

Identification of Bacteriophage K11 Genomic Promoters for K11 RNA Polymerase

Kyung Goo Han, Dong Hee Kim, Eunsung Junn[†], Sang Soo Lee* and Changwon Kang^{†,*}

Research Center for Bio-Medicinal Resources and Division of Life Science, Pai-Chai University, Daejeon 302-735, Korea

[†]Department of Biological Sciences, Korea Advanced Institute of Science and Technology, Daejeon 305-701, Korea

Received 11 October 2002, Accepted 23 October 2002

Only one natural promoter that interacts with bacteriophage K11 RNA polymerase has so far been identified. To identify more, in the present study restriction fragments of the phage genome were individually assayed for transcription activity *in vitro*. The K11 genome was digested with two 4-bp-recognizing restriction enzymes, and the fragments cloned in pUC119 were assayed with purified K11 RNA polymerase. Eight K11 promoter-bearing fragments were isolated and sequenced. We report that the nine K11 promoter sequences (including the one previously identified) were highly homologous from -17 to +4, relative to the initiation site at +1. Interestingly, five had -10G and -8A, while the other four had -10A and -8C. The consensus sequences with the natural -10G/-8A and -10A/-8C, and their variants with -10G/-8C and -10A/-8A, showed nearly equal transcription activity, suggesting residues at -10 and -8 do not regulate promoter activity. Using hybridization methods, physical positions of the cloned promoter-bearing sequences were mapped on *SalI*- and *KpnI*-restriction maps of the K11 genome. The flanking sequences of six cloned K11 promoters were found to be orthologous with T7 or T3 genomic sequences.

Keywords: K11 genome, K11 promoters, K11 RNA polymerase, Phage promoter consensus sequences, Sequence logos

Introduction

The bacteriophage K11 RNA polymerase is produced from the phage genome at an early stage of phage K11 infection of *Klebsiella* sp. 390 (03 : K11), and carries out transcription of

the middle (class II) and late (class III) phage genes (Dietz *et al.*, 1990). Bacteriophage K11 DNA-dependent RNA polymerase is similar to RNA polymerases encoded by bacteriophages T7, T3 and SP6, in that it consists of a single polypeptide with an approximate molecular weight of 100 kDa. Although these phage RNA polymerases recognize their own specific promoters, their promoter consensus sequences of 20-23 base pairs share high degrees of homology. The base pairs discriminating the promoters between phage types were identified through studies of cross-reactions between T7 and T3 transcription systems (Klement *et al.* 1990), and T7 and SP6 systems (Lee and Kang, 1993).

Eighteen T7 promoters, 14 T3 promoters and 11 SP6 promoters have been found and have facilitated the study of promoter-polymerase interactions (Lee and Kang, 1993). However, so far only one promoter has been identified that interacts with K11 RNA polymerase (Dietz *et al.*, 1990). Phage transcription systems have proved useful in expressing exogenous genes (Eun *et al.*, 2001; Roytrakul *et al.*, 2001). Unlike other polymerases, the K11 polymerase has a unique 20-amino-acid insertion at the N-terminus (Dietz *et al.*, 1990). It would be useful to identify more K11 promoters and study their interactions with the K11 polymerase, and compare these with other phage transcription systems.

In this study, eight new K11 promoters were isolated from the genome and characterized. Surprisingly, two prominent sequences were found in the promoter specificity region, showing equal promoter strength. The approximate positions of eight promoters were also mapped on the phage K11 genome.

Materials and Methods

Materials *E. coli* XL1-blue was used as the host strain for plasmid transformation, and *E. coli* BL21 for expression of the gene for phage K11 RNA polymerase cloned in pET-tacK11 (Han *et al.*, 1999). *Klebsiella* sp. 390 (03 : K11) and K11 phage were provided by R. Hausmann (Freiburg University, Germany). T4 DNA ligase

*To whom correspondence should be addressed.

Tel: 82-42-520-5616 (ssl) or 82-42-869-2628 (ck)

Fax: 82-42-520-5616 (ssl) or 82-42-869-2682 (ck)

E-mail: ssllee@mail.paichai.ac.kr or ckang@mail.kaist.ac.kr

and restriction endonucleases were obtained from New England Biolabs (Beverly, USA). All ribonucleotides, [α - 32 P]CTP (3,000 Ci/m mole) and [α - 35 S]dATP were from Amersham Pharmacia Biotech (Uppsala, Sweden).

Cloning and sequencing of K11 promoters After the K11 genome DNA was extensively digested by *RsaI* and *Tsp509I*, the fragments were inserted into the *SmaI* and *EcoRI* sites of pUC119, respectively. Recombinant plasmids having disrupted lacZ' were selected from the transformants, and named the pKR# and pKT# series, respectively. The recombinant plasmids were assayed for transcription activity with purified K11 RNA polymerase. Standard reaction mixtures contained 40 mM Tris-HCl, pH 7.9, 6 mM MgCl₂, 10 mM dithiothreitol, 4 mM spermidine-HCl, 0.5 mM each of ATP, GTP, CTP and UTP, 5 units RNasin, 2 μ Ci [α - 32 P]CTP, 1 μ g plasmid DNA, and 10 units K11 RNA polymerase. Incorporation of 32 P into RNA was measured by insolubility in 5% trichloroacetic acid (Kang and Wu, 1987). Cloned genomic insert fragments showing K11 transcription activity were sequenced using the Sanger method.

Genomic mapping of K11 promoters Cloned K11 promoters were physically mapped on the phage genome using Southern hybridization against *SalI*- and *KpnI*-digested fragments of phage DNA. One μ g K11 genomic DNA was cleaved with *SalI* and *KpnI*, fractionated on an agarose gel (1%), denatured and then transferred onto a nitrocellulose filter (Schleicher & Schuell, Einback, Germany). The K11 promoter-bearing clones were linearized and transcribed *in vitro* with K11 RNA polymerase, and the radiolabeled transcripts were used as probes.

Measurement of promoter strength In order to measure K11 promoter strength, run-off transcription assays were performed as previously described by Kang and Wu (1987). The reactions were stopped after 1 h incubation by adding 10 μ l gel-loading buffer containing 80% (w/v) deionized formamide and 10 mM EDTA (pH 8.0). The transcripts were separated by 6% polyacrylamide-8 M urea gel electrophoresis. Radioactivity of run-off transcript bands was counted using a Fuji phosphorimage analyzer Bas-1500.

Consensus sequence promoters and variants Two duplex DNA were constructed containing the natural consensus sequence having -10G/-8A and -10A/-8C. They covered the region from -20 to +5 and had 4-nt overhangs for *BamHI* and *HindIII* at the upstream and downstream ends, respectively.

Wild-type No. 1 (-10G and -8A)

5'-GATCAATAATTAGGGCACACTATAGGGAG-3'
3'-TTATTAATCCCCTGTGATATCCCTCTCGA-5'

Wild-type No. 2 (-10A and -8C)

5'-GATCAATAATTAGGACCCACTATAGGGAG-3'
3'-TTATTAATCCTGGGTGATATCCCTCTCGA-5'

Likewise, two variants with -10G/-8C and -10A/-8A were also constructed.

Mutant-type No. 1 (-10G and -8C)

5'-GATCAATAATTAGGGCCCACTATAGGGAG-3'
3'-TTATTAATCCCGGGTGTGATATCCCTCTCGA-5'

Mutant-type No. 2 (-10A and -8A)

5'-GATCAATAATTAGGACACACTATAGGGAG-3'
3'-TTATTAATCCTGTGTGATATCCCTCTCGA-5'

The eight deoxyoligonucleotides shown above were synthesized using an automatic DNA synthesizer. Complementary strands in a 1 : 1 mixture were annealed and sub-cloned into *BamHI* and *HindIII* sites of pUC118, as previously described (Lee and Kang, 1993).

Results

Isolation of novel K11 promoters and their sequences

Restriction fragments of the phage K11 genome were assayed individually for the presence of K11 promoters. The genomic DNA was extensively cleaved with 4-bp-recognizing restriction enzymes *RsaI* and *Tsp509I*, and the fragments cloned in pUC119 were assayed for transcription with purified K11 RNA polymerase. Eight recombinant plasmids that yielded strong K11 transcription were selected, and their inserts of K11 DNA were found to range from 0.25 to 2.1 kb. All inserts were sequenced for 400-500 base pairs at both ends, and all had a sequence homologous to the known K11 promoter sequence. Each cloned insert had only one promoter sequence, based on the transcription assay results described below. The promoter sequences are shown in Table 1.

Using the eight newly identified promoter sequences, plus the one already known, a consensus sequence was deduced (Table 1), and sequence logos are shown in Fig. 1. The consensus sequence is a stretch of 21 base pairs, from -17 to +4 with respect to the transcription start site +1. Interestingly, no particular base was predominant at -10 and -8. The -10 residues of the K11 promoters were either G (56%) or A (44%), and the -8 residues were either A (56%) or C (44%). However, the two residues appeared to be linked, at least in these nine promoters. Five promoters had -10G and -8A, and four had -10A and -8C, while no promoters had -10G/-8C or -10A/-8A.

Locations of phage promoters on the K11 genome

In order to locate the newly identified K11 promoters on the phage genome, radiolabeled RNA transcripts were produced from the promoter-bearing plasmids, and used as probes in hybridization assays against various digestion fragments of the genome. The plasmids were linearized such that the transcripts have only K11 genomic sequences, and do not include sequences of parent plasmid pUC119.

The *SalI* and *KpnI* restriction maps of the phage K11 genome were previously reported (Dietz *et al.*, 1985). The nine *SalI* fragments were named A to I in the order of size, and the 10 *KpnI* fragments, named a to j, as shown in Fig. 2. The radiolabeled transcripts were hybridized to the *SalI* and *KpnI* fragments of the K11 genome. By examining which genomic fragments were hybridized with a transcript probe, we were able to determine the ranges of possible locations of all the inserts on the genome, as shown in Fig. 2. The ranges are broader than the insert sizes because we did not attempt to

Table 1. Promoters for K11 RNA polymerase

Plasmid	Nucleotide sequence ^a
Previously known ^b	TCAGGCAATTAGGG <u>C</u> ACTATAGGGAACTTCGA
pKT55 (0.3 kb) ^c	TGTACAAATTAGG <u>A</u> CCACTATCAGGAACATACTC
pKR54 (1.8 kb)	GGTAAACATTAGGG <u>C</u> ACTACAGGGTCTGTTAAT
pKR25 (2.1 kb)	AGATGTAATTAGG <u>A</u> CCACTATAGGAGACACCAAA
pKT68 (0.5 kb)	CCAGTGAATTAGG <u>A</u> CCACTATAGGGAAGTGCCCA
pKT54 (0.4 kb)	AGCTCGAATTAGGG <u>C</u> ACTATAGGGAGACCATCA
pKT56 (1.1 kb)	CCAGTGAATTAGGG <u>C</u> ACTATAGGGAGACCACAC
pKR24 (1.0 kb)	GCATGTAATTAGGG <u>C</u> ACTATAGGGAGACCGATT
pKR2 (0.5 kb)	ACTTGGAAATTAGG <u>A</u> CCACTATAGGGAGAGACCT
Consensus ^d	-----aATTAGG <u>g</u> CaCACTAtagGga----- -----aATTAGG <u>a</u> CcCACTAtagGga-----

^aThe nucleotide sequences of the upper strands from -23 to +12 are shown. The transcription start site is numbered +1. Bases different from the conserved from -17 to +4 are underlined.

^bOne previously known promoter-bearing fragment was cloned (1).

^cThe approximate sizes of the K11 DNA fragments cloned in pUC119 shown in parentheses were determined by measuring the sizes of *Pvu*II-digested plasmids or the products of polymerase chain reaction with universal and reverse primers.

^dThe bases occurring in more than 66% of the non-coding strand sequence are shown by lower case letters, while those 100% conserved are shown by capital letters. Although no particular base is predominant at -10 and -8, all promoters have either -10G/-8A or -10A/-8C.

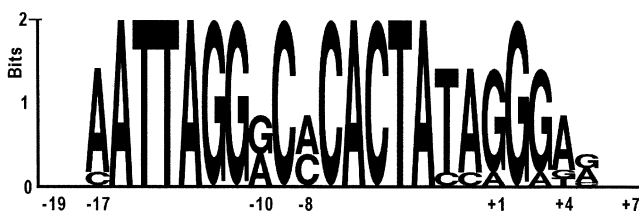


Fig. 1. Sequence logos of the K11 promoter. Logos were generated from the nine promoter sequences using programs at <http://www.bio.cam.ac.uk/seqlogo/logo.cgi>. The logo height indicates the amount of information each position contains.

determine where the insert sequences actually were within the range.

Comparison with the T7 genome map Using a DNA sequence homology search program (BLAST at NCBI), the entire GenBank sequence database was searched for sequences orthologous with the eight sequences of the K11 promoter-bearing inserts. Six K11 sequences matched with T7 or T3 sequences (65-87% identity, and the matched sequences were of 37-73 base pairs). All the K11 promoters have the same polarity with the corresponding T7 and T3 promoters, running from left to right on their linear genomes.

The clones in pKR54, pKT68, pKT56 and pKR2 matched with genes 3.8, 6.5, 13 and OR of the T7 genome, respectively. The T7 promoters are located on 28.0%, 46.4%, 68.3% and 98.2% positions of the T7 genome, respectively (Fig. 2). The clones in pKT54 and pKR24 matched with

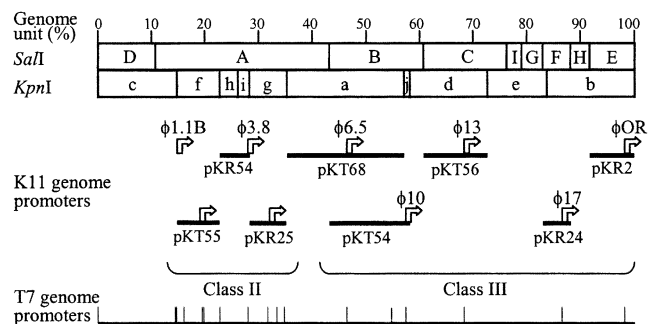


Fig. 2. Genomic mapping of the cloned K11 promoters. The *Sall*- and *KpnI*-restriction maps of the 43-kb K11 genome are shown in scale with genomic units (%). The *Sall*-*KpnI* fragmentation regions that commonly hybridized to the radiolabeled transcripts of each K11 promoter-bearing plasmid are shown by solid lines. The possible genomic locations of the cloned K11 promoters are shown by open arrows. The genomic locations of all 17 T7 promoters are shown as upright lines for the purpose of comparison.

genes 10 and 17 of the T3 genome, respectively. Their corresponding T7 promoters are located on 57.4% and 86.6% positions of the T7 genome, respectively (Fig. 2). The promoter positions assumed in this way are in good agreement with the ranges of their clone sequence locations.

The pKT55 and pKR25 insert sequences did not substantially match with T7 and T3 genomes. These two inserts were rather small. The pKT55 0.25-kb insert hybridized with the *KpnI* fragment f (15-23% position), and

one K11 promoter should be within the range. The pKR25 2.1-kb insert hybridized with the *KpnI* fragment g (28-35% position), and also had only one promoter. (Several T7 promoters are located within each of the corresponding regions of the T7 genome.) The previously identified K11 promoter, named ϕ 1.1B, is also shown in Fig. 2.

Transcription activity of the cloned promoters Four of the eight newly identified promoters were quantitatively assayed *in vitro* for transcription activity. The promoters in pKT54 and pKR24 have -10G and -8A, and those in pKR2 and pKR25 have -10A and -8C. All have the consensus sequence from -17 to +4 but their flanking sequences are different. These four promoters yielded roughly the same activity (within $\pm 20\%$).

In order to eliminate the effects of different flanking sequences, four synthetic promoters were constructed and cloned in pUC118 such that the sequences are identical, except for the residues in the -10 and -8 positions. Two promoters had the natural -10G/-8A and -10A/-8C, and two had variants -10G/-8C and -10A/-8A. Relative strengths of the four promoters were quantitatively assayed in transcription reactions containing two plasmids. One plasmid was digested with *HaeII* and the other with *PvuII*, so that the transcripts produced from the two promoters were different in size, 168 and 95 nucleotides, respectively. All six pair-wise comparisons showed roughly the same promoter strengths (within $\pm 20\%$). Even when transcription activities were measured in high salt conditions (20 mM MgCl₂ or 100 mM NaCl plus 6 mM MgCl₂), no significant difference was observed between the four promoters. Thus, the base pair differences at -10 (G or A) and -8 (A or C) did not appear to affect promoter strength. Also, transcription activity did not appear affected by the flanking sequences.

Discussion

The phage K11 system belongs to the family comprising the T7, T3 and SP6 systems. Although 18 T7 promoters (including one in T3 genome), 14 T3 promoters and 11 SP6 promoters have been identified, only one K11 promoter has been reported (see Lee and Kang (1993) for a compilation). Consensus sequences have been derived from the multiple phage promoters, but not for K11 for which only one promoter sequence was known. The results of the present study now mean nine K11 promoters are known, and a consensus sequence can be drawn from them (Table 1).

The K11 promoter consensus sequence is a stretch of 21 base pairs, from -17 to +4 with respect to the transcription start site +1. It is interesting that no particular base is predominant at -10 and -8, but the two residues appear to be linked, with all nine promoters having only -10G/-8A or -10A/-8C. However, this linkage does not seem to bear any functional significance since consensus sequence promoters containing the variations -10G/-8C and -10A/-8A have

similar activity as the natural -10G/-8A and -10A/-8C.

The promoter base pairs from -12 to -8 of T7, T3 and SP6 are highly conserved within a phage, but differ from one phage to another. These bases have been proposed to constitute a promoter identity region that results in specific interactions with their own RNA polymerases. We have previously shown that alteration of two base pairs in the T7 and SP6 promoters at -9 and -8 (CT \leftrightarrow GA) is necessary and sufficient to switch recognition from one polymerase to the other (Lee and Kang, 1993). Also, Klement *et al.* (1990) demonstrated that discrimination between T7 and T3 promoters depends on the region from -12 to -10, although there was some cross-reaction activity between the two systems.

However, not all residues from -12 to -8 determine the identity of phage promoters. Although the set of mutants studied here is limited, the residues at -10 and -8 of the K11 promoter do not seem to be critical for transcription activity (Fig. 1). Our previous study of saturation mutagenesis on the SP6 promoter suggested that the residue at -10 hardly contributes to transcription activity *in vivo* or *in vitro* (Shin *et al.*, 2000), although it is mostly T among the 11 known SP6 promoters. Also, the SP6 residues at -12 are highly diverse and contribute little to promoter activity.

According to the X-ray crystal structure of the T7 promoter-RNA polymerase complex, the promoter base pairs at -11, -9, -8 and -7 interact with Asn-748, Arg-756, Gln-758 and Arg-746 of RNA polymerase. The K11 RNA polymerase has Lys, Asn, Lys and Arg at the corresponding positions, respectively. If this alignment is correct, the -7C appears to be recognized by Arg in both T7 and K11 RNA polymerases. The T7 -8T interacts with Gln-758 of T7 RNA polymerase. The -8 residue of natural K11 promoters is either C or A, and the K11 RNA polymerase has Lys in the corresponding position of T7 Gln-758. When the Gln-758 was changed to Lys, the variant T7 polymerase recognized only the -8C variant of the T7 promoter (Rong *et al.*, 1998). This result could now suggest a possible interaction between Lys and -8C in the K11 system.

The phage T7 genes are grouped into three classes based on their genome location. Class I genes are located in the leftmost region of the linear genome and are transcribed by the host RNA polymerase. Class II genes, located in the middle, and class III genes, located in the rightmost region, are transcribed by the phage RNA polymerase. The class III promoters show the strongest transcription activity *in vivo* and *in vitro*, and mostly have the conserved sequence. The K11 promoter sequences in plasmids pKT68, pKT54, pKT56, pKR24 and pKR2 are identical to the consensus sequence. These five promoters might belong to class III since they are located in the rightmost two-thirds of the genome, based on the hybridization results and the locations of their orthologous genes in the T7 genome (Fig. 2).

On the other hand, the three promoters in pKT55, pKR54 and pKR25 are located towards the leftmost one-third region.

Their sequences differ from the consensus at 2-3 positions, and hence their promoter strengths may be weak. Thus, these three promoters might belong to class II (Fig. 2). The promoter in pKT55 differs at -1 and +1, positions that are generally important in phage promoters. The +1A has been found in two T7 promoters, although all other phage promoters have +1G. Six T3 promoters have -1C, and one T7 promoter and one SP6 promoter have -1T, while all other phage promoters have -1A. The promoter in pKR54 differs at the upstream (-17) and downstream (+4) ends, and also at -2. This is the only example of a base other than A or T at the -2 position among all 52 known phage promoters. The promoter in pKR25 differed only at the downstream ends, +3 and +4, positions that generally do not contribute much to promoter activity.

Acknowledgments We are grateful to Prof. R. Hausmann (Freiburg University, Germany) for phage K11 stock and its host *Klebsiella* sp. 390 (03:K11). This work was supported by grants from the Korea Science and Engineering Foundation (R01-1999-00111), from the National Research Laboratory Program (M1-0104-00-0179), and from the Brain Korea 21 Project.

References

- Dietz, A., Kössel, H. and Hausmann, R. (1985) On the evolution of the terminal redundancies of *Klebsiella* phage no. 11 and of coliphages T3 and T7. *J. Gen. Virol.* **66**, 181-186.
- Dietz, A., Weisser, H.-J., Kössel, H. and Hausmann, R. (1990) The gene for *Klebsiella* bacteriophage K11 RNA polymerase: Sequence and comparison with the homologous genes of phages T7, T3, and SP6. *Mol. Gen. Genet.* **221**, 283-286.
- Eun, S. H., Kim, D. J. and Kim, Y. S. (2001) *Acinetobacter calcoaceticus* glucose-1-phosphate thymidyltransferase: cloning, sequencing, and expression in *E. coli*. *J. Biochem. Mol. Biol.* **34**, 230-236.
- Han, K. G., Lee, S. S. and Kang, C. (1999) Soluble expression of cloned phage K11 RNA polymerase gene in *Escherichia coli* at a low temperature. *Protein Expr. Purif.* **16**, 103-108.
- Kang, C. and Wu, C.-W. (1987) Studies on SP6 promoter using a new plasmid vector that allows gene insertion at the transcription initiation site. *Nucleic Acids Res.* **15**, 2279-2294.
- Klement, J. F., Moorefield, M. B., Jorgensen, E., Brown, J. E., Risman, S. and MacAllister, W. T. (1990) Discrimination between bacteriophage T3 and T7 promoters by the T3 and T7 RNA polymerases depends primarily upon a three base-pair region located 10 to 12 base-pairs upstream from the start site. *J. Mol. Biol.* **215**, 21-29.
- Lee, S. S. and Kang, C. (1993) Two base pairs at -9 and -8 distinguish between the bacteriophage T7 and SP6 promoters. *J. Biol. Chem.* **268**, 19229-19304.
- Rong, M., He, B., McAllister, W. T. and Durbin, R. K. (1998) Promoter specificity determinants of T7 RNA polymerase. *Proc. Natl. Acad. Sci. USA* **95**, 515-519.
- Roytrakul, S., Eurwilaichitr, L., Suprasongsin, C. and Panyim, S. (2001) A rapid and simple method for construction and expression of a synthetic human growth hormone gene in *Escherichia coli*. *J. Biochem. Mol. Biol.* **34**, 502-508.
- Shin, I., Kim, J., Cantor, C. R. and Kang, C. (2000) Effects of saturation mutagenesis of the phage SP6 promoter on transcription activity, presented by activity logos. *Proc. Natl. Acad. Sci. USA* **37**, 3890-3895.