Volume 15 Number 5 1987 Nucleic Acids Research

Studies on SP6 promoter using a new plasmid vector that allows gene insertion at the transcription initiation site

Changwon Kang* and Cheng-Wen Wu+

Department of Pharmacological Sciences, State University of New York, Stony Brook, NY 11794, USA

Received November 11, 1986; Revised and Accepted January 8, 1987

ABSTRACT

All the phage-promoter containing subcloning vectors available for in vitro transcription reactions contain a polylinker away from the transcription initiation site. A new SP6 transcription subcloning vector, pCKSP6, has been constructed, in which a gene can be inserted precisely at the transcription initiation site. This was achieved by bringing the BamHI cleavage site into the initiation site. When DNA ends of both insert gene and BamHI cleaved pCKSP6 are made blunt-ended using a single strand specific nuclease, the in vitro transcripts of the recombinant DNA by SP6 RNA polymerase will contain only the gene sequence immediately after the initiation base G.

Mung bean nuclease was used to generate a series of mutants resulting from step-wise deletion of single base pairs around the initiation site. Transcription assays with these SP6 promoter mutants revealed that not only the sequence immediately upstream of the initiation site but also the six base pairs from position +1 to +6 are important elements for promoter binding and/or transcription initiation activity. Furthermore, there appears to be a hierachy of importance of each base pair in the order of position +1 > +2 > +3 > +4, +5, +6, -1, -2.

INTRODUCTION

Single-stranded RNA with defined sequence can be synthesized by the specific transcription of cloned DNA in vitro. Several in vitro bacteriophage transcription systems, such as T3, T7, and SP6, have proven very useful for this purpose and have recently become commercially available. These systems take advantage of the stringent promoter specificities of the phage RNA polymerases, and provide convenient and efficient means to prepare microgram quantities of defined RNA transcripts with high specific radioactivities. The RNA synthesized by these methods can be used as hybridization probes (1) or RNA processing substrates (2). Furthermore, the transcripts have been shown to

be biologically active as mRNA (3) and infective viroid RNA (4).

One of the key elements in the in vitro phage transcription systems of T3, T7 and SP6 is the promoter containing plasmid. Typically, a gene of interest can be inserted into any one of restriction endonuclease sites in the polylinker sequence that is located downstream of the promoter sequence. The cloned plasmid is then used as the template in the in vitro transcription reaction. In all three phages, transcription starts at G in the consensus CACTATAG sequence (5). Also, the polylinker sequences in all available vectors are located some base pairs away from the initiation sites, resulting in transcripts that contain a stretch of plasmid nucleotide sequence at the 5' end. When substantially long RNA molecules are synthesized, the extraneous sequence at the 5' end may not interfere with the RNA activity of interest. In many cases, however, the additional 5' end nucleotides may pose a serious problem, particularly when the desired RNA molecules are relatively short. It was observed recently that (6) the infective activity of the synthetic polioviral RNA increased about 50 fold when the 60 nucleotide-long extraneous sequence at the 5' end was reduced to 2 bases.

To circumvent the problem of the plasmid derived sequence at the 5' end of the transcript, one can remove the undesired sequence on the cloned plasmid by oligonucleotide-directed site-specific deletion (7), or by trimming the ends of linearized vector and gene to appropriate length before insertion. These procedures, however, are laborious and time-consuming. Thus, it is desirable to have a plasmid that will allow direct insertion of a gene exactly at the transcription initiation site. In this paper, we report the construction of such a plasmid. Furthermore, mung bean nuclease was used to generate a series of deletion mutants around the initiation site. Comparative studies on relative transcription activities of these promoter mutants have helped to define the downstream boundary of the SP6 promoter, as well as the relative importance of individual base pairs from positions -2 to +6 that are involved in promoter binding and/or transcription initiation.

MATERIALS AND METHODS

(a) Recombinant DNA techniques

Plasmid pSP64 (1) was obtained from Promega Biotec. All enzymes were purchased from New England Biolabs, except where mentioned otherwise. Cohesive-end ligation was carried out at 12°C overnight, and blunt-end ligation at 25°C overnight in 20 ul solutions containing 50 mM Tris-HCl (pH 7.6), 10 mM $MgCl_2$, 1 mM ATP, 10 mM dithiothreitol, 50 mg/l BSA, and 1 or 2 units of phage T4 DNA ligase.

Phage T4 DNA polymerase reaction was carried out in 33 mM Tris-OAc (pH 7.9), 10 mM $Mg(OAc)_2$, 66 mM KOAc, 0.5 mM dithiothreitol, 0.1 g/l BSA, 0.25 mM each of dNTP, and 1 unit of T4 DNA polymerase at 37°C for 10 min. All dNTPs were purchased from Pharmacia P-L, and neutralized to pH 7.5. The single-stranded extensions of DNA (1 ug), were removed with 2 units of mung bean nuclease in 10 ul of 50 mM NaOAc (pH 5), 30 mM NaCl and 1 mM $ZnSO_4$ at 37°C for 30 min.

The recombinant plasmids were screened and characterized by digestion with the appropriate restriction endonucleases. The *E. coli* strain HB101 was used as recipient in highly efficient transformation as described by Hanahan (8). Plasmid DNAs were prepared by the alkaline lysis method of Birnboim and Doly (9) and further purified by equilibrium centrifugation in CsCl-EtBr gradients.

(b) *In vitro* transcriptions

A typical run-off transcription was carried out in 20 ul of reaction mixture containing 40 mM Tris-HCl (pH 7.5), 6 mM $MgCl_2$, 2 mM spermidine, 10 mM dithiothreitol, 0.1 g/l BSA, 0.5 mM each rNTP, 1.5 unit/ul RNasin, 30 nM linearized plasmid DNA, 10 units of SP6 RNA polymerase, and 25 uM [α-^{32}P]GTP (410 Ci/mmol) at 40°C for 1 h. Following digestion with an appropriate restriction enzyme, the mixture of DNA fragments was isolated by two phenol-chloroform extractions, followed by three ether extractions and an ethanol precipitation. The dried DNA pellet was resuspended in buffer and used directly in the transcription reaction. The ice-chilled transcripts were twice ethanol precipitated in the presence of 2 M NH_4OAc before analysis on a sequencing gel.

SP6 transcription assay was carried out with uncut circular DNA and 0.7 uM [5,6-^{3}H]UTP (35 Ci/mmol) at 40^{o}C for 1/2 h. Transcripts were precipitated by 10% trichloroacetic acid on ice for 10 min, filtered, and the radioactivity of precipitates was counted.

SP6 RNA polymerase was a gift of Dr. William T. Windsor, III in our laboratory, and its specific activity was 700,000 units/mg as defined by Butler and Chamberlin (10). The RNase inhibitor, RNasin, was purchased from Promega Biotec.

(c) Rapid Sequencing of Double Stranded DNA

The rapid preparations of double-stranded plasmid DNA were sequenced directly by the Klenow fragment of E. coli DNA polymerase I using an oligodeoxynucleotide primer. The plasmid DNA extracted from 3 ml overnight culture according to Birnboim and Doly (9) was dissolved in 100 ul of water. A 6 ul aliquot was digested by 2 units of restriction enzyme RsaI in 10 ul of 6 mM Tris-HCl (pH 8), 50 mM NaCl, 12 mM $MgCl_2$ and 6 mM 2-mercaptoethanol at 37^{o}C for 25 min and reacted with 1 ul of DNase-free RNase A (1 ug/ul) for additional 5 min. The digested DNA (4.5 ul) was transfered into a screw-cap microtube, and mixed with 1 ul of 0.25 M dithiothreitol, 0.5 ul of 10 X polymerase reaction buffer (BRL), and 4 ul of 2.5 mg/l 15-mer primer for SP6 promoter (BRL). After being boiled for 5 min and quickly chilled on ice, 2 ul of [α-^{32}P]dATP (3000 Ci/mmol, 10 Ci/l) and 1 ul of Klenow fragment of E. coli DNA polymerase I (1 unit/ul) were added. A 2.5 ul aliquot of the mixture was mixed, on ice, with 2 ul of each of four ddNTP-dNTP mixtures. After incubation at 37^{o}C for 15 min, the elongation reactions were continued with 1 ul of 2.5 mM dNTP mixture at 37^{o}C for another 15 min. After mixing with 4 ul of 99% formamide, each sample was loaded on a sequencing gel. Composition of each ddNTP-dNTP mixture was as follows:

A: 0.17 mM ddATP, 63 uM dCTP, 63 uM dGTP, 63 uM dTTP;
C: 1.7 mM ddCTP, 8.3 uM dCTP, 83 uM dGTP, 83 uM dTTP;
G: 1.7 mM ddGTP, 83 uM dCTP, 8.3 uM dGTP, 83 uM dTTP;
T: 3.3 mM ddTTP, 83 uM dCTP, 83 uM dGTP, 8.3 uM dTTP.

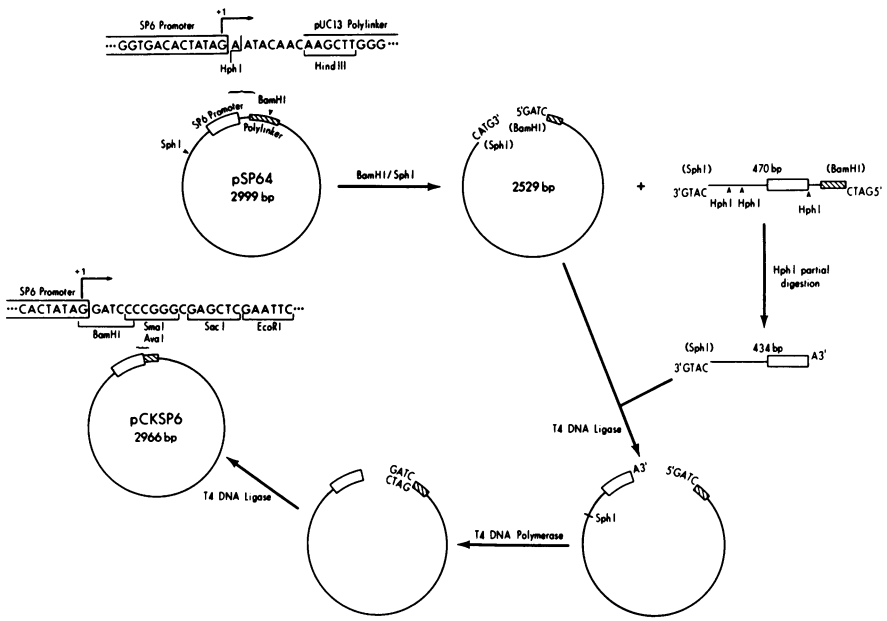

Figure 1: Construction of pCKSP6 from pCK64. The sequences around the transcription initiation site are shown. Transcription starts with G at position +1. The SP6 promoter is boxed and polylinker is shaded.

RESULTS

(a) Construction of pCKSP6 vector

The strategy for the construction of pCKSP6 vector is based on the observation that among many other sites, restriction endonuclease HphI cleaves the SP6 promoter-containing vector pSP64 at a base pair immediately following the transcription initiation site. This cleavage site is staggered about the A/T base pair located nine base pairs downstream from the recognition sequence, GGTGA, of HphI.

```
                    HphI
                  +1 |
5'...GGTGACACTATAGA|ATACAA...3'    non-coding strand
3'...CCACTGTGATATC|TTATGTT...5'      coding strand
     SP6 Promoter |
```

As outlined in Fig 1, the plasmid pSP64 was first cut by BamHI and SphI to yield two fragments of 2.5 kb and 470 bp, respectively. Partial digestion of the smaller fragment by HphI generated a 434 bp fragment containing the SP6 promoter. This fragment was easily resolvable from other partially digested fragments by 8% polyacrylamide gel electrophoresis.

The 434 bp fragment was ligated back onto the 2.5 kb pSP64 fragment through the SphI sticky ends using T4 DNA ligase. The resulting linear DNA was reacted with T4 DNA polymerase which removed the 3' overhang of the SphI cut and filled in the 5' overhang of the BamHI cut, making both ends flush. The blunt ended linear DNA was then circularized using T4 DNA ligase. The ligation mixture was used to transform E. coli HB101 cells, and the successful transformants were screened by analysis of plasmid DNA on an 8% polyacrylamide gel after digestions with SphI and EcoRI. In this circular recombinant DNA, the BamHI site was regenerated at the transcription initiation site (Fig 1).

The new vector of 2966 bp long, named pCKSP6, was proven to be an active template for SP6 RNA polymerase. When the plasmid pCKSP6 was linearized with SmaI (yielding flush ends), EcoRI (5' four-base overhang) or BstNI (5' one-base overhang), the transcripts synthesized by run-off transcription in vitro with SP6 RNA polymerase were precisely 8, 24 or 115 nucleotides long, respectively, as expected (data not shown). In all these transcripts, the initiation site was the first G/C base pair of the BamHI recognition sequence. The coding strands of both the 5' overhang and flush ends were fully transcribed to the last nucleotide.

(b) Usage of vector pCKSP6

There are two simple ways to insert a gene of interest into the BamHI site of pCKSP6. Any DNA fragment whose ends are compatible with 5'GTAC cohesive ends generated by restriction enzymes like BamHI, BclI, BglII, Sau3A, XboII, etc. can be easily inserted into the unique BamHI site. Otherwise, a gene can be inserted by blunt-end ligation after trimming the overhangs with mung bean or S1 nuclease.

A 193 bp fragment from HgiAI/BamHI digestions of the plasmid pXbs201 was chosen to demonstrate the usage of pCKSP6 vector. The

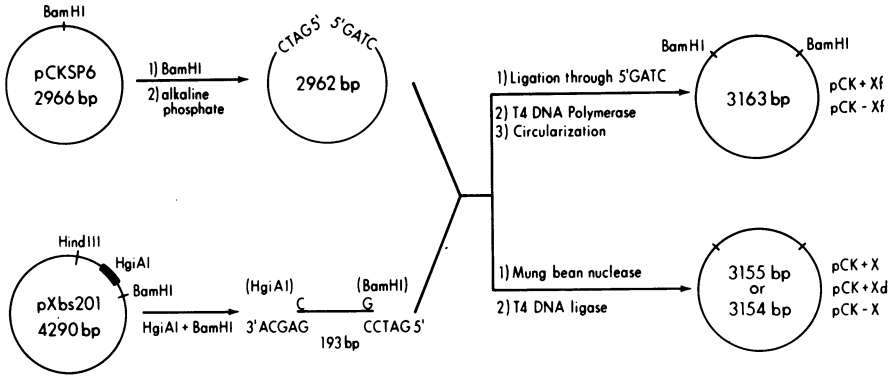

Figure 2: Insertion of Xenopus borealis somatic 5 S RNA gene isolated from pXbs201 into the transcription initiation site of pCKSP6. The DNA sequences around the initiation site of the recombinant DNAs are listed in Table I.

plasmid pXbs201 (4290 bp in length) contains a Xenopus borealis somatic (Xbs) 5 S RNA gene along with flanking sequences between the HindIII and the BamHI sites of pBR322 (11). The restriction enzyme HgiAI cleaved the plasmid pXbs201 near the 5' end of the Xbs 5 S RNA gene, removing the entire 5' flanking sequence except for a single C/G base pair. A double digestion of pCKSP6 by HgiAI and BamHI yielded a 193 bp fragment having both 5' and 3' protruding ends (Fig. 2). This fragment was treated with mung bean nuclease to remove the single-stranded overhangs. The blunt-ended fragment containing 5 S RNA gene was inserted into the destroyed (blunt-ended) BamHI site of pCKSP6, using phage T4 DNA ligase (Fig. 2). The blunt end insertion resulted in two types of recombinant DNA of 3155 bp containing only one insert in either orientation. The recombinant DNA with positive orientation was named pCK+X, and the other with negative one, pCK-X. Both recombinants were also active templates for SP6 RNA polymerase. The termination signal for Xenopus RNA polymerase III in the Xenopus borealis somatic 5 S RNA gene was not recognized at all by SP6 RNA polymerase.

Under the conditions described in Materials and Methods, single-stranded mung bean nuclease removed the single-stranded overhangs of the 193 bp DNA fragment to form flush ends. However, in some cases the enzyme could also penetrate into the double-stranded edges of linear DNA, probably because the flush ends of

Table I. Transcription activities of SP6 promoter mutants

Plasmid	Gene	Non-coding strand [a] Nucleotide sequence	Relative[b] activity
T3 consensus	5'	CCTCACTAAAGGGAGA 3'	
T7 consensus		ACTCACTATAGGGAGA	
SP6 consensus		TGACACTATAGAA	
		-10 +1 +10	
pCKSP6		TGACACTATAGGATCCCCGGGCGAGCTCGAATT	100%
pCKSP6d3		TGACACTATAGCCCCGGGCGAGCTCGAATTCGT	65%
pCKSP6d4		TGACACTATAGCCCGGGCGAGCTCGAATTCGTA	79%
pCKSP6d5		TGACACTATACCCGGGCGAGCTCGAATTCGTAA	19%
pCKSP6d6		TGACACTATCCCGGGCGAGCTCGAATTCGTAAT	18%
pCKSP6d7		TGACACTACCCGGGCGAGCTCGAATTCGTAATC	3%
		+123456789	
pCK+Xf	5S(+)	TGACACTATAGGATCCGCCTACGGCCATACCAC	87%
pCK+X	5S(+)	TGACACTATAGCGCCTACGGCCATACCACCCTG	29%
pCK+Xd	5S(+)	TGACACTATACGCCTACGGCCATACCACCCTGA	4%
		+123456789	
pCK-Xf	5S(-)	TGACACTATAGGATCCGGCTGTTCGGGGAAAGA	102%
pCK-X	5S(-)	TGACACTATAGCGGCTGTTCGGGGAAAGAAAAA	42%

[a] Only the 3' half of the consensus sequences are shown. The sequences shared by the three consensus sequences are underlined. Transcription of wild type promoter starts at G in position +1.
[b] The relative transcription activities were compared to pCKSP6 (100%).

DNA might be partially denatured or frayed under these conditions. Due to this "nibbling" activity of the enzyme, a few transformants contained the plasmid that lack one G/C base pair. This recombinant of 3154 bp was named pCK+Xd, and its partial sequence is shown in Table I.

The ends of the 193 bp DNA fragment could also be made blunt by T4 DNA polymerase that filled in the 5' protruding ends of the BamHI cut and removed the 3' protruding ends of the HgiAI cut. As shown Fig. 2, the two DNA fragments were first ligated through the BamHI cohesive ends and the linear recombinants were reacted with T4 DNA polymerase to make flush ends before

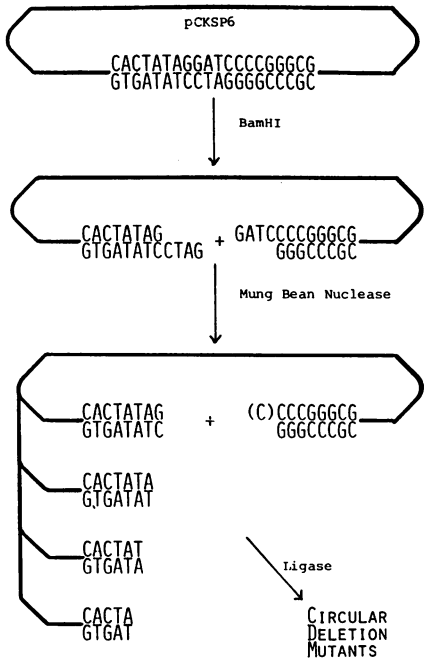

Figure 3: Deletion mutations of pCKSP6 using mung bean nuclease. The nuclease reaction was stopped by chilling on ice and extraction with phenol-chloroform. An aliquot of the ligation mixture was used to transform E. coli HB101 competent cells. Only the DNAs having blunt ends after the nuclease reaction, that are ligated to form circular DNAs and transformed into the competent cells, are shown. The base C in parenthesis in the non-coding strand at the downstream end indicates that in one mutant the residue survived in the nuclease reaction and the resulting gap in the coding strand was filled in by some unknown process.

recircularization by T4 DNA ligase. The resulting recombinant of 3163 bp was named pCK+Xf or pCK-Xf, depending on the orientation of the insert. A BamHI site was regenerated at the junction of the filled-in BamHI site and the removed HgiAI site. The partial sequences of these recombinant DNAs are shown in Table I. Plasmids pCK+Xf and pCK-Xf were as active as pCKSP6 when used as the template in the in vitro transcription by SP6 RNA polymerase, indicating that the sequence downstream from +7 is not essential for transcription initiation.

(c)Mutations using mung bean nuclease

Taking advantage of the capability of mung bean nuclease to penetrate into frayed DNA double-stranded edges, we have generated various deletion mutants that lack a small number of base pairs near the transcription initiation site. The plasmid pCKSP6 was cut at the unique BamHI site and treated with mung bean nuclease at 37°C for 30 min as described in Materials and Methods (Fig. 3). The linear DNA was recircularized by T4 DNA ligase and the ligation mixture was used to transform E. coli HB101 cells. Plasmid DNA isolated from each transformant was directly sequenced from a 15-mer primer, which hybridized to a coding strand sequence in the upstream region of the SP6 promoter using dideoxynucleotides and the Klenow fragment of E. coli DNA polymerase I. After sequencing 33 randomly selected plasmids, five deletion mutants were obtained. They were named pCKSP6d3, pCKSP6d4, pCKSP6d5, pCKSP6d6, and pCKSP6d7, the number at the end defining the size of deletion in base pairs. In one of these mutants, pCKSP6d3, mung bean nuclease did not remove all the overhangs, instead leaving one base (C), and the resulting gap in the coding strand was filled in. The pertinant partial sequences of all the mutants are listed in Table I.

It is interesting to note that the penetration of mung bean nuclease through the double-stranded DNA ends is asymmetric (Fig. 3 & Table I). No base pairs were removed at the downstream end of BamHI cleaved pCKSP6, but up to three base pairs were deleted at the upstream end. This is probably because the sequence of the downstream end has a stretch of 8 G:C base pairs and is therefore more stable than the sequence at the upstream end which is rich in A:T base pairs. Also, in the plasmids containing a 5 S RNA gene in both orientations (pCK+X and pCK-X), a stretch of 4 G:C base pairs at the downstream end (no base pairs removed) is more stable than the upstream end (one base pair deleted) under the conditions employed in our experiments (Table I).

(d) Transcriptional activities of mutants

The conditions for the maximum in vitro transcription activity of SP6 RNA polymerase has been established (1). The promoter mutants described above were assayed for their transcription activities under these conditions. The relative

efficacies of these mutants to serve as template in the SP6 *in vitro* transcription reactions compared to that of pCKSP6 are listed in Table 1.

The template activities of plasmids pCKSP6, pCK+Xf and pCK-Xf are very similar. Since these three plasmids share the identical sequence down to position +6 and their further downstream sequences are different, the measured activities indicate that the sequences downstream from position +7 do not measurably affect promoter function or transcription initiation.

The seven plasmids listed below share the identical sequence upstream from position -1. When these plasmid DNAs were used as template in *in vitro* transcription, their activities varied greatly. This can only be explained if promoter efficacy is dependent on the six base pairs from position +1 to +6, as well as the previously recognized consensus sequence upstream from the transcription initiation site (5,16,17).

plasmids	+123456	rel. act.
pCKSP6	GGATCC	100%
pCKSP6d4	GCCCGG	79%
pCKSP6d3	GCCCCG	65%
pCK-X	GCGGCT	42%
pCK+X	GCGCCT	29%
pCKSP6d5	CCCGGG	19%
pCK+Xd	CGCCTA	4%

By comparing the relative transcriptional activities of wild-type pCKSP6 and these mutants, it was possible to further define the relative importance of individual base pairs. Two pairs of plasmids differ only at one base pair. When the altered base is at +4 (pCK-X and pCK+X) or at +5 (pCKSP6d3 and pCKSP6d4), there is a relatively small difference in the efficiency of transcription. Those plasmids which contain a G at position +1 have significantly higher activities than those with a C at this position. From these observations, it appears that the ability of a change to affect transcriptional efficiency decreases as one moves away from the start site +1. Using this assumption, similar comparisons for the two base pairs at positions +1 and +2

would give the order of relative activities, GG > GC > CC > CG; and for the first three base pairs (+1 to +3), GGA > GCC > GCG > CCC > CGC. In no case does a downstream mutation override an additional mutation occurring further upstream. For example, pCKSP6d3 differs from pCK-X at positions +3, +4 and +6. The change from a G to a C at +4, in the absence of any other change (pCK-X → pCK+X), results in a loss of transcriptional activity. The transversion from a G to a C at +3, however, leads to increased activity (GCG → GCC). The overall effect of this triple mutation is increased activity, consistent with the net effect being dominated by the alteration occurring closest to the +1 position. Thus, this limited set of mutants is entirely consistent with the assumption that the importance of nucleotide sequence in transcriptional activity decreases as one moves downstream from position +1 to position +6. Moreover, most mutants used in the above comparisons are transversions between G/C and C/G base pairs, and so the observed activity differences are probably not due to the melting temperature of the double-stranded initiation region of the DNA template. The melting temperatures of short DNA duplexes (from position -1 to +6) calculated from nearest-neighbor thermodynamic parameters recently determined by Breslauer et al. (18) are not correlated with the order of relative transcriptional activities observed with the corresponding mutants.

The SP6 consensus promoter sequence contains the characteristic "TATA" box at position -4 to -1 (Table 1). Surprisingly, when the A/T and T/A base pairs at positions -1 and -2 are altered to a C/G base pair (pCKSP6d5 → pCKSP6d6, and pCK+Xd → pCKSP6d7), while keeping the two base pairs at positions +1 and +2 constant (CC and CG, respectively), there is little change in the transcriptional activities (19% → 18%, and 4% → 3%, respectively). This indicates that the base pairs at positions +1 and +2 are much more important than those at positions -1 and -2 for the promoter function and/or transcription initiation.

DISCUSSION

(a) New transcription subcloning vector pCKSP6

It has been difficult to obtain a transcription vector

containing a restriction enzyme cleavage site allowing the insertion of a gene right at the initiation site (1). We report here the construction of such a transcription vector pCKSP6 for the SP6 *in vitro* transcription system (Fig. 1). When a gene is inserted at the unique BamHI site of pCKSP6, either through cohesive or blunt ends, the transcript starts at the initiation base G and is immediately followed by the sequence of the cloned gene. The ability to synthesize RNA with the authentic 5' end sequence efficiently from the cloned DNA makes it possible to apply various *in vitro* mutagenesis techniques to the studies of RNA structure-function relationships. Sufficient RNA, either wild type or mutants, could be synthesized and purified in this manner for biological as well as physical studies.

Another usage of pCKSP6 is the synthesis of eukaryotic messenger RNA. When a full length cDNA is inserted into the initiation site, a "capped" mRNA can be obtained by adding $m^7G(5')ppp(5')G$ and all four ribonucleotides in the SP6 transcription mixture (15). SP6 RNA polymerase is capable of transcribing poly(A) stretches of many nucleotides (1). Thus, pCKSP6 would be a useful cloning vector for eukaryotic mRNA studies.

(b) Cleavage site of restriction endonuclease HphI

The restriction enzyme from *Haemophilus parahaemolyticus* (HphI) recognizes the asymmetric five base pair sequence, 5'GGTGA3', and cleaves the DNA at a site approximately one turn of the helix away from the center of the recognition sequence, leaving a one base extension at the 3' end (12). According to the latest restriction enzyme compilation by Roberts (13) and New England Biolab (the sole commercial supplier of the enzyme), the cleavage is always eight base pairs away from the recognition site regardless of the base pair at the cleavage site.

Near the transcription initiation site of plasmid pSP64, however, we have found that the cleavage site is staggered about the A/T base pair at nine base pairs downstream from the recognition sequence. This is consistent with an earlier report by Kleid *et al.* (14) that the cleavage site is located eight or nine base pairs away from the recognition site, and staggered about either an A/T or a T/A base pair.

(c) Mutations by mung bean nuclease

The "nibbling" property of a single strand specific mung bean nuclease can be taken advantage of in creating deletion mutants that lack few base pairs around a specific site. A circular DNA can be linearized at the desired site, reacted with mung bean nuclease, which occasionally removes a few extra base pairs at the ends of DNA, and then recircularized. When the linearized DNA has 5' protruding ends, treatment with T4 DNA polymerase to make DNA ends blunt before or after mung bean nuclease will yield a series of mutants resulting from the stepwise deletion of single base pairs.

Other exonucleases that are currently used for sequential deletions, like Bal 31, are too potent to control the deletion extent to only few bases. Under the conditions employed in this study, deletions of one, two and three base pairs were obtained. Larger deletions can be obtained at a higher reaction temperature, since the extent of deletion is expected to depend on the stability of the edges of double-stranded DNA. Thus alone or in combination with T4 DNA polymerase, mung bean nuclease should provide a useful tool for a series of single-base deletion mutations.

(d) SP6 transcription promoter

The consensus promoter sequences are similar for the phage T3, T7 and SP6 DNA-dependent RNA polymerases (5,16,17) as listed in Table 1. The DNA sequence including the consensus promoter sequence immediately upstream of the initiation site is believed to be responsible for the observed stringent transcription specificity. It was of interest to see whether the sequence downstream of the initiation site was also important for the transcription.

Utilizing the new transcription vector pCKSP6 and mung bean nuclease, we have constructed a series of SP6 promoter mutants containing varied DNA sequences downstream from position -2 (Table I). When the relative template activities of these mutants are compared, it is apparent that the base pairs at positions +1 and +2 (and +3) are more important for transcription initiation than other downstream sequences, and even more important than the base pairs at positions -1 and -2, which are

part of the consensus promoter sequence. Thus, there appears to be a hierachy of importance of each base pair at the initiation site, in the order of position +1 > +2 > +3 > +4, +5, +6, -1, -2.

The conclusion presented above is in good agreement with the recent DNA footprinting data by Ikeda & Richardson (19). T7 RNA polymerase protects DNA sequence from -21 to +8 when trinucleotide (3-mer) RNA is synthesized. The formation of 6-mer RNA further extends the protected region from -21 to +11. From these observations, one can assume that when only the first ribonucleotide is bound, the protected region would be from -21 to +6. Our results show that the sequence downstream from +7 do not appear to be involved in initiation. In the absence of RNA synthesis, T7 RNA polymerase protects the DNA from -21 to only -3 as forming a closed colmplex of initiation (19). This may explain why mutations at positions -1 and -2 are silent in <u>in vitro</u> transcription, since they are not involved in initial binding of RNA polymerase.

During transcription initiation, a closed complex is converted to an open complex and the two strands of the DNA template in the downstream part of promoter sequence (-2 to +6) must be separated. Since A:T base pairs can melt more easily than G:C base pairs, one would predict that A:T to G:C mutants should disfavor the transcription initiation. Our results, however, do not support this hypothesis. Changes from A:T base pair to G:C at positions -1 and -2 do not affect initiation activity at all, while transversions between G/C and C/G base pairs at positions +1 through +5 do affect it significantly. The RNA polymerase-promoter interaction which initiates specific transcription appears to be a complex process that cannot be explained simply by melting in of the enzyme.

ACKNOWLEDGEMENTS

This work was supported by NIH grant GM 28069 and American Cancer Society grant NP 309.

Abbreviations: A/T, A in the non-coding strand base paired with T in the coding strand; A:T, A/T or T/A; bp, base pairs; BRL, Bethesda Research Laboratories, Inc.; BSA, bovine serum albumin; DNase, deoxyribonuclease; EtBr, ethidium bromide; RNase, ribonuclease.

*Present address: Department of Biological Sciences, Korea Advanced Institute of Science and Technology, Seoul 131, Korea

+To whom correspondence should be addressed

REFERENCES

1. Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K. and Green, M.R. (1984) Nucl. Acids Res. 12, 7035-7056.
2. Green, M.R., Maniatis, T. and Melton, D.A. (1983) Cell 32, 681-694.
3. Krieg, P.A. and Melton, D.A. (1984) Nucl. Acids Res. 12, 7057-7070.
4. Tabler, M. and Sanger, H.L. (1985) EMBO J. 4, 2191-2199.
5. Brown, J.E., Klement, J.F. and McAllister, W.R. (1986) Nucl. Acids Res. 14, 3521-3526.
6. van der Werf, S., Bradley, J., Wimmer, E., Studier, F.W. and Dunn, J.J. (1986) Proc. Natl. Acad. Sci. USA 83, 2330-2334.
7. Sollazzo, M., Frank, R. and Cesareni, G.(1985) Gene 37, 199-206.
8. Hanahan, D. (1983) J. Mol. Biol. 166, 557-580.
9. Birnboim, H.C. and Doly, J. (1979) Nucl. Acids Res. 7, 1513-1523.
10. Butler, E.T. and Chamberlin, M.J. (1982) J. Biol. Chem. 257, 5772-5778.
11. Bogenhagen, D.F., Sakonju, S. and Brown, D.D. (1980) Cell 19, 27-35.
12. Kleid, D.G. (1980) Methods Enzymol. 65, 163-166.
13. Roberts, R.J. (1985) Nucl. Acids Res. 13 supplement, r165-r200.
14. Kleid, D.G., Humayun, Z., Jeffrey, A. and Ptashne, M. (1976) Proc. Nat. Acad. Sci. USA 73, 293-297.
15. Contreras, R., Cheroute, W., Degrave, W. and Fiers, W. (1982) Nucl. Acids Res. 10, 6353-6362.
16. Dunn, J.J. and Studier, F.W. (1983) J. Mol. Biol. 166, 477-535, and erratum (1984) J. Mol. Biol. 175, 111-112.
17. McGraw, N.J., Bailey, J.N., Cleaves, G.R., Dembinski, D.R., Gocke, C.R., Joliffe, L.K., MacWright, R.S. and McAllister, W.T. (1985) Nucl. Acids Res. 13, 6753-6766.
18. Breslauer, K.J., Frank, R., Bloecker, H. and Marky, L.A. (1986) Proc. Natl. Acad. Sci. USA 83, 3746-3750.
19. Ikeda, R.A. and Richardson, C.C. (1986) Proc. Natl. Acad. Sci. USA 83, 3614-3618.