Isolation and sequence determination of 5'-terminal oligonucleotide fragments of RNA transcripts synthesized by bacteriophage T3-induced RNA polymerase from T3 DNA

(T3 late RNA synthesis/specific start signals/bovine brain exonuclease/hydroxylapatite chromatography/RNA nucleotidyltransferase)

UMADAS MAITRA*, WARREN JELINEK†, ARTURO YUDELEVICH‡, HEMANTA K. MAJUMDER*, AND ARABINDA GUHA*†‡

*Department of Developmental Biology and Cancer, Albert Einstein College of Medicine, Bronx, New York 10461; and †Rockefeller University, New York, New York 10021

Communicated by Jerard Hurwitz, April 11, 1980

ABSTRACT The nucleotide sequence of the 5'-terminal oligonucleotides produced by pancreatic RNase digestion of bacteriophage T3 RNA polymerase (EC 2.7.7.6) transcripts of T3 DNA has been determined. The sequence determination is based upon a simple isolation procedure for the 5'-terminal oligonucleotides. This procedure involves treatment of pancreatic RNase digests of α-32P-labeled T3 RNA polymerase transcripts with bovine brain exonuclease to remove oligonucleotides with free 5'-hydroxyl termini and then chromatographing the products on hydroxylapatite to resolve the remaining oligonucleotides having 5'-phosphate termini. By application of standard two-dimensional separation and sequence techniques, the major 5'-end sequences deduced were pppGpGpApApCpY (Y = pyrimidine nucleoside) and pppGpGpApApGpCp. In addition, the sequences of other minor 5'-terminal oligonucleotides observed on homochromatograms were also determined. The sequences of these 5'-oligonucleotides were pppGpGpGpApApCpY, pppGpGpGpApApGpCp, pppGpGpApApCp, and pppGpGpCp. These results demonstrate that T3 phase-induced RNA polymerase possesses a high degree of specificity in the initiation of RNA chains.

Bacteriophage T3 RNA polymerase (nucleosidetriphosphate: RNA nucleotidyltransferase, EC 2.7.7.6) is a simple monomeric protein of molecular weight 105,000 (1–4). One striking property of the phage RNA polymerase is its remarkable template specificity. T3 RNA polymerase will transcribe only T3 DNA, and is ineffective with a wide variety of native DNAs tested so far. Even with DNA from the related bacteriophage T7, the rate of RNA synthesis by T3 RNA polymerase is only 5% of that observed with T3 DNA (1–5). These results suggest that there are unique nucleotide sequences or structural characteristics (or both) in T3 DNA that are recognized by the phage enzyme to initiate RNA synthesis. In agreement with this suggestion, it has been observed that, under standard conditions of in vitro transcription using T3 DNA as a template, all T3 RNA chains contain the sequence pppG(Gp)n(Ap)m at the 5' end (7). In addition to this selective initiation, the phage polymerase also terminates RNA synthesis at defined sites on the DNA template with the production of six to eight major transcripts, which vary in size from 0.21 to 6.2 X 106 daltons (8, 9). Sequence analysis of the T3 RNA polymerase products indicates that all (or a majority) of RNA chains synthesized by T3 RNA polymerase terminate in a unique sequence, (Gp)UpUpUpUpUpOH, at their 3' termini (10).

To examine in more detail the specificity with which T3 RNA polymerase initiates transcription on T3 DNA, we have isolated the 5'-terminal oligonucleotide fragments derived from T3 RNA polymerase transcripts and determined their nucleotide sequence. Our results indicate that a majority of the RNA chains initiate transcription with the sequence pppGpGpApApRp (R = purine nucleoside) at the 5' end, indicating a high degree of specificity of initiation.

EXPERIMENTAL PROCEDURES

Materials. T3 RNA polymerase was purified and characterized as described (4). Bovine brain exonuclease, an enzyme specific for cleavage of RNA and oligoribonucleotides having a free 5'-OH group to nucleoside 3'-monophosphates (10) was purified and characterized as described by Guha (11). α-32P-Labeled ribonucleoside triphosphates were obtained from ICN, and [γ-32P]GTP was prepared by the method of Glynn and Chappell (12).

Preparation of T3 RNA Polymerase Transcripts Labeled in Vitro. RNA synthesis was carried out in standard T3 RNA polymerase reaction mixtures as described (10), using [γ-32P]GTP, [α-32P]ATP, or [α-32P]GTP (4–5 X 104 cpm/mmol) as the labeled substrate. Labeled RNA products were isolated from polymerase reactions by a modification of the procedure described by Majumder et al. (10) as follows: After extraction of reaction mixtures with neutralized phenol, the labeled RNA was precipitated with 2.5 vol of ethanol at −20°C for 12 hr. The precipitated RNA was dissolved in 0.5 ml of a buffer containing 10 mM Tris-HCl, pH 7.5, 5 mM EDTA, and 50 mM KCl and chromatographed on a column (0.9 X 50 cm) of Sephadex G-50 equilibrated in the same buffer, to remove unincorporated precursor. The labeled RNA was then precipitated with 2.5 vol of ethanol at −20°C.

Digestion with Pancreatic RNase and with Bovine Brain Exoribonuclease. Samples of the 32P-labeled RNA were digested in 10 mM Tris-HCl, pH 7.5, containing 5 mM EDTA, using a pancreatic RNase-to-RNA substrate ratio of 1:10 (wt/wt). Incubation at 37°C for 45 min gave complete digestion. After digestion with pancreatic RNase, reaction mixtures were adjusted to 50 mM 2-(N-morpholino)ethanesulfonic acid buffer, pH 6.0, and 100 mM NaCl and were treated with purified bovine brain exoribonuclease, using an enzyme-to-substrate ratio of 1:10 (wt/wt). Incubation was at 37°C for 20 min.

Hydroxylapatite Chromatography. The procedure is similar to that described by Grohmann et al. (13). All operations were carried out at 0–5°C unless otherwise indicated. Products obtained after digestion of 32P-labeled RNA with pancreatic RNase followed by brain exoribonuclease were adjusted to pH 6.8 by the

Abbreviations: R, purine nucleoside; Y, pyrimidine nucleoside.
† Permanent address: Departamento de Biología Celular, Universidad Catolica, Santiago, Chile.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U. S. C. §1734 solely to indicate this fact.
addition of 0.5 M K$_2$HPO$_4$. The reaction mixture was diluted with 3 vol of 10 mM potassium phosphate buffer, pH 6.8, and then applied to a column (1-ml bed volume) of hydroxylapatite (grade HT from Bio-Rad) that had been equilibrated with 50 mM potassium phosphate buffer, pH 6.8. The column was then washed successively with 20-ml volumes of potassium phosphate buffer, pH 6.8, at the following concentrations: (i) 50 mM, (ii) 75 mM, (iii) 100 mM, (iv) 500 mM. The flow rate was approximately 10 ml per hour, and 1-ml fractions were collected. The radioactivity content of effluent fractions was measured in 10- to 20-µl aliquots in 10 ml of Aquasol (New England Nuclear) in a liquid scintillation counter.

DEAE-Cellulose Chromatography. The fractions from the 500 mM potassium phosphate wash containing the highest level of radioactivity were pooled, diluted with 10 vol of cold water, and applied to a 0.75-ml bed volume of DEAE-cellulose in a disposable Pasteur pipette column that had been equilibrated with 100 mM triethylammonium bicarbonate, pH 8.0. The column was washed with 10 ml of 100 mM triethylammonium bicarbonate (pH 8.0) and the radioactivity was then eluted with 3 ml of 2 M triethylammonium bicarbonate (pH 8.0). After removal of triethylammonium bicarbonate by repeated evaporation with 2-ml volumes of 1:1 mixture of methanol/ H$_2$O (1:1, wt/wt), the labeled oligonucleotides were dissolved in 10 µl of 5 mM EDTA for sequence determination.

Sequence Determination. Two-dimensional separation of oligonucleotides was carried out as described by Barrall (14). Procedures for recovery of spots after homochromatography on DEAE-cellulose thin-layer plates, redigestion with RNase, and characterization of products were performed as described (14, 15).

RESULTS

Evidence for Separation on Hydroxylapatite Columns of Oligonucleotides Bearing 5'-Triphosphate Termini. Chromatography on hydroxylapatite in potassium phosphate has been used as a simple procedure for rapid and quantitative recovery of all oligonucleotides containing triphosphate end groups (13). Using RNase T1 digests of phage φX174 mRNAs selectively labeled in vitro with [γ-32P]GTP and [γ-32P]ATP and uniformly labeled with [3H]UTP, Grohmann et al. demonstrated (13) that non-5'-terminal oligonucleotides bearing free 5'-hydroxyl groups elute at the beginning of the gradient, while oligonucleotides bearing 5'-phosphate end groups are selectively retarded and principally begin to elute with 0.25 M potassium phosphate buffer. Large nonterminal oligonucleotides (20 nucleotides in length or greater) are, however, also retarded on the column and elute with the 5'-terminal fragments.

We have previously demonstrated that all RNA chains formed in a T3 RNA polymerase reaction directed by native T3 DNA as a template contain the sequence pppGp(Gp)$_m$ - (Ap)$_n$ . . . at the 5' end (7). The presence of a purine-rich sequence at the 5' end prompted us to determine the oligonucleotide sequence at the 5' end of a pancreatic RNase digest of T3 RNA polymerase transcripts. For this purpose, we initially determined whether hydroxylapatite chromatography can be used to purify the 5'-terminal fragments from the total RNAse digest (Fig. 1). In vitro T3 RNA polymerase transcripts, selectively labeled at the 5' end with [γ-32P]GTP and uniformly labeled with [5-3H]UTP, were mixed, digested with pancreatic RNase, and chromatographed on hydroxylapatite, using increasing concentrations of potassium phosphate for elution of oligonucleotides (Fig. 1). Comparison of [3H] and [32P] profiles showed that a large majority of non-5'-terminal oligonucleotides (containing a free 5'-hydroxyl group) eluted at lower phosphate concentrations, while oligonucleotides containing 5'-triphosphate ends were selectively retarded and eluted at 0.5 M potassium phosphate. A small peak of [3H]-labeled oligonucleotides also eluted at this higher phosphate concentration; this peak presumably represents 5'-terminal fragments and non-5'-terminal oligonucleotides of chain length about 20 or higher.

Characterization of [γ-32P]GTP-Labeled 5'-Oligonucleotides. In vitro T3 RNA polymerase products, labeled at the 5' end with [γ-32P]GTP, were prepared and subjected to complete digestion with pancreatic RNase. The resulting 5'-terminal oligonucleotides were isolated by hydroxylapatite chromatography followed by DEAE-cellulose chromatography. Fig. 2 is an autoradiogram of a two-dimensional fingerprint of oligonucleotides produced by digestion of [γ-32P]GTP-labeled T3 RNAs with pancreatic RNase. Six major radioactive spots (labeled 1–6) were detected on the homochromatogram; they represent the major 5'-start sequences of T3 RNA polymerase on T3 DNA. Among these six spots, the oligonucleotide fragments represented by spots 2 and 3 constituted >80% of the radioactivity present in the homochromatogram. These two 5' fragments thus represented the major 5'-start sequences of T3 RNA polymerase transcripts; the remaining spots presumably represented the minor start sites for T3 RNA polymerase.

Isolation and Sequence Determination of 5'-Terminal Oligonucleotides Derived from α-32P-Labeled T3 RNA Polymerase Products. Preliminary fingerprint analysis of the 5'-terminal oligonucleotides isolated by the hydroxylapatite technique from pancreatic RNase digests of [α-32P]ATP and [α-32P]GTP-labeled T3 RNA polymerase products indicated the presence of many labeled oligonucleotides, which obscured the positions of the 5'-terminal oligonucleotides on the homochromatograms in a two-dimensional sequence analysis (data not shown). Presumably, large nonterminal oligonucleotides of chain lengths 20 nucleotides or higher that are retained by the hydroxylapatite column at 50–100 mM potassium phos-
phate buffer (13) also eluted with the 5'-terminal oligonucleotides in hydroxylapatite chromatography and were not subsequently separated in homochromatography.

To remove these large nonterminal oligonucleotides, the pancreatic RNase digests of \( \alpha-^{32}P \)-labeled T3 RNA polymerase products were treated with bovine brain exonuclease, an enzyme that specifically digests to 3'-mononucleotides any oligonucleotides containing 5'-hydroxyl group; 5'-phosphorylated oligonucleotides are resistant to such digestion (11). After digestion with exonuclease followed by hydroxylapatite chromatography, the 5'-phosphorylated terminal oligonucleotide fraction would be expected to be free of any nonterminal oligonucleotides and mononucleotides.

Thus to isolate and determine the sequence of 5'-terminal oligonucleotides from \( \alpha-^{32}P \)-labeled T3 RNA polymerase products, T3 RNA polymerase transcription products were synthesized in three separate reaction mixtures, each containing a different \( ^{32}P \)-labeled nucleotide triphosphate—\( \gamma-^{32}P \)GTP, \( \alpha-^{32}P \)ATP, or \( \alpha-^{32}P \)GTP. The isolated RNA products from each reaction were subjected to complete digestion with pancreatic RNase followed by treatment with bovine brain exonuclease as described in Experimental Procedures. The resulting 5'-terminal oligonucleotide fragments were purified by hydroxylapatite column chromatography followed by DEAE-cellulose chromatography.

Fig. 2 is an autoradiogram of two-dimensional fingerprints of 5'-oligonucleotides derived from RNA products labeled with \( \alpha-^{32}P \)ATP and \( \alpha-^{32}P \)GTP, respectively. The fingerprint for \( \gamma-^{32}P \)GTP-labeled 5'-oligonucleotides was similar to that in Fig. 2 and was repeated to demonstrate that bovine brain exonuclease treatment of the pancreatic RNase digest caused no alteration in the integrity of the \( \gamma-^{32}P \)-labeled pancreatic RNase oligonucleotides (data not shown).

A comparison of Fig. 3 A and B with Fig. 2 shows that we have isolated all six 5'-terminal oligonucleotides (labeled 1–6 in both figures) and several other minor spots. The existence of a doublet in each spot is due to the presence of ppppGp at the 5' end of the faster-moving oligonucleotides and ppGp at the end of slower-moving fragments. This was shown by eluting each spot and subjecting it to complete digestion with T1 RNase. Both labeled ppGp and ppGp were identified by electrophoresis on DEAE-paper. It should be noted that, as with

\[
\gamma-^{32}P \text{-labeled oligonucleotides (see Fig. 2), spots 2 and 3 were again found to be the major 5'-oligonucleotide fragments observed on the homochromatogram. These two oligonucleotides thus represented the major 5'-end sequences.}
\]

Nucleotide Sequences of the Major GTP-Terminated Oligonucleotides. In order to determine the nucleotide sequences of the major 5'-oligonucleotide fragments, all spots corresponding to those in the homochromatograms presented in Fig. 3 A and B were individually eluted, redigested with RNase T1, and analyzed on a DEAE-paper at pH 3.5 (14, 15). The digestion products derived from each eluted spot were ppGp, ppGp, nucleoside monophosphates, and short oligonucleotides. These short oligonucleotides were individually eluted from the DEAE-paper and redigested to mononucleotides by alkaline hydrolysis, and their nucleotide compositions were determined (Table 1). The data from the table and the known cleavage specificity of RNase T1 enable us to determine the sequences of the major ppGp . . . oligonucleotides.

**DISCUSSION**

Determination of the 5'-terminal sequences of messenger RNAs is of importance in studies on the control of transcription, particularly with regard to the specificity and efficiency of promoter utilization by RNA polymerase. Because of the relatively large size of T3 RNA polymerase transcripts, it is difficult to determine the 5'-oligonucleotide sequence from a pancreatic RNase digest of total T3 RNA polymerase products. The formation of many non-5'-terminal oligonucleotides obscures the positions of the 5'-oligonucleotides on the homochromatograms in a standard two-dimensional sequence analysis. Following the observation of Grohmann et al. (13), we have employed chromatography on hydroxylapatite as a simple procedure for the isolation of oligonucleotides bearing 5'-triphosphate and 5'-diphosphate groups at the 5' end. We have utilized the bovine brain exonucleases to remove the 5'-OH-terminated oligonucleotides of chain lengths 20 or higher, which heavily contaminate the 5'-triphosphate-terminated oligonucleotides. The
5'-tri- and diphosphate-terminated oligonucleotides are resistant to attack by this enzyme (11). The results presented in this paper show that there are two major initiation sequences, pppGpGpApGpApGpApY ... and pppGpGpApGpApGpApCp ..., with which T3 RNA polymerase initiates RNA chains. In addition to these two major 5'-end sequences, four other minor 5'-oligonucleotide fragments have been observed on the homochromatograms in the two-dimensional analysis. These oligonucleotides also contain pppGpGpAp ... at the 5' end. These results indicate that this phage polymerase, unlike the more complex *Escherichia coli* RNA polymerase, is geared to a unique starting signal, which probably determines the strict template specificity of T3 RNA polymerase towards T3 DNA. Because there are at least eight promoters per T3 genome (8, 9) this sequence must be repeated several times in T3 DNA. It is likely that the two major start sequences correspond to promoters of the class III late genes, which code for phage structural proteins and whose transcription accounts for most of the phage-specific RNAs in *vivo* and *in vitro* (16, 17), whereas the minor sequences observed here are start signals for the class II late mRNAs, which code for enzymes of T3 DNA metabolism and are synthesized with low efficiency (16, 17).

Recently, a number of laboratories have determined the DNA sequences of several T7 RNA polymerase promoters on T7 DNA (18–21). They have observed that all contain a nearly identical 23-base-pair sequence that includes the 17 bases prior to and the first 6 bases in the transcribed region. Because the major T3 RNA polymerase promoters of T3 DNA are not recognized by T7 RNA polymerase and vice versa (1, 2, 4, 9, 22), it will be of considerable interest to determine the major T3 RNA polymerase promoters on T3 DNA and compare the sequences of these promoters with those recognized by T7 RNA polymerase on T7 DNA. The observation that both the major T7 and T3 RNA polymerase promoters have similar start signals, pppGpGpApGp (refs. 19–21 and this paper) points to flanking sequences upstream or downstream (or both) to the initiation sequences as the factors which influence the efficiency of two phage polymerases. Further comparative studies of these promoters for the two monomeric phage polymerases should give important information regarding the mechanism of promoter recognition by RNA polymerases.

This work was supported by grants from the National Institutes of Health and the National Science Foundation.