BIOLOGIJA. 2006. Nr. 2. P. 5–8

Lysis inhibition gene *rI* of bacteriophage RB49: the primary structure and expression regulation

A. Driukas*, V. Klausa,

G. Petkūnienė, R. Nivinskas

Department of Gene Engineering, Institute of Biochemistry, Mokslininkų 12, LT-08662 Vilnius, Lithuania

The lysis inhibition phenomenon was examined predominantly in T4. However, recent studies of viral evolution have made it clear that T4 family is an extremely widespread group of phages and inspired an extensive research of the genomic divergence between T4 and its various relatives. In the present study, we identified the pseudo T-even phage RB49 gene homologous for the T4 lysis inhibition gene *rI*. Analysis of gene expression revealed that transcription of the RB49 gene *rI* is very similar to that of T4 *rI*. However, in the translation initiation region of the RB49 gene *rI* we identified a stable RNA secondary structure, which potentially could inhibit the synthesis of gprI from the early transcripts. This result indicates the different control of RB49 gene *rI* expression, which interferes with the gprI function in lysis inhibition.

Key words: bacteriophage RB49, gene *rI*, lysis inhibition, RNA secondary structure

INTRODUCTION

The usual process of lysis of T4-infected cells involves an endolysin encoded by the gene *e* [1] and holin encoded by the gene *t* [2]. The gpt holin permeabilizes the cytoplasmic membrane permitting the passage of the endolysin gpe to the periplasm where the muralytic enzyme degrades the peptidoglycan layer [3]. In the normal infection cycle of T4 phage, lysis of the bacterial cell and release of phage particles take place at 20–30 min postinfection at 37 °C. However, if secondary phage particles adsorb to the already T4-infected cell, the latent period can be prolonged for as long as several hours with the consequent increase in the burst size. This phenomenon is known as lysis inhibition [4]. The molecular basis for the T4 lysis inhibition is a specific inhibition of the holin T by antiholin encoded by the gene *rI* [5, 6].

The role played by the gene *rI* in the regulation of the length of the infection cycle requires production of gprI in the cell before holin as a late gene product is made [6, 7]. Previous reports have revealed that T4 gene *rI* can be transcribed from the two upstream promoters belonging to different temporal classes: the late promoter upstream the gene *rI.1* and the early promoter upstream the gene *regB* [7]. Recently, the middle promoter $P_M$tk.3 upstream the gene *tk.3* was identified [8], which potentially directs the transcription of the gene *rI*. The transcription of the gene *rI* from the early, middle and late promoters is reasonable, because lysis inhibition can be induced as early as 3 min after infection and even after the lysis of the infected culture starts [7]. The notion that genes involved in lysis inhibition should be expressed in an appropriate mode has been supported by the fact that the transcription regulation of the potential lysis inhibition gene *rIII* is virtually identical to that of *rI* [7]. Recently, we have shown that the gene *rIII* encoded protein is produced early as well as late in infection [9].

In this study, we determined the primary structure of the lysis inhibition gene *rI* of the distant T4-type phage RB49. The genome of the pseudo T-even RB49 had substantially diverged from T4. Seventy percent of the RB49 genome encodes protein homologues of T4 gene products with the levels of amino acid identity ranging between 50% and 70% [10]. Many nonessential T4 genes in the genome of RB49 are replaced by novel open reading frames (ORF) encoding proteins without homologues in the NCBI protein database. The DNA of RB49 does not have nucleotide modifications [11], and this phage uses only two classes of temporally regulated promoters (early and late) rather than the three (early, middle and late) employed by T4 [10]. Moreover, the RB49 early promoters have the same consensus sequence as the *E. coli* σ^{70}-dependent promoters.

The results obtained in this study revealed that the expression pattern of the lysis inhibition gene *rI* of RB49

* Corresponding author. E-mail: audrius@bchi.lt

differs from that of T4 because of the stable RNA secondary structure in the translation-initiation region of the RB49 gene *rI*.

MATERIALS AND METHODS

Bacterial and bacteriophage strains. *Escherichia coli* strain B^E (*sup*^{0}) was a gift from Dr. Lindsay Black. Bacteriophage RB49 was kindly supplied by Dr. William Wood.

Sequencing of the RB49 gene *rI*. DNA sequencing was performed using a CycleReader™ DNA sequencing kit (Fermentas AB). The oligonucleotide primers were 5'-end-labeled by T4 polynucleotide kinase (Fermentas AB) with [γ-^{32}P]ATP (Amersham Biosciences). The DNA template for the sequencing reactions was in the form of a purified genomic phage DNA. Based on the sequence of the gene *tk* of the pseudo T-even phage RB49 [10], we designed the primer 5'-GTCTCCGCACTGACTTTCAGGG. This primer and other RB49-specific primers, 5'-GATGGTACTACCAACTAAGAG and 5'-GCCCATATTATGAAGATGTATGCTCG, were used to determine the sequence of the RB49 gene *rI*.

RNA methods. *E. coli* B^E (sup^0) cells were grown at 30 °C to a density of 2×10^8 cells/ml in LB medium before the infection with bacteriophage RB49 at a MOI of 10. RNA isolation, primer labeling and RNA sequencing were performed as described by McPheeters [12]. An assay was carried out with the AMV reverse transcriptase (Promega) using the primer 5'-GTTGGATCATATTGGTCAGCAC, complementary to 62-83 nt of the RB49 gene *rI*. The samples were analysed on a denaturing polyacrylamide gel (6% acrylamide, 8 M urea, TBE).

RESULTS AND DISCUSSION

In order to identify the RB49 homolog of the T4 lysis inhibition gene *rI*, we designed a primer based on the sequence of the gene *tk* of RB49 for sequencing the *tk* downstream region (see Materials and Methods). Immediately downstream the gene *tk* we defined an ORF preceded by a short intergenic region (Fig. 1A). The determined ORF has a potential to encode the 100-amino acid residue protein with a predicted molecular weight of 11.4 kDa and a theoretical pI of 6.5. An alignment between the primary structures deduced for the protein encoded by this ORF and the known 97-amino acid sequence for T4 gprI showed that a gene located downstream *tk* in the genome of RB49 encodes a highly diverged gprI homolog. RB49 gprI exhibits a 20% identity to the gprI of T4 at the amino acid sequence level (Fig. 1B).

Analysis of the noncoding region between RB49 genes *tk* and *rI* allowed to identify the sequence for the late promoter upstream the gene *rI* (Fig. 1A). In addition, the RB49 gene *rI* can potentially be transcribed into the early polycistronic mRNAs initiated at several *E. coli* like promoters located upstream (Fig. 2). To investigate the transcription pattern of the RB49 gene *rI in vivo*, we performed a primer extension analysis of RNA isolated at 5 and 15 min post infection from RB49-infected cells. As shown in Fig. 3A, the 5' end of the late transcript was detected 5 bp downstream from the late promoter sequence in case of RNA isolated at 15 min post infection. On the other hand, the RB49 gene *rI* mRNA was also detected in the samples isolated at 5 min post infection, indicating that this gene is also transcribed from the upstream early promoters. The obtained results clearly demonstrated that the RB49 gene *rI* is transcribed early as well as late in infection.

The phage T4 uses RNA secondary structures to prevent or reduce early translation of at least five late genes, the chromosomal locations of which result in their cotranscription with early or middle genes [13]. In each case, the late promoter is in the potential stem-loop region or in a close proximity to it, so the late transcripts have no secondary structures. T4 *rI* is not the subject of such regulation, because no RNA secondary structures were found in the translation initiation region of this gene [7]. Furthermore, T4 *rI* is the second gene in the late transcription unit (Fig. 2). Thus, the transcripts for the gene *rI* of T4 appear likely to be translated both early and late in infection.

The RB49 genomic region containing the gene *rI* considerably diverged as compared with the analogous genomic region of T4 (Fig. 2). The RB49 gene *rI* is the first gene in a two-ORF late transcription unit due to the absence of *rI.1* in the genome of RB49. Moreover, we examined the untranslated region upstream the gene *rI* of RB49 and revealed a potential stem-loop structure that would sequester the proximal end of the ribosome-binding site of *rI* in the early polycistronic mRNAs (Fig. 3B). This secondary structure cannot be formed in the late transcripts, since the late promoter for the gene *rI* is in the middle of the stem-loop region. Primer extension analysis confirmed the existence of a stable secondary structure in the early mRNA of the RB49 gene *rI*. As shown in Fig. 3A, two 5' ends of the transcript were detected immediately downstream from the predicted stem-loop in the reverse transcriptase assay of early mRNA. The positions of these 5' ends indicate that they were generated by reverse transcriptase stops at the stable RNA stem-loop in the translation initiation region of the gene *rI* (Fig. 3B). Therefore, these results suggest that the synthesis of RB49 gprI product is inhibited at the early phase of infection.

In conclusion, detection of the gene *rI* homolog in the genome of the phylogenetically distant phage RB49 supports the notion that lysis inhibition contributes significantly to the widespread occurrence in nature of T4-type phages and to their competitive advantage [7]. On the other hand, although transcription of the RB49 gene *rI* is almost the same as in case of T4, the RB49 gprI seems to be synthesized in a pattern that is distinct from that of T4. Considering the observed concordance

A

```
                                                          P_L
   1 atggtactaccaactaagagttaatcataaataaaggccccagattataaataataactga
     M  V  L  P  T  K  S  *  (end of tk)
  61 atattttaaatctgggggaatgacttatgtcatttattcgtaaaactctaacagcggcaa
                                 M  S  F  I  R  K  T  L  T  A  A  I
 121 tatgctttactgcggtcggaagcgtcagtgctgaccaatatgatccaactggtgacttta
       C  F  T  A  V  G  S  V  S  A  D  Q  Y  D  P  T  G  D  F  T
 181 ctgattatttcgacggcgcaatgatagtatatgtaaacaaaatatctccttccattgaaa
       D  Y  F  D  G  A  M  I  V  Y  V  N  K  I  S  P  S  I  E  T
 241 ccagccagaaatttcatgcccatattatgaagatgtatgctcgtagtaaatgtatggatt
       S  Q  K  F  H  A  H  I  M  K  M  Y  A  R  S  K  C  M  D  S
 301 ctaagcagtgttttattatgggggaatatgccgcaaaccaatttgttgtgttgcataata
       K  Q  C  F  I  M  G  E  Y  A  A  N  Q  F  V  V  L  H  N  M
 361 tggagaatcgagaacatgaaattcagtgattttcgaaagacgtccgaagatgggaaagct
       E  N  R  E  H  E  I  Q  *
```

B

```
T4    MALKATALFAMLGLSFVLSPSI-EANVDPHFDKFMESGIRHVYMLFENKSV 50
RB49  MSFIRKTLTAAICFTAVGSVSADQYDPTGDFTDYFD-GAMIVYVNKISPSI 50
      *::   :* * : :  * * *  : :   .* .::: *   **      *:

T4    ESSEQFYSFMRTTYKNDPCSSDFECIERGAEMAQSYARIMNIKLETE---  97
RB49  ETSQKFHAHIMKMYARSKCMDSKQCFIMGEYAANQFVVLHNMENREHEIQ  100
      *.*::*: .: . * .. * .  :*   *   *.::. : *:: .
```

Fig. 1. (A) Sequence of the RB49 gene *rI*. The Shine–Dalgarno sequence and initiation codon of the gene *rI* are in boldface. The late promoter sequence is boxed. Horizontal arrows indicate sequences predicted to form a secondary mRNA structure. The termination codons of gene *tk* and *rI* are indicated by an asterisk. This sequence has been submitted to the EMBL/GenBank with accession number AJ517766. (B) Alignment of predicted amino acid sequence of RB49 gprI with the sequence of T4 gprI (GenBank accession no. AAB26962). The sequences were aligned using ClustalW program. An asterisk (*****) means that the residues in the column are identical; a double dot (**:**) indicates the conserved substitutions; a dot (**.**) indicates the semi-conserved substitutions. A dash indicates a space which was inserted in the sequence to preserve the alignment

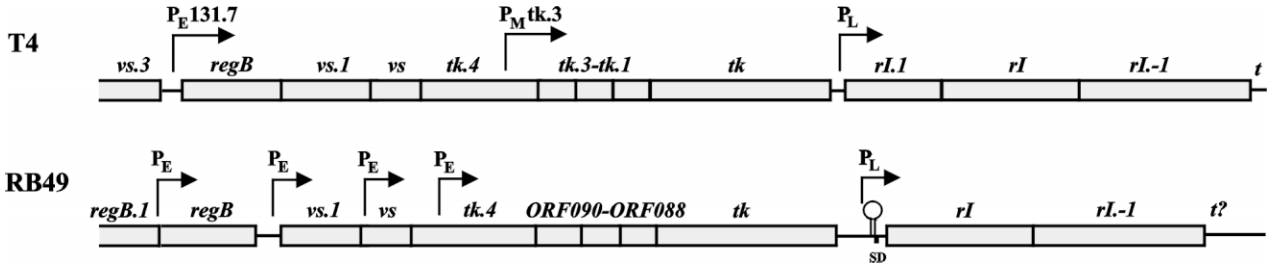

Fig. 2. The gene *regB–rI* region in the genomes of phages T4 and RB49. The name of the gene is indicated above the corresponding box. The arrows represent the promoters and the letter above the arrow indicates its temporal class. The promoters P_E131.7 [14] and $P_M$tk.3 [8] of T4, $P_E$regB, $P_E$vs.1 [15] and P_LrI (this study) of RB49 were shown to be active. The other promoters marked in the scheme were deduced from sequence analysis, including P_LrI.1 [16] of T4. Terminators and the RNA secondary structure in the case of RB49 are indicated as well. The scheme was designed according to the complete genomic sequences of T4 (AF158101) and RB49 (AY343333)

between the mode of expression and the function of the T4 gene *rI* [7], the shift of RB49 gene *rI* into the class of late genes that are transcribed into the early and late transcripts, though translated from the late ones because of inhibitory RNA stem-loop structures, may signify substantial alterations in lysis inhibition function. Work is in progress to test this assumption. We hope these experiments will shed more light on the lysis inhibition phenomenon in the phylogenetically distant T4 relatives such as pseudo T-even RB49.

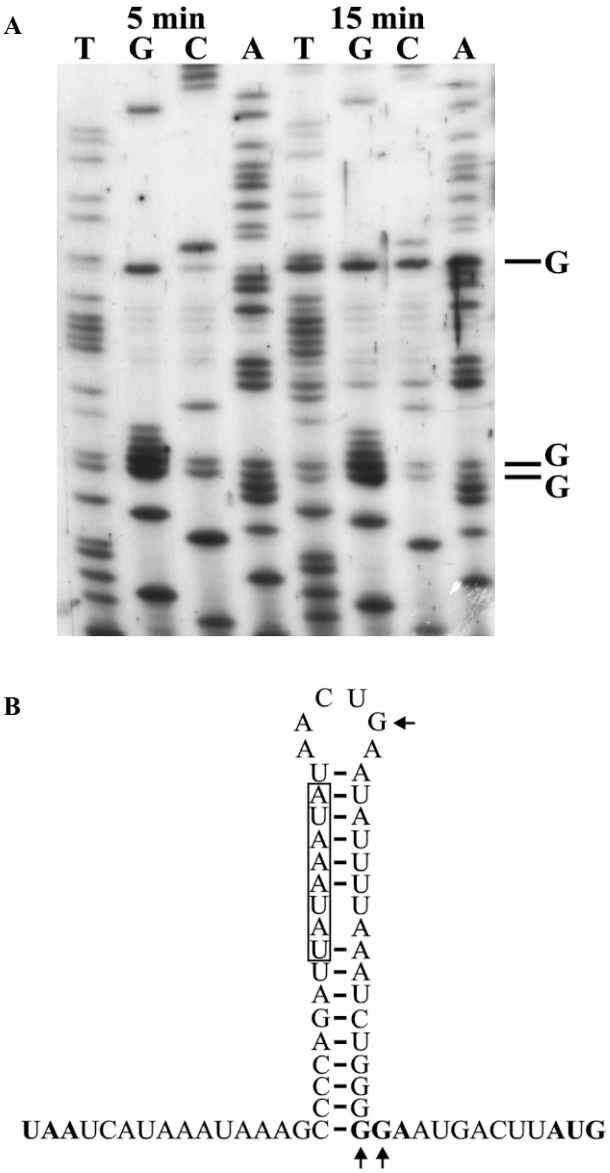

Fig. 3. (A) Primer extension sequencing of the transcripts for RB49 gene *rI*. Primer extensions were done on RNA isolated at 5 and 15 minutes post infection from *E. coli* B^E cells infected at 37 °C with RB49 phage. The sequencing lanes are labeled with the dideoxynucleotides used in the sequencing reactions. The initiating nucleotide G for the late transcript and the reverse transcriptase termination nucleotides downstream the potential stem-loop are noted. (B) Predicted RNA secondary structure in the intergenic region between RB49 genes *tk* and *rI*. Shine–Dalgarno sequence and the initiation codon of gene *rI* are in bold lettering. The late promoter sequence is boxed. Arrows indicate nucleotides at which RT stops were observed

ACKNOWLEDGEMENTS

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

Received 8 March 2006
Accepted 16 May 2006

References

1. Streisinger G, Mukai F, Dreyer WJ, Miller B, Horiuchi S. Cold Spring Harbor Symp Quant Biol 1961; 26: 25–30.
2. Lu M, Henning U. Mol Gen Genet 1992; 235: 253–58.
3. Young R. Microbiol Rev 1992; 56: 430–81.
4. Hershey AD. Genetics 1946; 31: 620–40.
5. Ramanculov E, Young R. Mol Microbiol 2001; 41: 575–83.
6. Tran TA, Struck DK, Young R. J Bacteriol 2005; 187: 6631–40.
7. Paddison P, Abedon ST, Dressman HK, Gailbreath K, Tracy J, Mosser E, Neitzel J, Guttman B, Kutter E. Genetics 1998; 148: 1539–50.
8. Truncaite L, Piešiniene L, Kolesinskiene G, Zajančkauskaite A, Driukas A, Klausa V, Nivinskas R. J Mol Biol 2003; 327: 335–46.
9. Driukas A, Nivinskas R. Biologija 2003; 3: 13–15.
10. Desplats C, Dez C, Tétart F, Eleaume H, Krisch HM. J Bacteriol 2002; 184: 2789-04.
11. Monod C, Repoila F, Kutateladze M, Tétart F, Krisch HM. J Mol Biol 1997; 267: 237–49.
12. McPheters DS. In Molecular Biology of Bacteriophage T4 (Karam JD, ed.) 1994; 469–72
13. Miller ES, Karam JD, Spicer E. In Molecular Biology of Bacteriophage T4 (Karam JD, ed.) 1994; 127–31.
14. Ruckman J, Parma D, Tuerk C, Hall DH, Gold L. New Biol 1989; 1: 54–65.
15. Piešiniene L, Truncaite L, Zajančkauskaite A, Nivinskas R. Nucleic Acids Res 2004; 32: 5582–95.
16. K.P. Williams, G.A. Kassavetis, D.R. Herendeen, Geiduschek EP. In Molecular Biology of Bacteriophage T4 (Karam JD, ed.) 1994; 161–75.

A. Driukas, V. Klausa, G. Petkūnienė, R. Nivinskas

BAKTERIOFAGO RB49 LIZĖS INHIBICIJOS *RI* GENAS: PIRMINĖ STRUKTŪRA IR RAIŠKOS REGULIACIJA

S a n t r a u k a

Lizės inhibicijos fenomenas ilgą laiką buvo tiriamas tik T4 arba labai jam gimininguose T lyginiuose bakteriofaguose. Tačiau pastarųjų metų fagų evoliucijos tyrimai rodo, kad gamtoje egzistuoja daugybė filogenetiškai susijusių T4 tipo bakteriofagų, ir tai paskatino plačius šios fagų grupės genetinės įvairovės tyrinėjimus. Šiame darbe buvo tiriamas pseudo T lyginis fagas RB49. Mes nustatėme, kad šio fago genome yra genas, homologiškas fago T4 lizės inhibicijos *rI* genui. Genų raiškos analizė rodo, kad, kaip ir fago T4 atveju, RB49 *rI* genas yra transkribuojamas tiek ankstyvuoju, tiek vėlyvuoju fago infecijos periodu. Tačiau, skirtingai nuo T4, RB49 *rI* geno transliacijos iniciacijos srityje mes aptikome patvarią iRNR antrinę struktūrą, kuri potencialiai gali slopinti ankstyvųjų šio geno transkriptų transliaciją. Toks RB49 *rI* geno raiškos reguliacijos mechanizmas yra nesuderinamas su lizės inhibicijos funkcija.