Identification of new middle promoters upstream genes *nrdC* and *tk.3* of bacteriophage T4

A. Zajančkauskaitė,
G. Kolesinskienė,
L. Piešinienė,
L. Truncaitė,
A. Driukas,
R. Nivinskas

Institute of Biochemistry,
Mokslininkų 12,
LT-2600 Vilnius, Lithuania

The consensus sequence of T4 middle promoters is defined by the MotA box (a/t)(a/t)TGCTT(t/c)A centered at position –30 and the *Escherichia coli* σ^{70} consensus sequence TAnnnT centered at –10. The results of primer extension sequencing of RNA isolated from T4D *motA*$^+$ and T4D *motA*$^-$ infections demonstrate the presence of T4 middle promoters just upstream of the genes *nrdC* and *tk.3*. Both new middle promoters have differences from the consensus MotA box, while their –10 regions match the σ^{70} consensus very well. Thus, our data support the recent idea that there is more flexibility in the sequence requirements for MotA than was previously appreciated.

Key words: bacteriophage T4, middle promoter, MotA box, transcription

INTRODUCTION

Three different classes of genes, early, middle and late, may be distinguished in the bacteriophage T4 genome according to their expression time. Temporal control of T4 gene expression is achieved mainly at the stage of transcription initiation by sequential activation of different classes of promoters.

The T4 middle promoters are characterized by the standard *E. coli* σ^{70} consensus sequence TAnnnT centered at –10, the MotA box consensus sequence (a/t)(a/t)TGCTT(t/c)A centered at the position –30 and the spacing of 11–13 bp between them [1, 2]. The –10 region is highly conserved among all known middle promoters, while the –30 sequence of some middle promoters deviates from the MotA box consensus sequence even at the most highly conserved positions [3, 4]. T4-modified host RNA polymerase (RNAP), transcription activator MotA, and co-activator AsiA are required for the transcription from T4 middle promoters [5]. The T4 early gene product MotA [6] is a DNA-binding protein that binds specifically to the MotA box sequence of middle promoter DNA and activates transcription [7]. AsiA is also a product of the early T4 gene [8], which binds tightly to the σ^{70} subunit of RNAP [9]. The AsiA-σ^{70} complex together with MotA recruit the host RNA polymerase to initiate from middle mode promoters.

The bacteriophage T4 thioredoxin gene *nrdC* [10] and gene *tk.3* [11] lie in two genomic regions which were known to contain only early and late genes but lack nearby middle promoters. Here we demonstrate the presence of T4 new middle promoters $P_M$tk.3 and $P_M$nrdC. Furthermore, both middle promoters contain the sequences which differ from the consensus MotA box.

MATERIALS AND METHODS

Bacterial and bacteriophage strains. *E. coli* strains B40 (*supD*, ser) and B^E (*sup^0*) were kindly provided by Dr L. W. Black. Bacteriophage T4D wild-type was a gift from Dr W. B. Wood. Phages T4D *motA*$^+$ (*33amN134–55amBL292–45amE10*) and T4D *motA*$^-$ (*33amN134–55amBL292–45amE10-tsG1*) were kindly supplied by Dr N. Guild.

RNA isolation. *E. coli* B^E (*sup^0*) cells were grown at 30 °C to a density of 3×10^8 cells/ml in LB medium before the infection with bacteriophage T4 wild-type at a m.o.i. of 10. In the case of mutant phages T4D *motA*$^+$ and T4D *motA*$^-$, *E. coli* B^E (*sup^0*) cells were grown at 42 °C. RNA was extracted essentially as described [12].

Primer extension analysis. Two synthetic oligonucleotides were used to prime reverse transcriptase: Pr. tk.3, a 24-mer, 5'-GTTCCCTTTACCTATGCC TGCACC, complementary to nucleotides 364–387 of the gene *tk.4* coding sequence; Pr. nrdC, a 27-mer, 5'-GACGTTTTGCATTATCGCAATACACAC, complementary to nucleotides 41–67 of the gene *nrdC* coding sequence. The oligonucleotides were 5'-end labeled by T4 polynucleotide kinase with [γ-^{32}P]ATP (Amersham Biosciences) and separated from the la-

beled nucleotides by precipitation with ethanol in the presence of 2 M ammonium acetate. Primer extension and RNA sequencing were carried out essentially as described by Uzan et al. [12].

RESULTS AND DISCUSSION

Figure 1A and B show the positions of bacteriophage T4 genes *tk.3* and *nrdC* on the physical map of T4. Almost all genes located in these genomic regions are expressed during the early period of infection. Only four genes, *rI.1*, *rI*, *rI.–1* and *49*, are under control of the late promoters. Surprisingly, no middle promoters have been shown to be located in these gene clusters. However, detailed analysis of the nucleotide sequences in these genomic regions revealed two potential middle promoters located just upstream of genes *tk.3* and *nrdC*.

In order to test the presence of these middle promoters, we performed primer extension sequencing of RNA extracted from T4D *motA*$^+$ and T4D *motA*$^-$ infected cells. The hard stops corresponding to the 5′ ends of transcripts directed from the putative middle promoters were observed when the *motA*$^+$ RNA templates were used in the sequencing reactions. In the case of the middle promoter upstream of gene *tk.3*, the 5′ ends of the transcripts indicated by reverse transcriptase stops (more intensive bands at the same position in all four sequencing lanes) start at nucleotides G and A (Fig. 2, A) positioned 5–6 bp downstream the promoter –10 region (Fig. 1, C). The 5′ ends of the transcripts directed from the promoter upstream of the gene *nrdC* start at nucleotides A and G (Fig. 3) positioned 5–6 bp downstream the –10 sequence TATTAT (Fig 1, C). In contrast, primer extension sequencing of transcripts from *motA*$^-$ infections did not reveal any 5′ ends (Fig. 2, A and Fig. 3). Thus, these results clearly indicate the presence of two middle promoters, $P_M$tk.3 and $P_M$nrdC.

To demonstrate the kinetics of mRNA synthesis directed from the middle promoter $P_M$tk.3, we performed primer extension analysis of RNA extracted from T4D-infected cells. Very week bands indicating the 5′ ends of transcripts directed from the promoter P_M tk.3 can be seen only two minutes after infection, and the intensities of the bands increase until six minutes after infection (Fig. 2, B). Thus, primer extension analysis shows that the kinetics of

Fig. 1. Genes and promoters in two genomic regions of bacteriophage T4. The schematic outline represents the genomic region between **(A)** 66 and 59 kb, and **(B)** 55 and 47 kb on the physical map of bacteriophage T4. Shown are the positions of genes, as well as the positions of early (P_E), late (P_L), and two middle promoters (P_M) revealed in this study. **(C)** Sequences of T4 middle promoters were identified by primer extension analysis. The consensus sequence of T4 middle promoters is given on the first line, matches between the promoters and this sequence are shown in bold face type. Initiating nucleotides for the MotA-dependent transcripts are indicated by vertical arrows

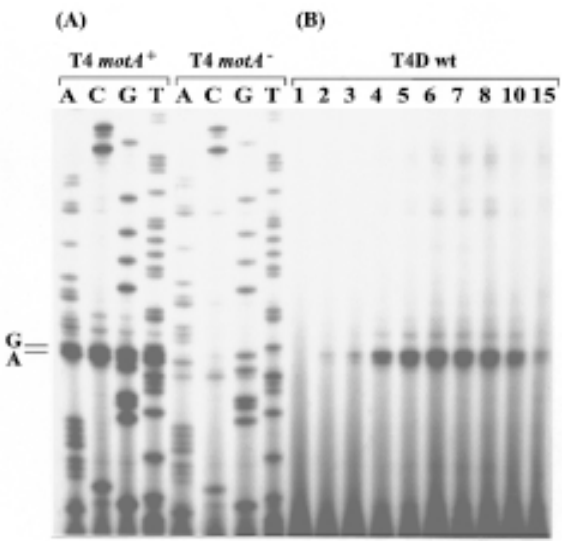

Fig. 2. Primer extension analysis of gene *tk.3* mRNA using RNA isolated from *E. coli* B^E cells infected with T4D wild-type or its mutants *motA*$^+$ (*33*$^-$ *55*$^-$ *45*$^-$) and *motA*$^-$ (*33*$^-$ *55*$^-$ *45*$^-$ *tsG1*). **(A)** Primer extension sequencing of RNA isolated six minutes post infection from cells that were infected at 42 °C with T4D *motA*$^+$ (*33*$^-$ *55*$^-$ *45*$^-$) and T4D *motA*$^-$ (*33*$^-$ *55*$^-$ *45*$^-$ *tsG1*). The sequencing lanes are labeled with the dideoxynucleotides used in the sequencing reactions. The initiating nucleotides, G and A, are noted. **(B)** Primer extension reactions on RNA isolated 1 to 15 minutes post infection from cells that were infected with T4 wild-type at 30°C. The time (minutes) of post infection that each RNA was isolated is noted at the top of the figure

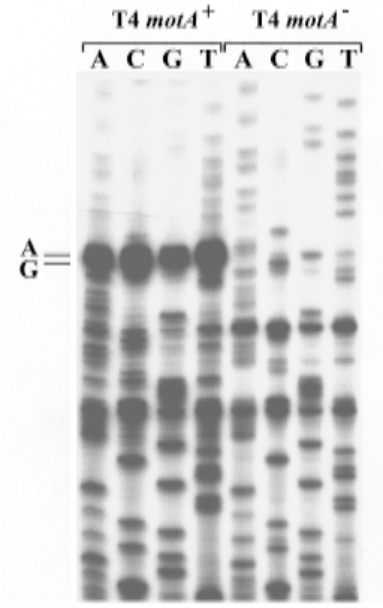

Fig. 3. Primer extension sequencing of gene *nrdC* mRNA using RNA isolated at six minutes post infection from *E. coli* B^E cells infected at 42 °C with T4D *motA*$^+$ (*33*$^-$ *55*$^-$ *45*$^-$) and T4D *motA*$^-$ (*33*$^-$ *55*$^-$ *45*$^-$ *tsG1*). The sequencing lanes are labeled with the dideoxynucleotides used in the sequencing reactions. The initiating nucleotides, A and G, are noted

accumulation of the 5′ ends of transcripts resulting from initiation at the promoter $P_M$tk.3 is in agreement with the kinetics of mRNA synthesis directed from T4 middle promoters.

Earlier transcription studies revealed 24 middle promoters on the T4 genome. Twenty three of them were listed in the review of Stitt & Hinton [2] and one, for gene *31*, was given by Nivinskas et al. [13]. Nineteen of these promoters carried the sequence motif TGCTT in their MotA boxes. On the basis of the sequences of these middle promoters it was suggested that the MotA binding required the highly conserved –30 region. However, recently Marshall et al. [3] have revealed seven new T4 middle promoters, of which only two contain the sequence motif TGCTT, while the other five contain differences even in the highly conserved center motif. It was proposed that an excellent match to the σ^{70}–10 consensus sequence, rather than an excellent match to the MotA box consensus sequence, is an invariant feature of MotA-dependent promoters. Furthermore, recently we have identified new T4 middle promoters, P_M30 and P_M30.2 [4], which significantly differ from the –30 consensus sequence. Both new middle promoters $P_M$tk.3 and $P_M$nrdC revealed in this study also differ from the consensus MotA box, while their –10 region match the σ^{70} consensus sequence very well. Promoter $P_M$tk.3 contains the sequence CGCTT and promoter $P_M$nrdC contains the sequence TGCTA in the highly conserved positions of the –30 region. So, now we know 35 T4 middle promoters, of which 14 deviate from the consensus sequence deduced previously. The main deviations reside in the MotA box sequence, while the –10 sequences match the host-like consensus very well. Thus, our data corroborate the idea of Marshall et al. [3] that MotA can tolerate changes at the most highly conserved positions of the MotA box.

ACKNOWLEDGMENTS

This work was supported in part by the Lithuanian State Sciences and Studies Foundation.

References

1. Guild N, Gayle M, Sweeney R, Hollingsworth T, Modeer T, Gold L. J Mol Biol 1988; 199: 241–58.
2. Stitt B, Hinton D. In: Molecular Biology of Bacteriophage T4. Karam JD. (Eds.) American Society for Microbiology, Washington, DC, 1994: 142–60.
3. Marshall P, Sharma M, Hinton DM. J Mol Biol 1999; 285: 931–44.
4. Truncaite L, Zajančkauskaite A, Nivinskas R. J Mol Biol 2002; 317 (2): 179–90.

5. Ouhammouch M, Adelman K, Harvey SR, Orsini G, Brody EN. Proc Natl Acad Sci USA 1995; 92: 1451–5.
6. Mattson T, Van Houwe G, Epstein RH. J Mol Biol 1978; 126: 551–70.
7. Schmidt RP, Kreuzer KN. J Biol Chem 1992; 267: 11399–407.
8. Orsini G, Ouhammouch M, Le Caer JP, Brody EN. J Bacteriol 1993; 175: 85–93.
9. Stevens A. Proc Natl Acad Sci USA 1972; 69: 603–7.
10. Tomaschewski J, Rüger W. Nucl Acids Res 1987; 15: 3632–3.
11. Valerie K, Stevens J, Lynch M, Henderson EE, de Riel JK. Nucl Acids Res 1986; 14: 8637–54.
12. Uzan M, Favre R, Brody E. Proc. Natl Acad Sci USA 1988; 85: 8895–9.
13. Nivinskas RH, Raudonikiene AA, Guild N. Mol Biol 1989; 23: 739–49 (in Russian).

A. Zajančkauskaitė, G. Kolesinskienė, L. Piešinienė, L. Truncaitė, A. Driukas, R. Nivinskas

T4 BAKTERIOFAGO NAUJŲ VIDURINIŲJŲ PROMOTORIŲ PRIEŠ *nrdC* IR *tk.3* GENUS NUSTATYMAS

S a n t r a u k a

T4 bakteriofago viduriniesiems promotoriams yra būdingos dvi konservatyvios sekos: –30 srityje esanti seka (a/t)(a/t)TGCTT(t/c)A, kurią atpažįsta MotA baltymas, bei –10 srityje išsidėsčiusi TAnnnT seka, būdinga *E. coli* σ^{70} promotoriams. Du vidurinieji promotoriai prieš *nrdC* ir *tk.3* genus buvo identifikuoti pradmens prailginimo metodu nustatant RNR, išskirtų iš T4D *motA*$^+$ bei T4D *motA*$^-$ fagais infekuotų ląstelių, nukleotidų sekas. Abiejų naujai išaiškintų promotorių MotA sekos skiriasi nuo tipiškos sekos, tuo tarpu –10 sritys išlieka konservatyvios. Taigi mūsų gauti rezultatai palaiko besiformuojančią nuomonę, kad galima didesnė, nei anksčiau buvo manyta, įvairovė sekos, kurią atpažįsta MotA baltymas.