* (29). The *ori*-region for preferred initiation by the pol α_2 and pol α-primase complex does not appear to correlate with the presence of hexanucleotide $^{\Psi}$-like consensus sequences, of the general formula C_2A_{1-2} (C_{2-3}/T_2), that was found to correlate with primase initiation siteson minute virus of mice DNA template (30). Further studies are required to elucidate the actual factors that are involved in the signals for the preferred initiations by DNA primase *in vivo*.

ACKNOWLEDGMENTS

This research was supported by National Institutes of Health Grants P30 12708, CA 15187, and CA 15579 from the National Cancer Institute and American Cancer Society Grant MV114. J.K.V. is a Fellow of the Leukemia Society of America. We are grateful to Lois Hager and Sandra Johnson for their assistance in preparation of this manuscript.

REFERENCES
1. Kornberg, A. (1980) DNA Replication. W.H. Freeman and Company, San Francisco.
2. DePamphilis, M.L. and Wassarman, P.M. (1980) Ann. Rev. Biochem. 49, 627-666.
3. Magnusson, G., Pigiet, V., Winnacker, E., Abrams, R. and Reichard, P. (1973) Proc. Natl. Acad. Sci. USA 70, 412-415.
4. Pigiet, V., Eliasson, R. and Reichard, P. (1974) J. Mol. Biol. 84, 197-216.
5. Anderson, S., Kaufmann, G. and DePamphilis, M.L. (1977) Biochemistry 16, 4990-4998.
6. Kaufmann, G., Anderson, S. and DePamphilis, M.L. (1977) J. Mol. Biol. 116, 549-567.
7. Tseng, B.Y. and Goulian, M. (1975) J. Mol. Biol. 99, 315-337.
8. Waqar, M.A. and Huberman, J.A. (1975) Cell, 6, 551-557.
9. Eliasson, R. and Reichard, P. (1978) Nature, 272, 184-185.
10. Hay, R.T., Hendrickson, E.A. and DePamphilis, M.L. (1984) J. Mol. Biol., 175, 131-157.
11. Hay, R.T. and DePamphilis, M.L. (1982) Cell, 28, 767-779.
12. Yagura, T., Kozu, T. and Seno, T. (1982) J. Biochem. (Tokyo), 91, 607-618.
13. Conaway, R.C. and Lehman, I.R. (1982) Proc. Natl. Acad. Sci. USA, 79, 2523-2527.
14. Tseng, B.Y. and Ahlem, C.N. (1983) J. Biol. Chem., 257, 7280-7283.
15. Chang, L.M.S., Rafter, E., Augl, C. and Bollum, F.J. (1984) J. Biol. Chem. 259, 14679-14687.
16. Plevani, P., Badaracco, G., Augl, C. and Chang, L.M.S. (1984) J. Biol. Chem., 257, 5700-5706.
17. Wang, T.S-F., Hu, S-Z. and Korn, D. (1984) J. Biol. Chem.,259, 1854-1865.
18. Gronostajski, R.M., Field, J. and Hurwitz, J. (1984) J. Biol. Chem., 259, 9479-9486.
19. Yamaguchi, M., Hendrickson, E.A. and DePamphilis, M.L. (1985) J. Biol. Chem. 260, 6254-6263.
20. Yamaguchi, M., Hendrickson, E.A. and DePamphilis, M.L. (1985) Mol. Cell. Biol. 5, 1170-1183.
21. Lamothe, P., Baril, B., Chi, A., Lee, L. and Baril, E. (1981) Proc. Natl. Acad. Sci. USA 78, 4723-4727.
22. Vishwanatha, J.K., Coughlin, S.A., Wesolowski-Owen, M. and Baril, E.F. (1986) J. Biol. Chem. 261, 6619-6628.
23. Pritchard, C.G. and DePamphilis, M.L. (1983) J. Biol. Chem.,258, 9801-9809.
24. Maxam, A.M. and Gilbert, W. (1980) Methods Enzymol. 65, 499-560.
25. Sanger, F. and Coulson, A.R. (1978) FEBS Lett. 87, 107-110.
26. Yoda, K. and Okazaki, T. (1983) Nucleic Acids Res. 11, 3433-3450.
27. Pritchard, C.G., Weaver, D.T., Baril, E.F. and DePamphilis, M.L. (1983) J. Biol. Chem. 258, 9810-9819.
28. Tseng, B.Y. and Ahlem, C.N. (1984) Proc. Natl. Acad. Sci. USA 81, 2342-2346.
29. Weaver, D.T. and DePamphilis, M.L. (1984) J. Mol. Biol. 180, 961-986.
30. Faust, E.A., Nagy, R. and Davey, S.K. (1985) Proc. Natl. Acad. Sci. USA 82, 4023-4027.
31. DiFrancesco, R.A. and Lehman, I.R. (1985) J. Biol. Chem. 260, 14764-14770.