Transcription Initiation Site Selection and Abortive Initiation Cycling of Phage SP6 RNA Polymerase*

Sang-Chul Nam and Changwon Kang‡

From the Department of Biological Science and Engineering, Korea Advanced Institute of Science and Technology, Seoul 130-650, Korea

The bacteriophage SP6, T7, and T3 transcription systems are not only popular for in vitro RNA synthesis (1, 2) and selective overproduction of proteins (3–5), but also excellent for molecular studies of transcription mechanisms. For example, the binding of the phage T7 RNA polymerase to the transcription promoter was studied recently by footprinting (6–8), filter binding (9), and mutagenesis (10).

Compared to the T7 system, relatively little is known of the SP6 system. It has been shown that mutations of SP6 DNA sequences at and immediately downstream of the transcription initiation site affect the SP6 transcription efficiency substantially (11). The objective in this paper is to determine the effects of these mutations on transcription initiation position. Although there have been a few examples where promoter mutations shift the transcription initiation site (12, 13), it has not been determined in any system whether mutations right at and around the initiation site shift the start site or not.

In this paper, we report the transcription initiation sites of the SP6 initiation sequence mutants. These shed some light on the mechanisms involved in the selection of the initiation site by the phage SP6 RNA polymerase. The initiation site of each mutant was determined by a novel method, using the nucleotide-specific pausings of the phage SP6 RNA polymerase under nucleotide-limiting conditions.

MATERIALS AND METHODS

RESULTS AND DISCUSSION

Nucleotide-specific Pausing of RNA Polymerase—The transcription initiation site of the pSP64 containing the wild-type promoter is precisely known to be +1 G (Table I) (11, 16). Transcription of the pSP64 in the presence of only ATP, GTP, and UTP produced mostly 5-mer, because C comes in the sixth nucleotide position, and shorter RNAs were also produced in low yield as abortive initiation products similar to those found for certain Escherichia coli promoters. When CTP was added up to 10 mM, the results were the same.

However, when CTP was added at 0.1 mM to 3 mM in a transcription mixture containing all the other nucleotides at 0.5 mM each and the plasmid pSP64, 9-mer, 15-mer, 18-mer, 23-mer, etc. were produced as well as smaller abortive initiation products up to 6-mer (Fig. 1, lane 2). These sizes of the RNA products are precisely matched with the noncoding strand sequence of the plasmid (Table I). CTP is required at the positions +6, +10, +16, +19, +24, etc. Therefore, the production of these oligomers must be due to the RNA polymerase pausing only at the C positions, caused by the limiting concentrations of CTP. The fidelity of this nucleotide-specific pausing was also demonstrated with 10-mer, 11-mer, 16-mer, 22-mer, etc. produced by UTP limitation (Fig. 1, lane 7).

Thus, when a ribonucleotide was at limiting concentration in the in vitro transcription reaction, the phage SP6 RNA polymerase stalled, or paused long enough only at the positions of the limiting nucleotide and the elongation was aborted. As a result, sequencing ladders of limited size range were obtained with nucleotide limitations. Naturally, at higher concentrations of the limiting nucleotide, the longer oligomers were more prevalent than the shorter ones.

In general, pausing of E. coli RNA polymerase is observed not only when the substrate concentrations are low or too high, but also when the stable RNA stem-loop structures are made (17). It was also proposed that the E. coli RNA polymerase may be stalled sometimes by downstream G-C-rich DNA sequences (18). The pausing of the phage SP6 RNA polymerase observed in this experiment, however, resulted only from the limitation of ribonucleotide concentration.

* This work was supported by a research grant from Korea Science and Engineering Foundation. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence and reprint requests should be addressed.

† Portions of this paper (including "Materials and Methods" and Table I) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are included in the microfilm edition of the Journal that is available from Waverly Press.
Transcription Initiation of SP6 RNA Polymerase

The limiting nucleotide was UTP, except lane 2 where CTP was limiting. The plasmid samples are, from the left, pCKSP6 (CK), pSP64 (SP), pCKSP6d3 (d3), pCKSP6d4 (d4), pCKSP6d5 (d5), pCKSP6d6 (d6), and pSP64 (SP). [α-32P]GTP was added to 2–5 μCi, depending on the transcription efficiency shown in Table I.

To see how stable the elongation ternary complexes of polymerase-DNA-RNA stalling at the limiting nucleotide positions are, chase experiments were performed. Transcription of pSP64 was carried out with radioactive GTP for 1 h or 10 min under UTP-limiting conditions. Then, a half of the mixture was further incubated with 0.5 mM UTP for 10 min, while the other half was not. There was no quantitative change in the distribution of pausing products by the chase (data not shown). If some of the pausing transcripts were still attached to the complex after the first 1-h or 10-min incubation, they could be further elongated when the limiting nucleotide was added fully.

Our results, however, suggested that during the first incubation nearly all the stalled complexes were dissociated. The paused ternary complexes of the phage SP6 RNA polymerase appear to be much less stable than the corresponding E. coli RNA polymerase complexes (19, 20), like those of phage T7 RNA polymerase (21).

Initiation Site Determination—The RNA polymerase pausing by limiting concentrations of a nucleotide was utilized to determine the transcription initiation site of all the plasmids carrying various sequences right at and around the initiation site shown in Table I.

For example, lane 1 of Fig. 1 shows the pausing products of the plasmid pCKSP6 with UTP limitation. When the transcription initiation site of the wild-type promoter (pSP64) is assigned the position +1, the noncoding strand sequence from −4 to +25 of the pCKSP6 is TATAGGATCCCC-GGGCGAGCTCGAATTCG (Table I). Since UTP is required at the underlined positions +17, +22, +23, etc., if transcription starts from the +1 G (underlined), the pausing products of UTP limitation should be only 16-mer, 21-mer, 22-mer, etc. This is the case as shown in Fig. 1, lane 1. If transcription started from other than +1 G, the pausing transcripts should have been longer or shorter than those. This start site was also confirmed by the pausing products of 13-mer, 19-mer, and 20-mer seen with ATP limitation.

The initiation sites of all the other plasmids were also determined in the same way. With the plasmid pCKSP6d3 (Fig. 1, lane 3), the three bands could not be sized unambiguously because they were positioned between the standard size marker bands in a usual sequencing gel. However, when the ratio of acrylamide to bisacrylamide was reduced from 19:1 to 10:1, and the total concentration of acrylamide and bisacrylamide was raised from 12% or 15% to 18% or 20%, they could be determined precisely as 13-mer, 18-mer, and 19-mer. This is probably because the size effect is more dominant on the mobility than the base composition variation effect in high cross-linked gels of smaller pore size (22).

The results with the plasmid pCK-X show a dramatic example of nucleotide-specific pausing of the RNA polymerase. The sequence of its noncoding strand has an interesting stretch of 8 As from +15 to +23 interrupted only by +18 G (Table I). ATP limitation yielded products of pausing at every A position, showing all the bands of 14-mer through 22-mer but 17-mer (Fig. 2). These results determined the initiation site of pCK-X to be the +1 G.

Another example that needs explanation is the plasmid pCKSP6d6 (Fig. 1, lane 6), where the TATA sequence from −4 to −1 is changed to TATC, and +1 is C. According to the sequence of the plasmid (Table I), if the transcription still starts from the +1 C, 15-mer and 16-mer are expected from UTP limitation, among others. The results were, however, 15-mer, 16-mer, and 17-mer, and the 16-mer was roughly twice more prevalent than each of the others. It can be explained by multiple initiations from +1 C (15-mer and 16-mer) and from −1 C (16-mer and 17-mer). Densitometric analysis of the three bands revealed that the two initiations occurred with the more or less same frequency, 50 ± 10%.

The initiation sites of the plasmids pCKSP6d7 and pCK+Xd could not be determined, because their transcription efficiencies were too low (11). Attempts to solve the problem by raising the reaction scale and using more radioactive nucleotide failed, because the transcripts appeared to be only
the small abortive initiation products.

Initiation Site Selection—The initiation sites determined for all of the plasmids are shown in Fig. 3. Again, the position is numbered such that the initiation site of the wild-type promoter (pSP64) is +1. Transcriptions of all the plasmids with G at the +1 start only from the +1 G, even if it is surrounded by purines at −1 and +2, regardless of the downstream sequence. When the +1 G is changed to either A (pCK-Xa) or C (pCKSP6d5) with its upstream sequence remaining the same, transcription still starts only at the +1 A or +1 C, respectively. This strongly suggests that the initiation site selection by the phage SP6 RNA polymerase depends on the precise physical position rather than the species of initiating nucleotide. This property of the SP6 RNA polymerase appears to be shared by the T7 RNA polymerase as recently reported (10).

However, the plasmid pCKSP6d6 shows a novel feature. The TATA sequence from −4 to −1 is changed to TATC, and the +1 is C in the noncoding strand. The transcription of the plasmid starts from either −1 C or +1 C, with roughly equal frequency. It can be interpreted as that the start position is located roughly halfway between the two nucleotide positions, as resulted from DNA sequence-dependent perturbation of helical DNA structure of the region.

Recently, Zhou et al. (23) suggested the structure of [d(GGTATACC)], in which the −TATA− moiety exists in the wrinkled D (wD) form and the GG and CC parts are in regular B form. The axial rise per residue is 0.304 nm in the D form and 0.338 nm in the B form. If the −TATA− moiety in the plasmid pSP64 or pCKSP6d5 exists in the wD form, the axial length from the −5 C to +1 G (pSP64) or +1 C (pCKSP6d5) would be 1.52 nm (5 × 0.304 nm). This distance would be made by 4.5 residue rises (1.52 nm + 0.338 nm) in B form DNA. If the −TATCC− sequence in the plasmid pCKSP6d6 exists in normal B form, the start site would be 4.5 residues away from the −5 C, which is halfway between the −1 and +1 residues.

Our hypothesis that the phage SP6 RNA polymerase starts to read the coding strand at a certain distance from a direct contact point on the promoter seems reasonable, because the region between the putative contact point and the transcription start point, namely all or part of the TATA sequence from −4 to −1, does not seem to be in direct contact with the RNA polymerase. Kang and Wu (11) reported that the base pairs at positions −1 and −2 can be changed without altering the SP6 transcription efficiency. Taking into consideration the similarity between the SP6 and T7 systems, it is also supported by the report by Chapman and Burgess (24) that the point mutations at −4, −3, and −2 of the phage T7 promoter have little effect on RNA polymerase binding and transcription efficiency.

Abortive Initiation Cycling—When a nucleotide was at limiting concentration, transcriptions with the plasmid pSP64 produced small oligomers up to 6-mer, regardless of which nucleotide was limited, while longer RNAs were the products of nucleotide-specific pausings (Fig. 4). All the other plasmids also produce the small oligomers up to 6-mer. These results indicate that phage SP6 RNA polymerase gets into the stable elongation mode, escaping from the abortive initiation cycling after synthesizing 6-base-long RNA. Recently, Ikeda and Richardson (6) found that the T7 RNA polymerase maintains the contact with the upstream promoter until synthesizing 6-mer RNA, based on their chemical footprinting data. Milligan

---

**Fig. 3.** Transcription initiation sites of the phage SP6 initiation sequence mutants. All mutants start transcription only at the +1 position, except the pCKSP6d8 which starts at both −1 and +1 positions with roughly equal frequency.

**Fig. 4.** Abortive initiation and elongation transcripts from the pSP64 with 10 μM GTP (G), with 3 μM CTP (C), or without the addition of UTP (U) to residual amounts of ribonucleotides in plasmid preparations. Each transcription contained only one nucleotide at the limiting concentration and the other three nucleotides at 0.5 mM each. The transcripts were labeled by incorporation of [α-32P]GTP.
et al. (10) also found the abortive initiation products of 2- to
6-base length with the T7 RNA polymerase.

The abortive initiation products were less prevalent at a
higher concentration of nucleotides, although quantitative
analysis was not attempted because the ethanol precipitation
of the small oligomers was not presumed to be quantitative.
When all four ribonucleotides were added fully (0.5 mM each),
the abortive initiation products were negligible compared to
long transcripts. On the other hand, when one or all of the
four ribonucleotides were at limiting concentrations, the small
abortive products were made in high yield. This suggests that
the abortive release of transcripts competes with the productive
formation of the stable elongation complexes.

Transcriptions of the plasmids pCKSP6d3, pCKSP6d4,
pCKSP6d5, and pCKSP6d6, which require neither ATP nor
UTP during the synthesis of the 6-mer, produced relatively
small amounts of the abortive initiation products under the
ATP- or UTP-limiting conditions. On the other hand, under the
CTP- or GTP-limiting conditions, the small products
were made in higher yield. Thus, the amount of the abortive
products depends on the sequence from +1 to +6, when a
ribonucleotide is at a limiting concentration.

These facts should be noted in order to achieve efficient
productive incorporation of the radioactive ribonucleotide,
which is usually at a low concentration for high specific
activity of transcripts, when one chooses which radioactive
ribonucleotide to use. For example, the use of radioactive
CTP is the best choice for the transcriptions of the recombi-
nant derivatives of the plasmids pSP64 and pSP65, because
the 5' end sequence of their transcripts will be 5'GAAUAC . . . ,
and so the CTP will be used only once for the last during
the abortive initiation cycling.

The abortive initiation products from the multiple initia-
tions at -1 and +1 of the plasmid pCKSP6d6 were up to only
6-mer. The 7-mer was not observed at all, which would have
been produced if the transition from the abortive to the
productive initiation occurred between positions +6 and +7.
Instead, it suggests that the transition depends on the suc-
cessful incorporation of the 7th ribonucleotide, and/or the
stability of the ternary complex of the 6-mer RNA, RNA
polymerase, and DNA coding strand template. The stability
of the ternary complex does not appear to depend on the
RNA-DNA hybrid stability, because transversion mutations
between C:G and G:C base pairs in the region between +1
and +3 still affect the transcription efficiency substantially
(11). Therefore, it appears to be the interactions of the polym-
erase with the RNA oligomers or with DNA template that
determine the extent of the abortive initiation and the tran-
scription initiation efficiency.

REFERENCES
397-415
Press, Washington, D. C.
U. S. A. 82, 1074-1078
130
5. Rosenberg, A. H., Lade, B. N., Chui, D.-S., Lin, S.-W., Dunn, J.
U. S. A. 83, 5614-5618
Biochemistry 26, 1539-1546
14, 2811-2827
10. Milligan, J. P., Groebe, D. R., Witherell, G. W., and Uhlenbeck,
2294
Biol. 157, 619-633
257, 5772-5778
15. Birnboim, H. C., and Doly, J. (1979) Nucleic Acids Res. 7, 1513-
1523
Salmonella typhimurium. Cellular and Molecular Biology (Neid-
hardt, F. C., ed) Vol. 2, pp. 1241-1275, American Society for
Microbiology, Washington, D. C.
18. Gilbert, W. (1976) in RNA Polymerase (Losick, R., and Cham-
berlin, M., eds) pp. 173-205, Cold Spring Harbor Laboratory,
Cold Spring Harbor, New York
177
Biol. 196, 85-100
527-534
23. Zhou, N., Bianucci, A. M., Pattabiraman, N., and James, T. L.
(1987) Biochemistry 26, 7905-7913
15, 5419-5432
Supplementary Material to:
TRANSCRIPTION INITIATION SITE SELECTION
AND ABORTIVE INITIATION CYCLING
OF PHAGE SP6 RNA POLYMERASE
Sang-Chul Nam and Changwon Kang

MATERIALS AND METHODS

Reagents - The phage SP6 RNA polymerase preparation was described previously (11), and its specific activity was 700,000 units/mg as defined by Butler and Chamberlin (14). The N\textsubscript{2}O\textsubscript{4} inhibitor, RNase, was purchased from Procter Blanc. (\textsuperscript{-32P}GTP (800 Ci/mmol) was purchased from American Corp. and New England Nuclear. All the ribonucleoside triphosphates were purchased from Pharmacia P-L and neutralized to pH 7.5.

DNA Templates - All the SP6 promoter/initiation sequence mutants listed here were described previously (11) except pCK-Xa, and their partial sequences are shown in Table 1. The plasmid pCK-Xa has an insertion of a 22 base pair sequence (from position 794 to 777 of pSP64) at the transcription initiation site of pCK-X. All the plasmids were prepared by the alkaline lysis method of Birckholz and Daly (15). The plasmid pSP64 was further purified by equilibrium centrifugation in CsCl-ethidium bromide gradients, and all the plasmids were purified by 5 time extraction with phenol/chloroform (1:1) mixture and 4 time extraction with ether and precipitated with ethanol.

In Vitro Transcription - Transcriptions were carried out in 30 \mu l of reaction mixture containing 40 mM Tris-\textsubscript{HCl} pH 7.5, 20 mM KCl, 6 mM MgCl\textsubscript{2}, 2 mM spermidine, 10 mM dithiothreitol, 0.5 mM each ribonucleoside triphosphate, 2 \mu g/ml RNase, 0.2% - 1 mg circular plasmid DNA, 1 unit of phage SP6 RNA polymerase, and 10 \mu Ci [\textsuperscript{32P}]GTP (800 Ci/mmol) at 40 \degree C for 1 h. For nucleotide-specific pausing, a nucleotide was either depleted or added to less than 5 mM. The reaction was stopped by addition of 30 \mu g RNA, N\textsubscript{2}O\textsubscript{4}OAc to 1 M, and ethanol to 70%. The tubes were chilled at -20 \degree C for 15 min, and centrifuged for 30 min in a swinging-boundary centrifuge at 4 \degree C. The pellets were washed with 70% ethanol, dried and dissolved in 10 \mu l of gel loading buffer containing 80% (v/v) formamide, 40 mM Tris-\textsubscript{HCl}, pH 5.3, 2.5 mM EDTA, and 0.25% (w/v) xylene cyanol.

Transcription Initiation Site Determination - Transcriptions were carried out in vivo with non-preparations of the plasmids that contain residual amounts of the ribonucleosides. Each transcription contained only one nucleotide at the limiting (residual) concentration and all the other nucleotides at 0.5 mM each. The transcripts were labeled by incorporation of [\textsuperscript{\gamma-32P}]GTP, which was chosen because the GTP is usually the initiating nucleotide. The pausing products were separated by nuclease on high-resolution polyacrylamide-urea gels. The precise sizing of the transcripts using those of the pSP64 as size markers precisely determined the initiation site of each plasmid. Because of the small oligomers produced from the abortive initiation cycling and lower resolution of longer oligomers in a gel, only the oligomers between 7-mer and 25-mer were analyzed for initiation site determination, although longer oligomers were also observed in some cases.

Electrophoresis of RNA - To accurately determine the sizes of transcripts smaller than 30 nucleotides, the 17% to 20% polyacrylamide-50% urea gels were used. The ratio of acrylamide to bisacrylamide was either 9:1 or 10:1 (w/w). Gels were gared for at least 1 h at 50 W before loading. Samples in the formamide loading buffer were heated at 65 \degree C for 2 min and quickly chilled immediately prior to loading. After electrophoresis, the gel was covered with Saran-Wrap and exposed to an Agfa-Curix AP-4 X-ray film with a Kodak X-Omat screen at -70 \degree C for 12 - 36 h.

Table 1

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Non-coding strand * nucleotide sequence</th>
<th>Relative ** activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>pSP64</td>
<td>TGACACTATGAGATCAAG25GCTGCGATCTCGATCT</td>
<td></td>
</tr>
<tr>
<td>pCKSP6</td>
<td>GATACATGAGATCAAG25GCTGCGATCTCGATCT</td>
<td>100%</td>
</tr>
<tr>
<td>pCKSP64-3</td>
<td>GCGCGGCAGCTGACTTCTAATGCTAATGCT</td>
<td>65%</td>
</tr>
<tr>
<td>pCKSP64-4</td>
<td>GCTGCGGCAGCTGACTTCTAATGCTAATGCT</td>
<td>75%</td>
</tr>
<tr>
<td>pCKSP64-5</td>
<td>GCGCGGCAGCTGACTTCTAATGCTAATGCT</td>
<td>15%</td>
</tr>
<tr>
<td>pCKSP64-6</td>
<td>GCGCGGCAGCTGACTTCTAATGCTAATGCT</td>
<td>18%</td>
</tr>
<tr>
<td>pCKSP64-7</td>
<td>GCGCGGCAGCTGACTTCTAATGCTAATGCT</td>
<td>3%</td>
</tr>
<tr>
<td>pCKX1</td>
<td>GGATCCCTCAAGCTATCTGCTATCTGCT</td>
<td>100%</td>
</tr>
<tr>
<td>pCKX</td>
<td>GGATCCCTCAAGCTATCTGCTATCTGCT</td>
<td>87%</td>
</tr>
<tr>
<td>pCKXd</td>
<td>GGATCCCTCAAGCTATCTGCTATCTGCT</td>
<td>7%</td>
</tr>
<tr>
<td>pCKY</td>
<td>GGATCCCTCAAGCTATCTGCTATCTGCT</td>
<td>65%</td>
</tr>
</tbody>
</table>

* The position +1 is defined as the transcription initiation site of the wild type promoter (pSP64). Only the sequences from -10 to +30 are shown. The hyphens indicate the sequences identical to the wild type.

** The relative transcription activities are compared to pCKSP64 (100%). These are taken from Kang and Wu (11).