Nucleotide sequences of the overlapping genes *30.3* and *30.3'* of T4-related bacteriophages

G. Kolesinskienė,
A. Zajančkauskaitė,
A. Arlauskas,
V. Klausa,
R. Nivinskas

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

Our previous studies showed the existence of a pair of bacteriophage T4 overlapping genes, *30.3* and *30.3'*, of which the smaller, *30.3'*, is completely embedded within the other by one position downstream. Now we have determined the primary structures of the overlapping genes *30.3* and *30.3'* for 15 T4-related phages. The nucleotide sequences of genes *30.3* and *30.3'* in most tested phages, as well as the predicted amino acid sequences of the gene products, showed a more than 85% homology to those of T4. Only in case of gp30.3' of phage RB69 the conservation level of the amino acid sequence showed only a 64.7% similarity to T4. It is interesting to note that the gene *30.3'* of phage Ox2 contains ACG as a putative initiation codon. The sequence of middle promoter P_M30.2 located in the coding region of gene *30.3* is conserved absolutely in all phage genomes studied.

Key words: bacteriophage T4, T4-related phages, genes *30.3* and *30.3'*, middle promoter

INTRODUCTION

T4, the archetype of the T-even phages, has been the subject of intensive study and is one of the best-characterized phages. Phage T4 has numerous relatives, which are closely related genetically but can infect different bacterial species [1]. More than 140 phages with a morphology resembling that of T4 have been revealed in the catalogue of T4-type bacteriophages [2].

We have previously showed that the genomic region between T4 early promoter P_E128.2 and DNA ligase gene *30* contains seven genes, *30.7–30.1* [3]. Moreover, in this region we have established the existence of one more gene, *30.3'*, which is completely embedded within gene *30.3* by one position downstream [4, 5]. Recently, two new T4 middle promoters have been identified: promoter P_M30 which is located just upstream of gene *30* and promoter P_M30.2 located in the coding region of gene *30.3* [6].

Most genetic and sequence work has been done with T2, T4 and T6, but interest has been growing rapidly in the broader T-even phage family [7]. Comparative analysis of T4-related bacteriophages is being used to gain the evolutionary origins and interrelationships of these phages genes, and the functions of their proteins [8].

The aim of our current work was to show the existence of the pair of overlapping genes *30.3* and *30.3'* in various T4-related phages and to evaluate the conservation level of genes and their products.

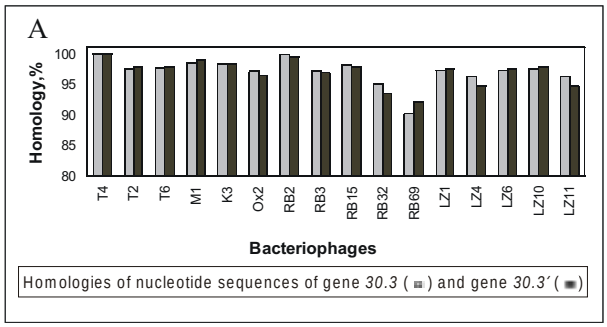
MATERIALS AND METHODS

Phages and bacterial strains. T4-related phages T2, T6, M1, K3, Ox2 were kindly provided by Dr. U. Henning (Tübingen, Germany), while phages RB2, RB3, RB15, RB32, RB69, LZ1, LZ4, LZ6, LZ10 and LZ11 were from Dr. K. Carlson (Uppsala, Sweden). All phages were grown in *Escherichia coli* B^E (sup^0) (provided by Dr. L W. Black), except for RB69 grown in *E. coli* CR63 (supD) (supplied by Dr. K. N. Kreuzer).

Amplification and sequencing of genes *30.3* and *30.3'* of T4-related phages. Three synthetic oligonucleotides based on the T4 genome were used to amplify [9] DNA fragments with genes *30.3* and *30.3'* of T4-related phages: Pr. 1, a 20-mer, 5'-GGACCCTAACTGGCCAGTTG, corresponds with 123-142 nt of the gene *30.4* coding sequence; Pr. 2, a 21-mer, 5'-CAAGATACTCCCCGGCGTGG, corresponds with 254–273 nt of the gene *30.3* coding sequence; Pr. 3, a 20-mer, 5'-AGGAAGGCCTGATTGCCAGC, complementary to nucleotides 44–63 of the gene *30.2* coding sequence. The oligonucleotides were 5'-labeled by T4 polynucleotide kinase with [γ-^{33}P]ATP (Amersham Bioscience). Sequencing was performed with a CycleReader DNA sequencing kit (Fermentas AB).

RESULTS AND DISCUSSION

The nucleotide sequences of the pair of overlapping genes, *30.3* and *30.3*', located in the genomic region between P_E128.2 and ligase gene *30* (Fig. 1) have been determined in 15 T4-related bacteriophages. Nucleotide sequences of gene *30.3*, as well as gene *30.3*', were more than 90% identical to T4 in all phages tested (Fig. 2A). The predicted amino acid sequences of gp30.3 were highly conserved (>90%) in all cases (Fig. 2B). In case of gene *30.3* the occasional single-base differences were observed mostly in the third base of the codon; that is why it did not cause changes of amino acids. It should be mentioned that 15-aa residues in gp30.3 of phage RB69 are absent (Fig. 3A), and homologies at nucleotide and amino acid levels of phage RB69 were showed to the corresponding sequence segments of T4 (Fig. 2). In case of gene *30.3*', which is completely embedded within gene *30.3* by one position downstream, most base differences were in the second base of the codon and therefore, as we can see from the results (Fig. 2B), it caused a change of amino acids. The predicted amino acid sequences of gp30.3' in most phages tested showed homology from 85.3% to 98.6% to those of T4, while in case of phage RB69 a homology of only 64.7% has been shown. The low level of conservation of gp30.3' of phage RB69 coincides with the observations that phage RB69 seems to occupy an "intermediate" position between T-even and pseudo T-even phages [7]. Neither of the 7-aa residues of phage T2 and the 41-aa residues of RB69 are available in case of gp30.3' (Fig. 3B) and the nucleotide and predicted amino acid sequences of those phages were shown to the corresponding sequence segments of T4 (Fig. 2).



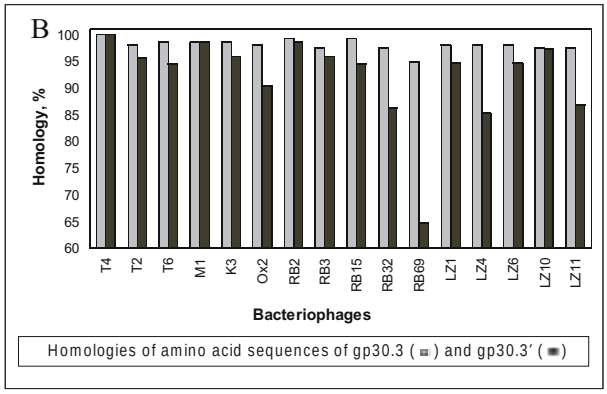


Fig. 2. Homologies of genes *30.3* and *30.3*', as well as gp30.3 and gp30.3', between T4 and 15 T4-related phages at the nucleotide (A) and amino acid level (B)

			–30	11-13 bp	–10	
Middle Promoter Consensus Sequence:			**a a t t TGCTT c t A**		**TA**nnn**T**	
						↓↓
T4	**P_M30.2**	GCACAGCATTGCT	**AGTGGTCTA**	ACAAAAGATGA	**TA**CAA**T**	ACTAAC**CG**AA
T2, T6, M1, K3, Ox2, RB2, RB15, RB32, LZ1, LZ4, LZ6, LZ10, LZ11	**P_M30.2**	GCACAGCATGGCT	**AGTGGTCTA**	ACAAAAGATGA	**TA**CAA**T**	ACTAACCGAA
RB3	**P_M30.2**	GCACAGCATGGCT	**AGTGGTCTA**	ACAAAAGATGA	**TA**CAA**T**	ACTAACTGAA
RB69	**P_M30.2**	GCACAGCATGGCT	**AGTGGTCTA**	ACAAAAGATGA	**TA**CAA**T**	ACTAACTTGA

Fig. 1. Genes and promoters in the genomic region between early promoter P_E128.2 and the DNA ligase gene *30* of phage T4. A schematic outline of T4 DNA fragment carrying overlapping genes *30.3* and *30.3'* is presented at the top of the figure. Shown are the positions of genes, as well as the positions of the early and middle promoters. The consensus sequence of T4 middle promoters [11] is given below. The sequences of putative P_M30.2 in 15 T4-related phages are presented at the bottom of the figure. The matches between all given promoter sequences and the consensus sequence of T4 middle promoters are shown in bold face type. Initiation nucleotides for the transcripts directed from T4 P_M30.2 are indicated by vertical arrows. The nucleotides that differ from T4 are underlined

It is well established that the initiation codon, the Shine–Dalgarno sequence, the spacing between those two elements and the secondary structure of the mRNA are important determinants of translation initiation [10]. The translation initiation regions of the overlapping genes *30.3* and *30.3*' of T4 and 15 T4-related phages are presented in Table. In case of gene *30.3* most of the phages tested have UAAG as a Shine–Dalgarno sequence and AUG as the translation initiation codon. The translation initiation region of gene *30.3*'

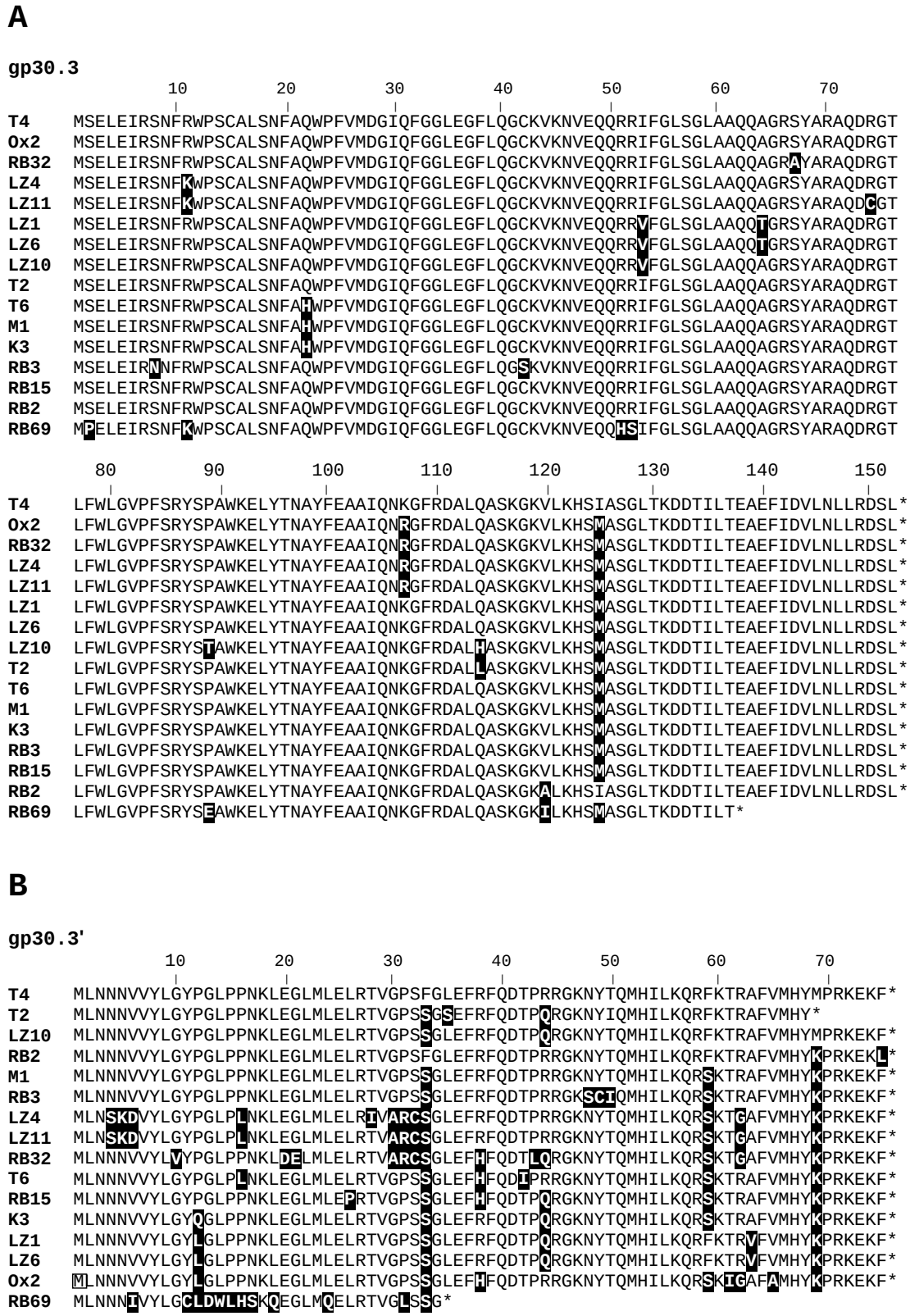


Fig. 3. Comparison of amino acid sequences of gp30.3 (A) and gp30.3' (B) of 16 T4-type bacteriophages. The protein sequences of the 15 phages were aligned with the T4 sequence using the ClustalW program. Amino acids common to all T4-related phages are indicated by white backround. Black background indicates non-conserved amino acids. An asterisk (*) shows the end of the gene. A box indicates methionine coded by the putative translation initiation codon ACG of phage Ox2. The nucleic acid sequences have been deposited in the EMBL/GenBank. Their accession numbers are: AJ457990-AJ457994 (for phages T2, T6, M1, K3 and Ox2, respectively), AJ458392-AJ458395 (RB2, RB3, RB15 and RB32, respectively), AJ458397-AJ458401 (LZ1, LZ4, LZ6, LZ10 and LZ11, respectively) and AJ439452 (RB69)

Table. **Translation initiation regions of genes *30.3* and *30.3'* of phage T4 and 15 T4-related bacteriophages**

Bacteriophage	Gene	Translation Initiation Region *
T4, T6, M1, K3, Ox2, RB2, RB3, RB15, LZ6	*30.3*	UUGAAAUUA**UAAG**CAUUGAAA**AUG**UCUGAG***
T2	*30.3*	UUGAAAUUAUA**AAG**UAUUAAAG**AUG**UCUGAG
RB32, LZ1, LZ4, LZ10, LZ11	*30.3*	UUGAAAUUA**UAAG**UAUUGAAA**AUG**UCUGAG
RB69	*30.3*	UUGAAAUUA**UAAG**UGUUAAAG**AUG**CCUGAG
T4, K3, RB2, LZ4, LZ11	*30.3'*	CAAGGGUG**UAAGG**UGAAAA**AUG**UUGAACA
RB69	*30.3'*	CAAGGAUG**UAAGG**UGAAAA**AUG**UUGAACA
T6, RB15	*30.3'*	CAAGGGUGC**AAGG**UGAAAA**AUG**UUGAACA
Ox2	*30.3'*	CAAGGAUGC**AAGG**UGAAAA**ACG**UUGAACA
RB3	*30.3'*	CAAGGUAGC**AAGG**UGAAAA**AUG**UUGAACA
T2, M1, LZ1, LZ6, LZ10	*30.3'*	CAAGGAUGC**AAGG**UGAAAA**AUG**UUGAACA
RB32	*30.3'*	CAAGGAUGCA**AAG**UGAAAA**AUG**UUGAACA

* SD sequences and the initiation codons of genes *30.3* and *30.3'* are in the underlined bold lettering. The grey shading indicates nucleotides which differ from T4.
*** indicates stop codon of T4 gene *30.4*.

contains the Shine–Dalgarno sequences UAAGG, AAGG and AAG. It should be noted that gene *30.3'* of bacteriophage Ox2 starts with the unusual translation initiation codon ACG. It is likely that the alternate initiation codon ACG is a poor start codon, temperature-sensitive and extremely rare [11]. On the other hand, we can not exclude that in case of phage Ox2 translation of mRNA could be initiated from the second codon of gene *30.3'*, UUG (Table).

Although the middle promoter sequence AGTGGTCTA shows a weak similarity to the –30 consensus sequence [12], the primer extension sequencing of RNA confirmed the presence of the T4 middle promoter P_M30.2 in the coding region of gene *30.3* [6]. The –30 and the –10 sequences of P_M30.2 were identical in T4 and all the phages tested (Fig. 1). Only a few nucleotide differences just upstream the –30 sequence were detected. In case of bacteriophages RB3 and RB69, the potential initiation nucleotides for the transcripts directed from promoter P_M30.2 differed from those of T4.

Received 4 October 2002

ACKNOWLEDGMENTS

This research was supported in part by grant No. 219/2490-1 from the Lithuanian State Program "Molecular Background of Biotechnology".

References

1. Tetart F, Desplats C, Kutateladze M, Monod C, Ackermann HW, Krisch HM. J Bacteriol 2001; 183: 358–66.
2. Ackerman HW, Krisch HM. Arch Virol 1997; 142: 2329–45.
3. Nivinskas R, Zajančkauskaite A, Raudonikiene A, Viteniene I. DNA Sequence. J DNA Sequencing and Mapping 1992; 2: 405–9.
4. Zajančkauskaite A, Malys N, Nivinskas R. Gene 1997; 194: 157–62.
5. Zajančkauskaite A, Raudonikiene A, Nivinskas R. Gene 1994; 147: 71–6.
6. Truncaite L, Zajančkauskaite A, Nivinskas R. J Mol Biol 2002; 317: 179–90.
7. Monod C, Repoila F, Kutateladze M, Tetart F, Krisch HM. J Mol Biol 1997; 267: 237–49.
8. Kutter E, Gachechiladze K, Poglazov A, Marusich E, Shneider M, Aronsson P, Napuli A, Porter D, Mesyanzhinov V. Virus Genes 1996; 11: 285–97.
9. Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, Erlich HA. Science 1988; 239: 487–91.
10. Miller E, Karam J, Spicer E. In:Bacteriophage T4. Karam JD, Editor-in Chief. ASM, Washington, DC 1994; 193–205.
11. Shinedling S, Gayle M, Pribnow D, Gold L. Mol Gen Genet 1987; 207: 224–32.
12. Guild N, Gayle M, Sweeney R, Hollingsworth T, Modeer T, Gold L. J Mol Biol 1988; 199: 241–58.

G. Kolesinskienė, A. Zajančkauskaitė, A. Arlauskas, V. Klausa, R. Nivinskas

T4 GIMININGŲ BAKTERIOFAGŲ PERSIDENGIANČIŲ GENŲ *30.3* IR *30.3*' NUKLEOTIDŲ SEKOS

S a n t r a u k a

Mūsų ankstesniuose tyrimuose buvo nustatyta persidengiančių genų *30.3* ir *30.3*' pora T4 bakteriofage. Mažesnis genas *30.3*' yra visiškai išsidėstęs gene *30.3* kitu atviru skaitymo rėmeliu. Šiame darbe mes nustatėme persidengiančių genų *30.3* ir *30.3*' pirminę struktūrą penkiolikoje T4 giminingų bakteriofagų. Daugelio tirtų fagų šių genų nukleotidų sekos bei šių genų produktų amino rūgščių sekos buvo homologiškos daugiau kaip 85% lyginant su T4. Tik fago RB69 gp30.3' atveju amino rūgščių konservatyvumas lyginant su T4 buvo 64,7%. Įdomu pastebėti, kad fago Ox2 geno *30.3*' galimas transliacijos iniciacijos kodonas yra ACG. Geno *30.3* koduojančioje srityje esanti vidurinio promotoriaus P_M30.2 seka buvo absoliučiai konservatyvi visuose tirtuose faguose.