Volume 15 Number 21 1987 Nucleic Acids Research

# Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates

John F.Milligan, Duncan R.Groebe, Gary W.Witherell and Olke C.Uhlenbeck

Department of Chemistry and Biochemistry, University of Colorado, Boulder, CO 80309-0215, USA

Received July 13, 1987; Revised and Accepted October 1, 1987

## ABSTRACT

A method is described to synthesize small RNAs of defined length and sequence using T7 RNA polymerase and templates of synthetic DNA which contain the T7 promoter. Partially single stranded templates which are base paired only in the -17 to +1 promoter region are just as active in transcription as linear plasmid DNA. Runoff transcripts initiate at a unique, predictable position, but may have one nucleotide more or less on the 3' terminus. In addition to the full length products, the reactions also yield a large amount of smaller oligoribonucleotides in the range from 2 to 6 nucleotides which appear to be the result of abortive initiation events. Variants in the +1 to +6 region of the promoter are transcribed with reduced efficiency but increase the variety of RNAs which can be made. Transcription reaction conditions have been optimized to allow the synthesis of milligram amounts of virtually any RNA from 12 to 35 nucleotides in length.

## INTRODUCTION

The recent developments in *in vitro* transcription methods using the highly active phage polymerases represent a major advance in RNA synthesis (1). The desired sequence is cloned into one of several available phage promoter vectors and the plasmid DNA is linearized with a restriction enzyme. When the reaction conditions are optimized, runoff transcription from these templates yields hundreds or even thousands of moles of RNA per mole of DNA (2). Since large amounts of T7, T3, and SP6 RNA polymerases are available from overproducing strains (3-5), multimilligram amounts of virtually any RNA sequence can be prepared.

The use of synthetic DNA templates for the synthesis of relatively short RNA fragments by T7 polymerase would be expected to have several advantages. First, the unwanted 5' flanking sequences that are often introduced as part of the cloning procedure can be eliminated. Second, the sequence of the 3' end of the transcript is not limited by the need for a restriction site to linearize the template (1). Finally, the time consuming cloning, sequencing, and plasmid DNA preparation can be avoided entirely. During our investi-

© IRL Press Limited, Oxford, England.

gations into the activity of synthetic DNA templates in transcription reactions, it became clear that only the promoter portion of the template had to be double stranded. This results in an additional, unexpected advantage of requiring the synthesis of only a single DNA fragment for each different RNA oligonucleotide desired.

In this paper in vitro transcription of synthetic DNA templates by T7 RNA polymerase is characterized in detail. The products of the reaction are determined, the sequence and the length dependence of the templates are established, and the optimal reaction conditions are given for the synthesis of RNA fragments from 12 to 35 nucleotides in milligram amounts.

## MATERIALS AND METHODS

Deoxyoligonucleotides were synthesized on an Applied Biosystems 380B DNA synthesizer and purified by preparative 20% polyacrylamide gel electrophoresis (6). The purified DNA was stored in 10 mM Tris-HCl, pH 7.0 at -20°C. Templates were prepared by heating the two DNA strands together to 65°C for 3 minutes and then cooling the DNA on ice. A complete list of deoxyoligonucleotides used in this study is given in Table 1. The top strands (T-N) contain the promoter region but are not templates. Bottom strands (B-N) contain both the promoter and template sequences and are given 3' to 5' in order to clarify pairing with the top strands. Residue +1 indicates the expected transcription start site.

Nucleoside triphosphates were purchased from Pharmacia (P-L) Biochemicals, dissolved in water to a concentration of 25 mM in each NTP, adjusted to pH 8.1 so as not to affect the final pH of the transcription reaction, and stored at -20°C. [$\alpha$-$^{32}$P] NTPs, in tricine, were purchased from New England Nuclear Research Products. [$\gamma$-$^{32}$P] ATP and [$\gamma$-$^{32}$P] GTP were prepared from $^{32}PO_4^{3-}$ by the method of Johnson and Walseth (7). [5'-$^{32}$P] pCp was prepared from [$\gamma$-$^{32}$P] ATP, and T4 kinase by the method of England and Uhlenbeck (8).

T7 RNA polymerase was isolated from E. coli strain BL21, containing the plasmid pAR1219, according to the procedure of Davanloo et al. (3). Purified T7 RNA polymerase was assayed on a linearized plasmid containing the wild type T7 promoter by the method of Chamberlin et al. (9) and found to contain approximately 300,000 units per mg of protein.

The standard transcription reactions were run in 40 mM Tris-HCl (pH 8.1 at 37°C), 1 mM spermidine, 5 mM dithiothreitol, 50 ug/ml bovine serum albumin, 0.01% (v/v) Triton X-100 and 80 mg/ml polyethylene glycol (8000 MW). The concentrations of NTPs, $MgCl_2$, DNA template, and T7 RNA polymerase are varied

Table 1: Deoxyoligonucleotides used for transcription.

```
TOP STRANDS

    5'                    +1                  3'
T-1  AATTTAATACGACTCACTATAGGGAGGAGGATTACCCCTCG
T-2      TAATACGACTCACTATAG

BOTTOM STRANDS
        3'                +1                        5'
B-1      ATTATGCTGAGTGATATCCCTCCTCCTAATGGGGAGCTTAA
B-2              "        CCTGTACTCCTAATGGGTACA
B-3              "        CCCTGACTGCTAATGCGTCAGATA
B-4              "        CCCTCTCTCCTAGTGGGAGAG
B-5              "        CCCTAGGGGGGCCTAGGT
B-6              "        CCCGTATCCTAATGGGTACGATA
B-7              "        CCGTTGTACTCCTTATGGGTACA
B-8              "        CCTGTACTCCTAATGGGTACA
B-9              "        CTCTGACTGCTAATGCGTGAGATA
B-10             "        TCCTGACTGCTAATGCGTCAGATA
B-11             "        CCCTGACTCCTAATGGGTCAGATA
B-12             "        CCCTGAGTGCTAATGCCTCAGATA
B-13             "        CCCTGACTACTAATGTGTCAGATA
B-14             "        CCCTATGGGGGGCATAGGC
B-15             "        CCGAGCTGACTACTCCGCG
B-16             "        CGCGGCTTTGGGCCAGAGCTCG
B-17             "        CCGTGGCTTTGTGGTACAGACAGG
B-18             "        CCCGAGACTCTCCGGCAATCCGGCTTTGTTG
B-19             "        CCTGCTTTGGTCCAGTAG
B-20             "        CCCTTCAGTCGG
B-21             "        CCTGAGAGAATTCCATC
B-22             "        CCGGACATTTTT
B-23             "        CCGCGTATTTTT
B-24             "        CCCGAGATTTTT
B-25             "        CCTAGTGGCAAA
B-26             "        GCTAGTGCCAAA
B-27             "        CGTAGTCGCAAA
B-28             "        CCTGTACTCCTA
B-29      TTATGCTGAGTGATATCCCTGACTGCTAATGCGTCAGATA
B-30    TATTATGCTGAGTGATATCCCTGACTGCTAATGCGTCAGATA
```

and are as indicated in each experiment. Reaction mixtures were incubated at 37°C for the indicated times.

Following the transcription reactions the RNA was heated in loading buffer containing 7 M urea and dyes for 2 minutes at 65°C immediately prior to loading on an analytical 20% denaturing polyacrylamide gel. The products were separated on a 46 cm vertical gel apparatus, the bands located by autoradiography, cut out and Cerenkov counted in a Beckman LS 3801 Liquid Scintillation counter to determine yields.

The sequence of the transcription products was determined by several methods. First, transcripts which had been labeled by one of each of the four different [$\alpha$-$^{32}$P] NTPs were subjected to complete enzymatic hydrolysis with

RNase $T_1$, RNase A, or RNases $T_1$, $T_2$, and A. The smaller digestion products were separated and identified by two-dimensional thin layer chromatography (2D-TLC) on Kodak cellulose plates which contain a fluorescent indicator (10). Second, the 3' end terminal nucleotide was determined by the addition of [5'-$^{32}$P] pCp to the free 3' hydroxyl using T4 RNA ligase (8) followed by total enzymatic hydrolysis using RNases $T_1$, $T_2$, and A, with product analysis by 2-D TLC (10). Finally, transcripts were sequenced in a method similar to that of Randerath et al. (11) in which a purified RNA containing a 5' triphosphate is partially hydrolyzed at 85°C for 10 minutes in $H_2O$. Since the 5' terminus is a triphosphate only those molecules which have been hydrolyzed may be labeled on their 5' termini by T4 polynucleotide kinase and [$\gamma$-$^{32}$P] ATP. The labeled RNAs are separated on a sequencing gel, cut out, soaked in 0.5 M NaOAc (pH 5.2), and ethanol precipitated in the presence of carrier tRNA. The recovered RNA is then digested with RNase $P_1$ and the 5' termini determined by one dimensional TLC on PEI cellulose 300 plates (Macherey-Nagel) in 0.25 M NaFormate at pH 3.5.

## RESULTS

### Transcription Using Synthetic DNA

Three different types of templates were compared for their activity with T7 RNA polymerase (Fig. 1A). Template 1 is a chemically synthesized Eco R1 restriction fragment, made by annealing strand T-1 and B-1, which contains 17 base pairs of the T7 promoter and a 24 base pair downstream region corresponding to a variant of the R17 coat protein binding sequence (12). Template 2 was created by cloning template 1 into the Eco R1 site of pUC13. Since the upstream site is not an Eco R1 restriction site, cleavage of the cloned DNA with Eco R1 results in template 2 which has the same runoff sequence as template 1 but has an additional 3 kb of flanking DNA upstream of the transcription start site. Template 3 was created by annealing the bottom strand, B-1, with an 18 nt top strand, T-2, to create a fragment which is base paired only from -17 to +1 of the promoter and has the remaining 23 nucleotides of B-1 as a 5' overhang. All three templates contain the same sequences downstream of the promoter and should be capable of producing an identical 24 nt runoff RNA.

*In vitro* transcription of the three templates under identical conditions resulted in virtually the same pattern of oligoribonucleotide products (Fig. 1B). The two longest product bands ($P_L$ and $P_U$) appear in about equivalent amounts and both are approximately the correct length to be runoff tran-

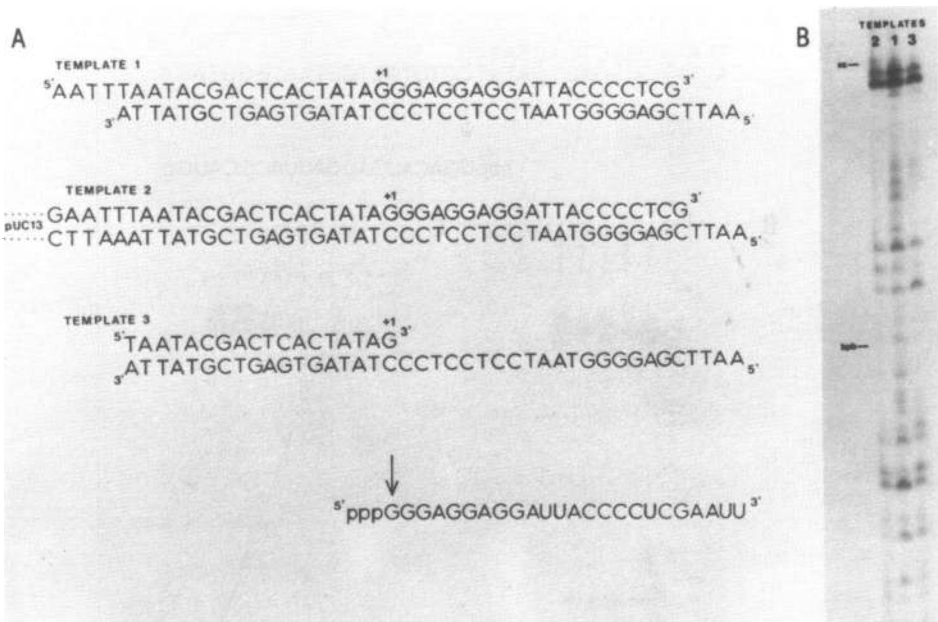


FIGURE 1: A.) Three templates designed to produce the same runoff transcript. B.) Autoradiogram of a 20% denaturing polyacrylamide gel analyzing the transcription products of the three templates. Each reaction contained 1.5 mM of each NTP (0.5 uCi/mmole [$\alpha$-$^{32}$P] CTP), 9 mM $MgCl_2$, 50 nM DNA, 30 U/ul T7 RNA polymerase in transcription buffer and was incubated 2 h at 37°C.

scripts. In addition, all three templates produce numerous smaller products. Relatively large amounts of RNA are produced in these reactions. Each reaction contained 0.25 pmoles of DNA template and produced approximately 15 pmoles each of $P_L$ and $P_U$. Thus, multiple rounds of transcription make T7 RNA polymerase an effective means of producing short RNA fragments.

As a preliminary test for the purity of the oligonucleotides and the fidelity of the polymerase, bands $P_L$ and $P_U$ were isolated from the gel and tested for their ability to bind R17 coat protein. The products from all three templates were found to bind R17 coat protein with equal efficiency and with a Kd similar to that of an oligoribonucleotide synthesized by T4 RNA ligase methods (data not shown). Since RNA binding by the R17 coat protein is very sensitive to changes made in the nucleotide sequence (12), this is a good preliminary indication that the oligoribonucleotides made by T7 RNA polymerase have the correct sequence.

Several conclusions can be made from these experiments. First, it

A

+1
TAATACGACTCACTATAG
ATTATGCTGAGTGATATCCTGTACTCCTAATGGGTACA

pppGGACAUGAGGAUUACCCAUGU

B

C

0 5 10 30 60 120 240

FIGURE 2: A.) Template prepared with T-2 and B-2. B.) Transcription of five identical reactions containing different $^{32}P$ labeled nucleotides. C.) Kinetics of [α-$^{32}P$] GTP labeled reaction. Each reaction contained 0.75 mM of each NTP, 6 mM $MgCl_2$, 50 nM DNA, 30 U/ul T7 RNA polymerase in transcription buffer and was incubated 2 h at 37°C, reactions were analyzed on 20% denaturing polyacrylamide gels.

appears that synthetic DNA provides as good a template for *in vitro* transcription as does bacterial plasmid DNA. The possible small amounts of residual blocking groups do not appear to alter the amounts or types of products. Second, under the reaction conditions used, the nucleotides upstream of position -17 do not contribute to the efficiency of T7 transcription. Kinetic analysis of similar templates also indicates that nucleotides upstream of -17 are not necessary for *in vitro* transcription (13). Although the consensus promoter sequence ends at -17 (14), footprinting experiments (15) suggested protein-DNA contacts as far back as -21. Third, unlike *E.*

Table 2: Analysis of RNase digestion products of $P_L$ and $P_U$.

| [α-$^{32}$P] NTP | RNase A Expected | | RNase A Identified | | RNase T$_1$ Expected | | RNase T$_1$ Identified | |
|---|---|---|---|---|---|---|---|---|
| ATP | Up | (1) | Up | 1.0 | Gp | (2) | Gp | 2.0 |
| | Cp | (1) | Cp | 1.1 | ACAUGp | (2) | Long | 3.8 |
| | pppGGACp | (1) | Long | 3.2 | AUUACCCAUGp | (2) | | |
| | GAGGAUp | (2) | | | | | | |
| | | | [GUp] | 0.2 | | | | |
| CTP | Cp | (1) | Cp | 1.0 | ACAUGp | (1) | Long | all |
| | ACp | (2) | ACp | 1.5 | AUUACCCAUGp | (3) | | |
| | pppGGACp | (1) | Long | 0.9 | | | | |
| | | | [GUp] | 0.3 | | | | |
| GTP | AUp | (2) | AUp | 2.0 | pppGp | (2) | pppGp | 2.4 |
| | pppGGACp | (2) | Long | 3.6 | AGp | (2) | AGp | 2.0 |
| | GAGGAUp | (2) | | | ACAUGp | (1) | Long | 1.8 |
| | | | | | AUUACCCAUGp | (1) | | |
| UTP | AUp | (2) | AUp | 2.2 | AACAUGp | (1) | Long | all |
| | GGAGGAUp | (2) | Long | 1.9 | AUUACCCAUGp | (4) | | |
| | GU | (1) | (GU) | 1.0 | | | | |
| | | | [GUp] | 1.3 | | | | |

Products identified by 2-D TLC after digesting the sample with ribonuclease. Numbers in parenthesis indicate the relative number of radiolabeled phosphates expected for a 24 nucleotide runoff transcript. The numbers after the identified products indicate the experimental values. (GU) is unique to $P_L$. [GUp] is unique to $P_U$. The longer digestion products (Long) were not identified.

coli RNA polymerase (16) the presence of long, non-specific 5' flanking DNA does not lead to significant rate enhancements. Perhaps the extreme terminal location of the promoter or the high enzyme concentrations used here obscure this effect. Finally, the non-template strand does not appear to be needed for efficient transcription provided that the promoter region is double stranded.

Product Analysis

Since the type 3 templates have the obvious advantage of only requiring a single new DNA fragment for each new RNA desired, we have analyzed the transcription products of one of these templates in greater detail. The template composed of T-2 and B-2 was transcribed in five separate reactions containing one each of the four [α-$^{32}$P] NTPs or [γ-$^{32}$P] GTP. The products of a time course of the [α-$^{32}$P] GTP reaction and a 2 hr incubation of all five reactions are shown in Fig. 2. Reactions where the DNA template was omitted or [γ-$^{32}$P] ATP was used gave no radioactive products. Each of the five

Table 3: 3' end analysis of $P_L$ and $P_U$.

| | Percentage of nucleoside | | | |
|---|---|---|---|---|
| Band | A | G | C | U |
| $P_L$ | 1.6 | 0.9 | 1.7 | 95.8 |
| $P_U$ | 42.6 | 1.9 | 46.2 | 9.3 |

reactions produced the same pair of longer products, $P_L$ and $P_U$, which have an approximate length consistent with a transcript starting at +1 and running until the end of the template DNA. In addition, a series of shorter products, $A_2$ - $A_6$, are produced at approximately a 10-fold greater level than $P_L$ or $P_U$. These shorter products appear with similar kinetics and show a labelling pattern that is approximately consistent with the sequence of the template and the length of the fragment. For example, $A_6$ is the smallest product labelled by [$\alpha$-$^{32}$P] UTP and the first product labelled in the [$\alpha$-$^{32}$P] UTP transcript is at nucleotide 6.

The method of Randerath et al. (11) was used to show that the sequence of residues 2-16 of $P_L$ is consistent with Fig. 2 (data not shown). Since no heterogeneity is seen at any positions, $P_L$ is not a mixture of transcripts resulting from initiating and terminating at different positions.

In order to confirm the sequence, each radiolabeled species of $P_L$ and $P_U$ generated in Fig. 2 was subjected to total digestion with RNases A or $T_1$. The smaller digestion products were separated and identified by 2D TLC. As summarized in Table 2, the products of both bands were identical and consistent with the sequence predicted in Fig. 2 with one exception. In the ribonuclease A digestions of [$\alpha$-$^{32}$P] UTP labeled molecules, $P_L$ gave GpU and $P_U$ gave GpUp. Since GpU is the expected 3' terminal dinucleotide in the molecule, the data indicate that $P_L$ has the expected sequence for a runoff transcript and that $P_U$ has an additional nucleotide on the 3' termini. Since small amounts of GpUp can also be seen in RNase A digestions of [$\alpha$-$^{32}$P] CTP and [$\alpha$-$^{32}$P] ATP, it appears that $P_U$ may have 3' termini of C and A.

In order to determine the 3' termini of $P_L$ and $P_U$, both molecules were 3' end labeled with [5'-$^{32}$P] pCp and RNA ligase and subjected to total digestion with ribonuclease $T_2$. As shown in Table 3, the 3' terminal nucleoside of $P_L$ is U as expected, while the 3' nucleoside of $P_U$ is a mixture of C, A, and some U. Thus, it appears that T7 polymerase can add a non-template encoded nucleotide at the 3' end of the transcripts, as has been reported for SP6 RNA polymerase (1).

Table 4: Nearest neighbor analysis of abortive initiation products of template T-2/B-2.

| BAND | $[\alpha^{32}P]$ NTP GTP | ATP | CTP | UTP | SEQUENCE |
|---|---|---|---|---|---|
| $A_2$ | pppGp | pppGp | pppGp | pppGp | pppGN |
| $A_3$ | pppGp | Gp | ---- | ---- | pppGGA |
| $A_4$ | pppGp | Gp | Ap | ---- | pppGGAC |
| $A_5$ | pppGp | Gp,Cp | Ap | ---- | pppGGACA |
| $A_6$ | pppGp | Gp,Cp | Ap | Ap | pppGGACAU |

Since a majority of the nucleotides are incorporated into the smaller products (Fig. 2B), their identity was determined. Each band $A_2$-$A_6$ was isolated after labeling with each of the four $[\alpha\text{-}^{32}P]$ NTPs. They were subjected to total enzymatic hydrolysis with RNase $T_2$ and the products identified by 2-D TLC (Table 4). Band $A_2$ shows a mobility change upon digestion, releasing labeled pppGp. Since pppGp is released regardless of the $[\alpha\text{-}^{32}P]$ NTP used to label the reaction, the product at $A_2$ is pppGN, where N is G, A, C, or U. The absence of any other digestion products indicates that $A_2$ is pppGN and not a degradation product containing a 3' phosphate. Similar nearest neighbor analyses of $A_3$-$A_6$ indicate that they have lengths and sequences consistent with a transcript which had been initiated at +1, transcribed three to six nucleotides, and then prematurely terminated. Similar abortive initiation products have been observed with E. coli RNA polymerase and many promoters (17,18). It appears that the E. coli enzyme can initiate transcription and elongate a few nucleotides without dissociating from the promoter. If the promoter clearance step is slow, the short oligomers are released and reinitiation can occur many times before a runoff transcript is produced (19).

Promoter Size

If the ratio of abortive initiation products to full length product is governed by the effectiveness of the promoter clearance step, then decreasing the stability of the initiation complex may increase the propensity of the polymerase to elongate. Thus, the effect of removing nucleotides from the promoter specific portion of the T-2/B-3 template upon the amount and type of products was examined. Figure 3 shows the effect of reducing the size of the top strand upon transcription yields. As the top strand is shortened from its

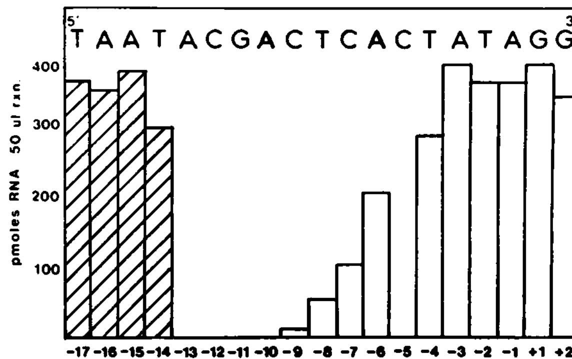


FIGURE 3: Effect of top strand length on transcription efficiency. Open bars indicate transcription yield with a top strand from -17 to the position indicated. Shaded bars indicate the yield with a top strand from +1 to the position indicated. All reactions used B-3 and contained 0.75 mM in each NTP, 6 mM $MgCl_2$, 50 mM DNA template, 15 U/ul T7 polymerase in transcription buffer and were incubated 1 h at 37°C.

3' end, neither the site of initiation nor the amount of products is significantly altered until nucleotide -3 is removed. Further elimination of nucleotides causes a decrease in the amount of full length product, but not any change in the point of initiation, until the nucleotide at -8 is eliminated at which point there are no detectable full length products. A similar, less complete, experiment removing nucleotides from the 5' terminus of the top strand shows no change in the transcription products or yields until nucleotide -14 is removed. Thus, the top strand does not have to cover the entire T7 consensus promoter, but can be up to 3 nucleotides shorter on either the 3' or 5' terminus.

The effect of changing the length of the 3' end of the bottom strand was examined with B-29 and B-30 which are identical to B-3, but have, respectively, one less nucleotide and an additional A residue. The T-2/B-30 template transcribes equally well as the T-2/B-3 template, while the T-2/B-29 templates gives at least a 50% reduction in yield. Thus, as implied by the experiment in Fig. 1, elongating the template strand beyond -17 has no effect on transcription. In addition, it is not possible to reduce the size of the template strand at the 3' end without reducing the reaction yield.

The results of the promoter truncation experiments indicate that the T7 RNA polymerase does not appear to require the DNA in the non-template strand in the region -17 to -14 and -3 to +6, since removing these nucleotides has little effect on the transcription reaction. A recent kinetic analysis using

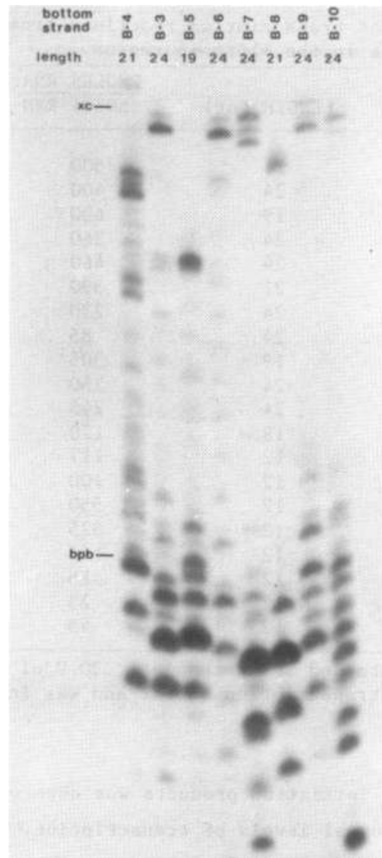


FIGURE 4: Transcription from eight different templates. Each reaction contained 2.5 mM in each NTP (1 uCi/mmole [$\alpha$-$^{32}$P] GTP), 15 mM $MgCl_2$, 50 nM DNA template, 30 U/ul T7 RNA polymerase in transcription buffer and was incubated 3 h at 37°C.

similar transcription templates also indicate that the region from -14 to -17 can be deleted without affecting the point of initiation, although both the $K_m$ and $k_{cat}$ are affected (13). Surprisingly, the absence of the base pairs appears to have no effect on the interaction between the enzyme and the template strand which extend from -17 onward. Recent experiments with the SP6 RNA polymerase indicate that the base pairs at positions -1 and -2 can be changed without altering the amount of product or the point of initiation (20), indicating that the very similar SP6 polymerase also does not recognize its promoter in this region.

In all the above experiments, no improvement in the relative amounts of

Table 5: Yields of transcription reactions from templates which are variants in the +1 to +6 region.

| BOTTOM STRAND | LENGTH (nt) | PMOLES RNA / 50 ul RXN | RNA SEQUENCE +1 to +6 |
|---|---|---|---|
| B- 4 | 21 | 500 | GGGAGA |
| B- 3 | 24 | 400 | GGGACU |
| B- 5 | 19 | 600 | GGGAUC |
| B- 6 | 24 | 360 | GGGCAU |
| B- 7 | 24 | 460 | GGCAAC |
| B- 8 | 21 | 390 | GGACAU |
| B- 9 | 24 | 230 | GAGACU |
| B-10 | 24 | 85 | AGGACU |
| B-15 | 19 | 305 | GGCUCG |
| B-16 | 24 | 250 | GCGCCG |
| B-17 | 24 | 245 | GGCACC |
| B-19 | 18 | 170 | GGACGA |
| B-20 | 12 | 117 | GGGAAG |
| B-22 | 12 | 100 | GGCCUG |
| B-23 | 12 | 550 | GGCGCU |
| B-24 | 12 | 325 | GGGCUC |
| B-25 | 12 | 143 | GGAUCA |
| B-26 | 12 | 15 | CGAUCA |
| B-27 | 12 | 73 | GCAUCA |
| B-28 | 12 | 45 | GGACAU |

Each reaction contained 50 nM template, 30 U/ul polymerase, 4.0 mM of each NTP, 20 mM $MgCl_2$, transcription buffer and was incubated for 4 hrs. at 37°C.

full length to abortive initiation products was observed. For all the templates that showed normal levels of transcription, the amount and pattern of abortive initiation products were unchanged. For those truncated templates which showed a reduced level of transcription, the relative amount of abortive products was actually greater. Thus, changing the promoter size and sequence was not successful in reducing the amount of NTP incorporated into the shorter products.

The +1 to +6 Region

Since the T7 promoters are highly conserved from -17 to +6 (14) all natural T7 transcripts initiate with similar sequences. Since it is desirable to prepare RNAs with a variety of sequences near the 5' terminus, it was important to determine the range of permissible sequences in the +1 to +6 region of the promoter. Therefore a series of bottom strands was synthesized which contain the highly active T7 class III promoter in the -17 to -1 region but with differing sequences +1 to +6, as well as some differing downstream regions. After annealing with T-2, each was tested in an *in vitro* transcription reaction under identical reaction conditions. As can be seen in

Fig. 4 for each of the eight different bottom strands, substantial amounts of RNA are produced in each reaction. In nearly every case, two or more bands were obtained at the approximate length expected for a full length transcription product. As before, the extra bands correspond to non-template encoded nucleotides added to the end of the expected full length transcript. It is also interesting to note that the distribution and amounts of the abortive initiation products vary considerably among the different templates. To a certain extent this is a consequence of the differing nucleotide content in the +1 to +6 region resulting in an altered labeling pattern, but it is also clear that the amount of abortive initiation varies considerably as a consequence of the +1 to +6 sequence.

The yields of full length transcripts obtained from reactions with large number of different bottom strands are shown in Table 5. While it is clear that all reactions produce useful amounts of RNA fragments, the yields vary substantially. Attempting to relate the yields to the +1 to +6 sequence is complicated by several factors. First, as will be discussed in the following section, each template may have a different set of optimal reaction conditions and the reaction conditions for each template in Table 5 were identical, not optimal, for each. Second, in order to produce useful RNA fragments, the sequences downstream of +6 were quite different in each molecule which might also influence the yield. Finally, the lengths of the transcripts varied. However, several interesting conclusions can be drawn from the survey of the transcription properties of a large number of bottom strands.

Most of the shorter templates produce lower molar yields of transcription products. Much higher concentrations of DNA (0.5 - 1 uM) are required for optimal activity of the shorter templates, suggesting that their interaction with T7 polymerase is altered in some fashion. Perhaps the competition for the promoter by the polymerases is hampered if an incoming polymerase cannot initiate until the previous polymerase has already dissociated from the template.

By comparing templates of approximately the same length, one can conclude that the sequence beyond +2 has little effect upon the yield of the reaction. However, substitutions at +2 can reduce the yield of transcript. Substituting $C_{+2}$ with a G or an T reduces the amount of product approximately 2-fold (compare B-27 with B-25, and B-9 with B-3). It is interesting to note that template B-9 shows a greater amount of [$\alpha$-$^{32}$P] GTP label incorporated into the smaller products versus template B-3, which resembles a wild-type promoter. The relative importance of the position +2 over positions +3 to +6 has also

been demonstrated for the SP6 RNA polymerase-promoter interaction (20).

Finally, two bottom strands were synthesized to test the ability of T7 RNA polymerase to initiate with other nucleotides at position +1. B-10 differs from B-3 by a $C_{+1}$ to an T substitution and produces about 20% as much transcript of identical length. Since the transcript could be labeled with [$\gamma$-$^{32}$P] ATP but not [$\gamma^{32}$P] GTP, it is clear that initiation is occurring at the +1 position with an A residue. This was not unexpected since one of the T7 class II promoters also initiates with an A residue (14). More surprising is that B-26 with a G at +1 also produces a product, although the yield is reduced 10-fold with respect to B-25 which has a C at +1 but is otherwise identical. The 5' nucleotide was confirmed to be a C by removal of the 5' triphosphate, 5' $^{32}$P labeling with polynucleotide kinase, and total digestion with nuclease $P_1$. It thus appears that different nucleotides can be inserted at +1 although the reaction yields are reduced substantially.

Optimization of Yields

The requirement of many biochemical and biophysical researchers for large amounts of homogeneous RNA prompted a relatively extensive survey of transcription reaction conditions with several different templates. Conditions were sought that would maximize the amount of RNA produced for a given amount of template and enzyme while minimizing the 3' terminal heterogeneity that is often observed. Buffer conditions were varied using constant template and enzyme concentrations. The concentration (40 mM) and pH (8.1 at 37°C) of the Tris-HCl buffer used previously was found to be optimal and changing to phosphate buffer did not alter the reaction products. Addition of NaCl or other monovalent ions decreased the yields and did not alter the product distribution. As reported, dithiothreitol (5 mM) and spermidine (1 mM) are needed for optimal transcription (12,9). Yields are increased by 50% by the addition of 0.01% Triton-X 100 and 80 mg/ml polyethylene glycol (8000 MW). Finally, raising the concentrations of NTPs up to 4 mM in each and 20 mM $MgCl_2$ resulted in a larger increase in the amounts of products. Lower NTP concentrations reduce the yield proportionally and higher NTP concentrations inhibit the reaction.

Although the above optimal buffer conditions seem to be the same for all templates, both the template and the enzyme concentrations appear to have optima that are different for each template. Therefore, in order to find optimal reaction conditions, template concentrations from 50 nM to 1 uM and enzyme concentrations from 1 U/ul to 50 U/ul must be tested. Once optimal conditions are found, the reactions can be scaled up to preparative levels and

Table 6: Summary of several large scale transcription reactions.

| BOTTOM STRAND | POLYMERASE (units/ul) | DNA (nM) | REACTION VOL. (ml) | RNA YIELD (milligrams) |
|---|---|---|---|---|
| B- 3 | 30 | 200 | 40 | 23.8 |
| B-10 | 10 | 50 | 10 | 1.5 |
| B-11 | 10 | 50 | 10 | 0.6 |
| B-12 | 10 | 50 | 10 | 0.7 |
| B-13 | 2 | 50 | 10 | 0.6 |

All reactions were 4.0 mM in each NTP, 20 mM $MgCl_2$, and incubated at 37°C for 4 hrs. The product of template B-3 was isolated by preparative HPLC and was a heterogeneous RNA consisting of a 24 and a 25 mer. All other products were isolated by preparative polyacrylamide gel electrophoresis and were homogeneous products.

the yields remain consistent. Several preparative reactions are summarized in Table 6.

## DISCUSSION

Despite the apparent complexity of the T7 RNA polymerase-promoter interaction, this system can be used as a relatively simple method for making RNA. The small size of the templates makes them easy to synthesize and isolate, and only one new DNA strand is needed for each new RNA. The system offers enough sequence flexibility that almost any RNA desired can be synthesized. Since the reactions can be scaled up, large quantities of RNA can be produced on a routine basis.

Although chemical synthesis of oligoribonucleotides may become more generally available within the next few years (21), the transcription method described here still retains many advantages. Side reactions which can produce biologically inactive RNAs do not occur in enzymatic transcription reactions. Multiple rounds of transcription allow the amount of synthetic DNA to be amplified as much as 250 fold when converting it to RNA. Finally, unlike stepwise chemical synthesis, there is no decrease in yield with increasing length of the RNA. In fact, longer transcripts actually appear to be made more efficiently than shorter transcripts. In a recent experiment, a 76 nucleotide transcript was amplified nearly 1500 times from a 93 nucleotide template. Therefore, synthesis of oligoribonucleotides using T7 RNA polymerase and synthetic DNA templates provides an immediately available method for efficiently making small RNAs for both biochemical and biophysical studies.

ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health Grant GM 36944. We wish to thank Dr. John Dunn for encouraging us to use T7 RNA polymerase and for the generous gift of the cloned gene.

REFERENCES

1. Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, R., and Green, M.R. (1984) Nuc. Acids Res., 12, 4035-7056.
2. Chamberlin, M. and Ryan T. (1982) in "The Enzymes," Boyer, P.D., ed., 15, 85-108.
3. Davanloo, P., Rosenburg, A.H., Dunn, J.J. and Studier, F.W. (1984) Proc. Nat. Acad. Sci., 81, 2035-2039.
4. McAllister, W.T., unpublished results.
5. Kontani, H., Ishizaki, Y., Hiraoka, N., and Obayashi, A. (1987) Nuc. Acids Res., 15, 2653-2661.
6. Sealey, P.G. and Southern, E.M. (1982) in "Gel Electrophoresis of Nucleic Acids," Rickwood, D. and Hames, B.D., eds., IRL Press., 39-75.
7. Johnson, R.A. and Walseth, T.F. (1979) in "Advances in Cyclic Nucleotide Research," 10, Brooker, G., Greengard, P., and Robinson, G.A., eds., pp. 135-167.
8. England, T.E. and Uhlenbeck, O.C. (1978) Nature, 275, 560-561.
9. Chamberlin, M. and Ring, J. (1973) J. Biol. Chem., 248, 2235-2244.
10. Nishimura, S. (1979) in "Transfer RNA: Structure, Properties, and Recognition," Shimmel, P.R., Soll, D., and Abelson, J., eds., Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 551-561.
11. Randerath, K., Gupta, R.C., and Randerath, E. (1980) in "Methods in Enzymology", 65, Academic Press Inc., 638-680.
12. Uhlenbeck, O.C., Carey, J., Romaniuk, P.J., Lowary, P., and Beckett, D. (1983) J. Biomol. Struct. and Dyn., 1, 539-552.
13. Martin, C.T. and Coleman, J.E. (1987) Biochem., 26, 2690-2696.
14. Dunn, J.J. and Studier, F.W. (1983) J. Mol. Biol., 166, 477-535.
15. Ikeda, R.A. and Richardson, C.C. (1986) Proc. Nat. Acad. Sci., 83, 3614-3618.
16. Von Hippel, P.H., Bear, D.G., Morgan, W.D., and McSwiggen, J.A. (1984) Ann. Rev. Biochem., 53, 389-446.
17. Carpousis, A.J. and Gralla, J.D. (1980) Biochemistry., 19, 3245-3253.
18. Levin, J.R., Krummel, B., and Chamberlin, M.J. (1987) J. Mol. Biol., 196, 85-100.
19. McClure, W. (1985) Ann. Rev. Biochem., 54, 171-204.
20. Kwang, C. and Wu, C.W. (1987) Nuc. Acids Res., 15, 2279-2294.
21. Caruthers, M.H., Dellinger, D.J., Prosser, K., Barone, A.D., Dubendorff, J., Kierzek, R., and Rosendahl, M. (1986) Chemica Scripta, 26, 25-30.