
Essential structure of *E. coli* promoter II. Effect of the sequences around the RNA start point on promoter function

Takashi Aoyama* and Mituru Takanami

Institute for Chemical Research, Kyoto University, Uji, Kyoto 611, Japan

Received 9 October 1984; Revised and Accepted 13 May 1985

ABSTRACT

Starting from a synthetic *E. coli* promoter with the consensus sequences at -35 and -10 regions, a sequence CAT frequently occurred in the RNA start points of natural promoters was introduced in the downstream of the consensus sequences, and the sequences around the RNA start points as well as the relative positions of CAT from the consensus sequences were altered. Analysis of the RNA start points and strength of these synthetic promoters in an *in vitro* transcription system provided evidence that the RNA start point was principally fixed by distance from the -10 consensus sequence (TATAAT). Neither the promoter strength nor the RNA start point was significantly influenced by the CAT sequence. The sequences around the RNA start points rather seemed to exert influence on the response of promoter to temperature and salts.

INTRODUCTION

Comparative analysis of *E. coli* promoter sequences and accumulation of promoter mutants have indicated that two hexanucleotide sequences, centered around -35 and -10 positions with respect to the RNA start point, are extremely important for interaction with RNA polymerase(1-5). Although the sequences at these two regions fluctuate with promoter, most frequently occurring sequences at -35 and -10 regions are TTGACA and TATAAT, respectively, and have been named "consensus sequences". Nevertheless, information is still not enough to predict the site and strength of promoter from a given sequence. There are considerable amounts of evidence suggesting that the sequences in other regions also play important roles for promoter function (6-9). It has been noted that a sequence CAT frequently occurs in the RNA start points of natural promoters, but little is known about the functional role of this sequence.

In order to define the structure directing promoter function more precisely, we initiated a series of investigation by constructing an ideal promoter having the two consensus sequences at the -35 and -10 regions (10). We first examined the effect of spacing between the two consensus

sequences, and found that not only the sequences but also the spacing of the two consensus sequences are important determinant of promoter function(10). In this paper, we investigated the effect of the sequences around the start point of transcription on promoter function by altering the sequences systematically.

MATERIALS AND METHODS

Synthetic oligonucleotides and enzymes

Two septamers, dGTCTAGA and dTCTAGAC, and four decamers, dTCGAGGGCAT, dCTAGATGCCC, dTCGAGAACAT and dCTAGATGTTC, were provided by Takara Shuzo Co. RNA polymerase was products of New England Biolabs. All other enzymes were purchased from Takara Shuzo Co. and Bethesda Res. Lab.

Strains and plasmids

E.coli C600 and GM33(dam⁻) were used as recipients. Plasmids pPSC7, pSPI57 and pPSC8 containing synthetic promoters with the two consensus sequences, TTGACA and TATAAT, at -35 and -10 regions has been constructed previously(10).

Preparation of plasmids and restriction fragments

Plasmids and restriction fragments were prepared according to the methods described previously(10), except that plasmids were propagated in E.coli GM33 by the reason that XbaI recognizing unmethylated restriction sequences was used in this study.

In vitro RNA synthesis and analysis of products

The reaction mixtures used for RNA synthesis were composed of 50mM tris-HCl (pH7.9), 8 mM MgCl₂, 150 mM KCl, 0.1 mM dithiothreitol, 0.1 mM (α -³²P)UTP, and 0.4 mM each of three other NTPs, and when examined the effect of salts, KCl concentrations were altered. For determination of RNA start points and relative activity of promoter, RNA synthesis was done for 15 min at 37°C in the reaction mixtures(0.1ml each) containing 1 pmol DNA and 10 pmol RNA polymerase. When the influence of temperature and salts on promoter function were examined, two sets of reaction mixtures(20 μ l each) were prepared; one containing 0.25 pmol DNA fragments and the other containing 5 pmol RNA polymerase. Both sets of the reaction mixtures were pre-incubated for 5 min at different temperatures(for the temperature effect) or at 37°C(for the salt effect), and then mixed to initiate RNA synthesis. After holding 1 min at the pre-incubation temperatures, heparin was added at 100 μ g/ml to prevent re-initiation(11), and RNA synthesis was

allowed to continue at 37°C for another 15 min. In all cases, reaction was terminated by phenol treatment. Synthesized products were precipitated by ethanol, dissolved in 80% formamide-dye, heat-quenched for 3 min at 90°C and analyzed by sequencing gels.

Other methods

Conditions for T4 DNA polymerase reaction, S1 nuclease digestion and ligation have been described previously(10). DNA sequencing was carried out by the dideoxy method which uses supercoiled plasmid DNA as template(12). S1 mapping was performed by the method of Berk and Sharp(13).

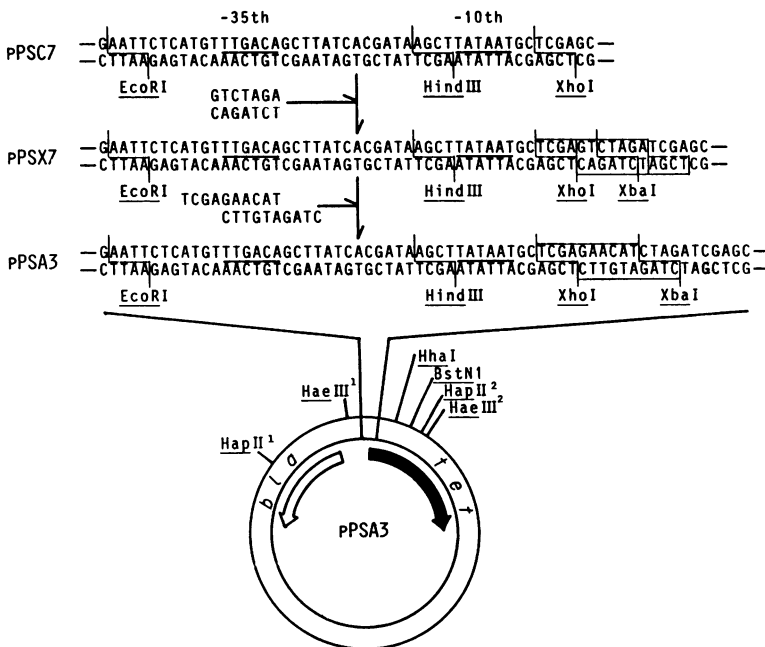


Fig.1 Construction of promoters having different sequences in the downstream from the -10 consensus sequence. pPSC7 is a derivative of pBR322 carrying a synthetic promoter(10). The promoter sequences, the insertion sites in the plasmid, and the steps of sequence modification are schematically shown. The -35 and -10 consensus sequences were indicated by underlines, and inserted sequences in derivatives were boxed. When constructed pPSG series, the indicated decamer duplex was replaced by the other decamer-duplex(see text). Restriction sites indicated on the circular map of pPSA3 are those used for generation of fragments. The HapII¹-BstNI fragment and HaeIII¹-BstNI fragment were used as template for *in vitro* transcription, and the HhaI-HapII² fragment was used as primer for dideoxy-supercoil-sequencing.

RESULTS

Construction of promoters having different sequences in RNA start points

The synthetic promoter carried by plasmid pPSC7 has the two consensus sequences at an optimal distance(17 base-pairs), and exhibits a strong activity in an in vitro transcription system(10). Starting from this promoter, the sequences around the RNA start point were altered according to the scheme shown in Fig.1. Plasmid pPSC7 was linearized by XhoI, and after repairing by T4 DNA polymerase, the synthetic duplex, dGTCTAGA:dTCTAGAC, was introduced by blunt-end ligation. As a consequence, plasmid, pPSX7, having a XbaI site at the outside of the XhoI site was constructed. Then, two kinds of synthetic duplexes, dTCGAGGGAT:dCTAGATGCC and dTCGAGAACAT:dCTAGATGTTC, were inserted between the XhoI and XbaI sites. The resulting two plasmids, pPSG3 and pPSA3, contained promoters carrying the CAT sequence at the same position, but differing in GC-contents of the sequences around the RNA start points.

As the next step, insertion and deletion were introduced at the XhoI sites of pPSG3 and pPSA3 by the following steps: linearization by XhoI digestion, partial repairing of the resulting ends by T4 DNA polymerase and appropriate dNTPs, conversion of the ends to blunt-ends by S1 nuclease treatment, and circularization by blunt-end ligation(10). The products were digested by XhoI to destroy any intact molecules and transformation was carried out. Plasmids were isolated from randomly selected clones, and the sequences around the modified sites were

Plasmids	Sequences	Relative Activities
pPSA303	-AAGCTTATAATGCTCGCGAGAACATCT-	10.5
pPSA3	-AAGCTTATAATGCTCGAGAACATCTAG-	12.2
pPSA323	-AAGCTTATAATGCTGAGAACATCTAGA-	11.6
pPSA213	-AAGCTTATAATGCGAGAACATCTAGAT-	14.2
pPSA302	-AAGCTTATAATCGAGAACATCTAGATC-	9.5
pPSA403	-AAGCTTATAATGCGAACATCTAGATCG-	13.2
pPSG10	-AAGCTTATAATGCTCGCGAGGGCATCT-	11.1
pPSG3	-AAGCTTATAATGCTCGAGGGCATCTAG-	10.0
pPSG105	-AAGCTTATAATGCTCAGGGCATCTAGA-	7.2

Fig.2 Sequences, initiation points and relative activities of constructed promoters. The -10 consensus sequences were boxed, and start points of transcription were indicated by asterisks. Relative activities of promoters to the bla promoter are given in the right column.

determined by dideoxy-supercoil-sequencing. As the result, a series of promoters having different sequences in the RNA start points were obtained(Fig.2). The relative positions of the CAT sequence from the consensus sequences are also different with promoters.

Activity of constructed promoters

The constructed promoter was cut out together with the bla gene promoter by HapII and BstNI as a fragment of about 600 base pairs, and the promoter activity was determined in the in vitro transcription system. Products synthesized are shown in Fig.3 A, in which RNA-1 corresponds to the transcript from the constructed promoter and RNA-2 to the transcript from the bla promoter, respectively. This was confirmed by analysis of truncate transcripts from DNA fragments shortened by digestion with other enzymes. Analysis was performed by using 0.3mm-thin 8% polyacrylamide sequencing gel, so that resolution was high enough to determine the product sizes by one-base length, and the relative positions of the RNA start points could be assigned on the template from the product sizes. Three bands having slightly different mobility are seen at the RNA-2 position in each lane of Fig.3 A. Analysis of truncated products indicated that all of them originated from a single site at the bla gene promoter, and appearance of three bands was attributed to the variation of the RNA stop position at the HapII terminus. However, this was not the case for RNA-1, which was transcribed in the reverse direction. This is not surprising, for RNA stop positions are markedly influenced by stability of the helical structure at termini(14).

Radioactivities of RNA-1 and RNA-2 in each lane were measured by Cerenkov counting of the bands, and the values of RNA-1 corrected for the U contents were compared by molar ratios relative to those of RNA-2. The result is given in the right column in Fig.2. The levels of the constructed promoters were not significantly different, except for the one on pPSG105. This promoter differed from the pPSG3 promoter only by one base, but it gave a low value beyond the measurement errors. As discussed later, we assume that this is due to the presence of a G cluster in the RNA start point. The data in Fig.2 reveal that the position of the CAT sequence bears little relation to the level of promoter.

RNA start points of constructed promoters

In order to allocate the start point of transcription precisely, S1 mapping was carried out. The transcript from the HaeIII fragment of pPSA303 (see Fig.1) was hybridized to the terminal labeled HapII fragment from pPSA303.

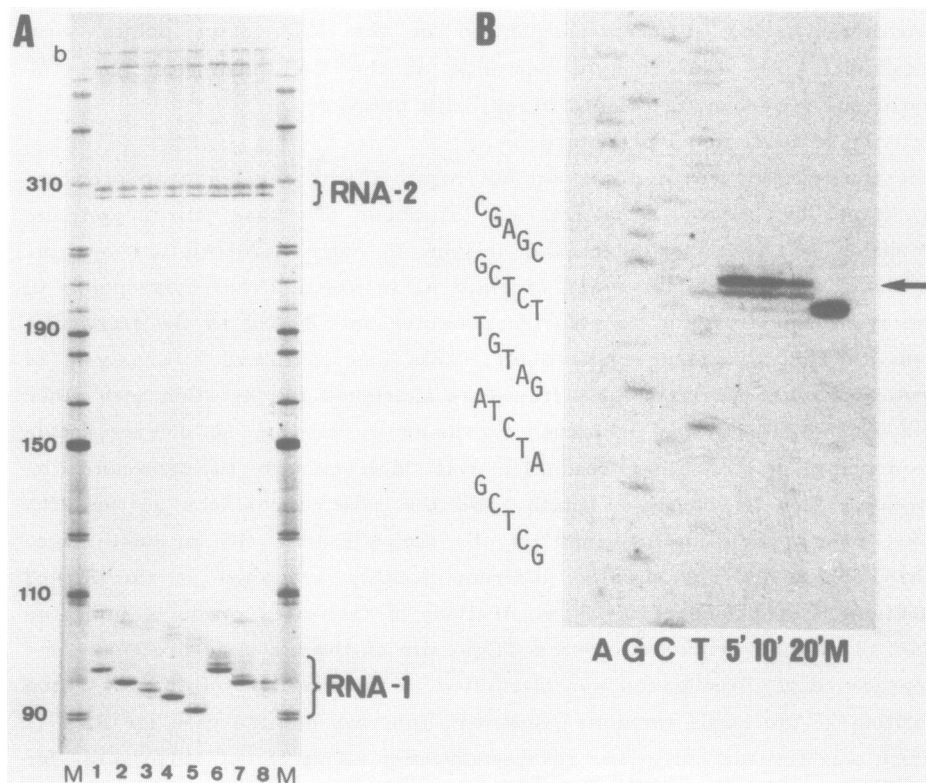


Fig.3 (A) Transcripts formed on HapII¹-BstN1 fragments from constructed plasmids. Products from pPSA303, pPSA3, pAPSA323, pPSA213, pPSA302, pPSG10, pPSG3 and pPSG105 are shown in lanes 1 to 8 (see Fig.2 for plasmid names). Marker fragments are shown in lanes M. The band regions for RNA-1 (transcripts from constructed promoters) and RNA-2 (transcripts from the bla gene promoter) were indicated by the side of lanes. RNA-1 and RNA-2 are transcribed in different directions. A few bands seen above RNA-2 are read-back transcripts. (B) S1 mapping of transcripts formed on the HaeIII fragment generated from pPSA303. Protected bands at different periods (min) of S1 digestion and the XhoI-HapII² fragment as a fragment maker (M) are given in right four lanes. The major protected band is shown by arrow. Left four lanes are sequence ladders, obtained from pPSA303 by dideoxy-supercoil-sequencing. The primer used is the HhaI-HapII² fragment labeled at the 5'-terminal of the HapII end.

After S1 nuclease treatment, protected DNA was analyzed by sequencing gels. The result is shown in Fig.3 B. The protected bands at different periods of S1 digestion and the XhoI-HapII² fragment (see Fig.1) as a fragment marker are shown in right four lanes. Left four lanes are

Plasmids	Sequences
pPSA303-I57	-GT TTGACA GCTTATCATCGATAAGCT TATAAT GCTCGC G *AGAA-
pPSA303	-TGT TTGACA GCTTATCACGATAAGCT TATAAT GCTCGC G *AGAA-
pPSA303-C8	-ATGT TTGACA GCTTATCATATAAGCT TATAAT GCTCGC G *AGAA-
pPSA3-I57	-GT TTGACA GCTTATCATCGATAAGCT TATAAT GCTCGC G *AACA-
pPSA3	-TGT TTGACA GCTTATCACGATAAGCT TATAAT GCTCGC G *AACA-
pPSA3-C8	-ATGT TTGACA GCTTATCATATAAGCT TATAAT GCTCGC G *AACA-

Fig.4 Sequences and initiation points of spacer-mutant promoters. The -35 and -10 consensus sequences were boxed, and the start points of transcription were indicated by asterisks.

sequence ladders of pPSA303 by dideoxy-supercoil-sequencing. From the relative position to markers, the prominent protected band was assigned to the C residue of which the position is two bases longer than the *XhoI-HapII*² fragment, indicating that the constructed promoter of pPSA303 initiated transcription at the complementary G residue. The start points of the other promoters were assigned by positions relative to that of pPSA303. The starting points assigned are shown by asterisks in **Fig.2**. The validity of these assignments was confirmed by other experiments, in which only products starting with A were labeled by using (γ -³²P)ATP (data not shown). In 9 out of 11 promoters, initiation occurred by Pu(purine) at the 7th position from TATAAT, regardless of the sequences around the initiation sites. Initiation of the remaining two from pPSA302 and pPSG105 occurred by Pu at the 8th and 6th positions, respectively. The pPSA302 promoter contains Py(pyrimidine) at the 7th position, so that it is likely that RNA starts at the neighboring Pu, if no Pu is present at the 7th position. Unexplained is the case of the pPSG105 promoter containing Pu at the 7th position. As mentioned in the previous section, activity of this promoter was also low. We assume that the abnormality is due to the occurrence of a G-cluster around the RNA start point.

In order to determine by which of the consensus sequences the starting point of transcription was determined, pPSA3 and pPSA303 in this paper were recombined with pSPI57 and pPSC7 previously constructed (10) at the *HindIII* site between -35 and -10 regions, and two sets of spacer-mutant promoters shown in **Fig.4** were constructed. Transcripts from these mutant promoters were analyzed as in the experiment in **Fig.3**. The data are shown in **Fig.5**, as indicated by asterisks in **Fig.4**, RNA synthesis was initiated in the constant position from the -10 consensus sequence at -10 region.

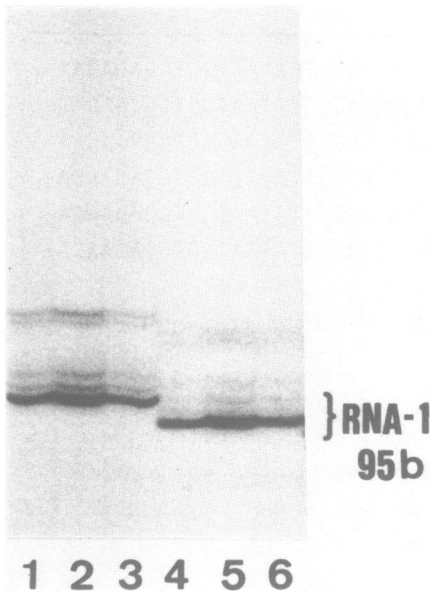


Fig.5 Transcripts directed by spacer-mutant promoters listed in Fig.4. Patterns in the RNA-1 band regions are shown. Products from HapII¹-BstN1 fragments of pPSA303-I57, pPSA303, pPSA303-C8, pPSA3-I57, pPSA3 and pPSA3-C8 are shown in lanes 1 to 6.

Effect of the sequences around the RNA start points on properties of promoter

Dependency of promoter function on temperature and ionic strength were compared between the promoters on pPSA3 and pPSG3 or between those on pPSA3 and pPSG10. The promoters in these plasmids initiate RNA synthesis at the 7th position with about the same efficiencies, but the GC contents of the sequences around the start points are significantly different each other(see Fig.2). The HaeIII¹-BstN1 fragments carrying only the constructed promoter were generated from these plasmids, and RNA synthesis was carried out either at different temperatures or at different salt concentrations(KCl). Other reaction conditions, including the amounts of template and enzyme, were equalized as much as possible, and the promoter strengths were compared by measuring the amounts of RNA initiated within a definite time of incubation.

The major band formed on the template was RNA-I of about 95 bases long, but a few additional bands of read back-transcripts were formed especially at low salts. The formation of such read-back transcripts is due to the switching of the template strand by RNA polymerase at the end of the restriction fragment, and greatly depends on the structure of termini(14). Relative amounts of RNA were estimated by Cerenkov counting of respective bands on gels, and correction for read-back transcripts was

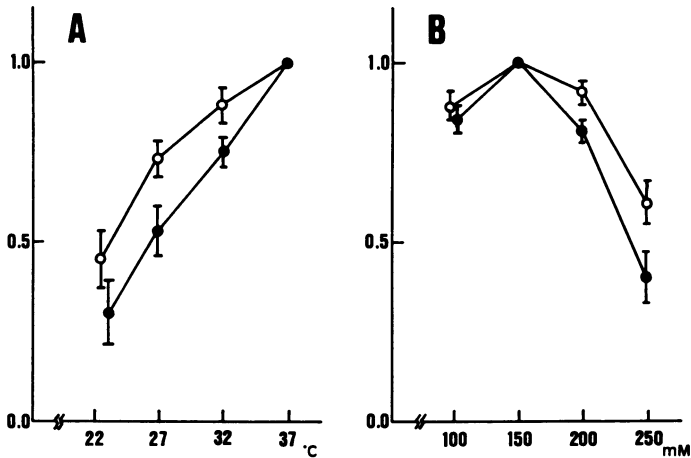


Fig.6 Effects of temperature and salt-concentrations on activities of promoters from pPSA3(O) and pPSG3(●). *Hae*III¹-*Bst*NI fragments were generated from these plasmids, and RNA synthesis was carried either at indicated temperatures(A) or at indicated KCl concentrations(B) under the conditions described in the text. The data were compared by normalizing to the values at 37°C(A) or to the value at 0.15 M KCl(B). Vertical bars represent standard deviations in the experiments.

made on the basis of U-contents as labeling was done with (α -³²P)UTP. In **Fig.6**, the amounts of RNA formed with promoters on pPSA3 and pPSG3 at different temperatures(**Fig.6 A**) or at different salts(**Fig.6 B**) are shown by normalizing the values to those at 37°C or those at 0.15 M KCl, respectively. Compared with the promoter on pPSA3, that on pPSG3 seemed to be more sensitive to temperature and salts. A similar correlation was observed between the promoters on pPSA3 and pPSG10, and that of pPSG10 was more intensely depressed at low temperature and at high salts(data not shown).

DISCUSSION

A sequence CAT frequently appears in the RNA start points of natural promoters, so that this sequence has been defined as the consensus sequence in the RNA start point(1,2,4). In this paper, we constructed a series of promoters having CAT in different positions from the -10 region. However, neither the promoter strength nor the RNA starting point was significantly influenced by the CAT sequence. Our data rather demonstrate that the -10 consensus sequence is the determinant of the RNA start point,

Table 1 Start points of transcription in 85 known promoters having different bases(Pu or Py) at the 7th position from TATAAT. Occurrence of Pu and Py at indicated positions (number from TATAAT) and number of promoters initiating by boxed bases were indicated. Data were taken from ref.4.

Bases at 7th position (Total Number)	6th	7th	8th	9th	Number of Initiation
	(Positions from TATAAT)				
P u (4 2)	.	Pu	.	.	3 1
	Pu	Pu	.	.	3
	.	Pu	Pu	.	4
	.	Pu	.	Pu	4
P y (4 3)	.	Py	Pu	.	2 6
	Pu	Py	.	.	5
	Py	Py	Py	Pu	6
	Py	Py	Py	.	3
	Py	Py	Py	.	1
	Pu	Py	Py	.	2

and that initiation predominantly occurs by Pu base at the 7th position from the -10 consensus sequence, TATAAT. Based on these observations, the start points of known promoters for *E.coli* RNA polymerase were compared according to the data compiled by Hewley and McClure(4). The data are summarized in Table 1. Among 85 promoters of which the initiation sites had been assigned, occurrences of Pu and Py bases at the 7th position are 42 and 43, respectively. Most of the promoters having Pu at the 7th position initiate at this position, but the remaining promoters initiate by Pu at adjacent positions. Although little is known about the mechanism of such exceptions, we noted that initiation by Pu at other positions often occur if a stretch of the same Pu base encompasses the 7th position. This is also the case of the constructed promoter in pPSG105, which initiates RNA synthesis by G at the 6th position within the G cluster. On the other hand, initiation of promoters having Py bases at the 7th position mostly occurs by Pu at adjacent positions, especially at the 8th position. The constructed promoter in pPSA302 initiating at the 8th position belongs to this category (see Fig.2). Only a few promoters initiate by Py bases, but none of them contain Pu at the 7th and 8th positions. Combining with the result obtained in this paper, we deduced a general conclusion that initiation principally occurs by Pu base at the 7th position from the -10 consensus sequence and that if Pu base is not present at the 7th position, RNA synthesis is initiated by Pu bases at adjacent positions.

Since the sequences around the RNA start points did not exert influence on the promoter strength significantly, we examined the sequence effect on other properties of promoters, and found that promoters having relatively GC-rich sequences in the RNA start points are more sensitive to temperature and salts. It has generally been thought that unwinding of the DNA helix is a prerequisite to expose the template bases for initiation of transcription(15,16). In fact, local melting of the promoter region from positions -9 and +3 by binding of RNA polymerase has been demonstrated from the sensitivity to modification agents and S1 nuclease(2,17). Dependency of promoter on temperature and salt concentrations could be influenced by the GC-contents of the region around the RNA start point, as GC-contents and ionic strength relate to thermal stability of DNA helix(18).

The entire promoter regions constructed here can be generated by EcoRI and XbaI, so that these promoters would be useful for manipulation of transcription by inserting them into appropriate sites.

ACKNOWLEDGEMENTS

This work was supported by research grants from the Ministry of Education, Science and Culture of Japan.

*To whom correspondence should be addressed

REFERENCES

1. Rosenberg, M. and Court, D. C. (1979) *Ann. Rev. Genet.* 13, 319-353.
2. Siebenlist, U., Simpson, R. B. and Gilbert, W. (1980) *Cell* 20, 269-281.
3. Youderian, P., Bouvier, S. and Susskind, M.M. (1982) *Cell* 30, 843-853.
4. Hawley, D. K. and Mc Clure, W. R. (1983) *Nucleic Acids Res.* 11, 2237-2255.
5. von Hippel, P. H., Bear, D. G., Morgan, W. D. and McSwiggen, J. A. (1984) *Ann. Rev. Genet.* 53, 389-446.
6. Horn, G. T. and Wells, R. D. (1981) *J. Biol. Chem.* 256, 2003-2009.
7. Klein, R. D. and Wells, R. D. (1982) *J. Biol. Chem.* 257, 12954-12961.
8. Munson, L. M., Mandeck, W., Caruthers, M. H. and Reznikoff, W. S. (1984) *Nucleic Acids Res.* 12, 4011-4017.
9. Travers, A. A. (1984) *Nucleic Acids Res.* 12, 2605-2617.
10. Aoyama, T., Takanami, M., Ohtsuka, E., Taniyama, Y., Marumoto, R., Sato, H. and Ikehara, M. (1983) *Nucleic Acids Res.* 11, 5855-5864.
11. Walter, G., Zillig, W., Palm, P. and Fuchs, E. (1967) *Eur. J. Biochem.* 3, 194-201
12. Wallace, R. B., Johnson, M. J., Suggs, S. V., Miyoshi, K., Bhatt, R. and Itakura, K. (1981) *Gene* 16, 21-26.
13. Berk, A. J. and Sharp, P. A. (1977) *Cell* 12, 721-732.
14. Sugimoto, K., Sugisaki, H., Okamoto, T. and Takanami, M. (1977) *J. Mol. Biol.* 111, 487-507.

15. Wang, J. C., Jacobsen, J. H. and Saucier, J. -M. (1977) *Nucleic Acids Res.* 4, 1225-1241.
16. Melnikova, A., Beabealashvili, R. and Mirzabekov, A. D. (1978) *Eur. J. Biochem.* 84, 301-309.
17. Siebenlist, U. (1979) *Nature* 279, 651-652.
18. Marmur, J. and Doty, P. (1962) *J. Mol. Biol.* 5, 109-118.