Cloning and DNA sequence of the 5′-exonuclease gene of bacteriophage T5

A.V. Kaliman, A.I. Krutilina, V.M. Kryukov and A.A. Bayev

Institute of Biochemistry and Physiology of Microorganisms, USSR Academy of Sciences, Pushchino, Moscow region 142292, USSR

Received 30 October 1985

The nucleotide sequence of the *BalI-PstI* fragment of T5 DNA, 1347 bp in length, coding for 5′-exonuclease (D15 gene), has been determined. A coding region of the gene contains 873 bp and is preceded by a typical Shine-Dalgarno sequence. The D15 gene belongs to a cluster, consisting of at least 3 genes, in which a termination codon of a preceding gene overlaps an initiation codon of the following one. The sequence contains an open reading frame for 291 amino acid residues. The molecular mass of the 5′-exonuclease calculated from the predicted amino acid sequence is 33400 Da.

Bacteriophage T5 *DNA sequencing* *D15 gene* *IS1 element*

1. INTRODUCTION

During infection of *Escherichia coli* by bacteriophage T5, the gene D15-encoded nuclease is produced which belongs to early proteins [1]. Enzymatic studies of purified D15 nuclease showed that the enzyme possesses 5′ → 3′ exonuclease activity on both single- and double-stranded DNA and can function (in vitro) also as a T5 DNA specific endonuclease [2]. This product performs 2 functions essential for the phage T5 life cycle: it is necessary for T5 DNA replication [3] and initiation of late transcription [4]. However the exact role of 5′-exonuclease in this process is still obscure. The availability of this protein in the transcription-replication enzyme complex makes it possible to suggest its necessity for the formation and maintenance of the structural integrity of the complex or for the creation of a modified DNA template which serves as a site of nucleation for all the components of the complex [5]. In such a case, the effect of the D15 gene on the turn-on of late genes may be of an indirect character.

This paper deals with the nucleotide sequence of the T5 DNA *BalI-PstI* fragment, 1347 bp in length, coding for 5′-exonuclease (D15 gene), and the amino acid sequence determined from its structure.

2. MATERIALS AND METHODS

Construction and analysis of recombinant plasmids pBR322-T5 and pUC-T5 were performed as in [6,7]. The nucleotide sequence was determined according to Maxam and Gilbert [8]. Recombinational analysis of plasmids by marker rescue with amber mutants of T5 was performed as in [9]. Assay of exonucleolytic activity of T5 D15 nuclease in the crude cell extract and purification of 5′-exonuclease were as described in [2].

3. RESULTS AND DISCUSSION

Cloning of the bacteriophage T5 *PstI*-J fragment and mapping of intact genes D14 and D15 as well as a part of the D12 gene on this fragment were reported earlier [9].

To obtain material for determining the D15 gene sequence, we have selected from our collection of recombinant molecules 2 clones giving hybridiza-

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/86/$3.50

tion with the *Pst*I-J fragment of bacteriophage T5. Restriction enzyme digestion and Southern blot analyses of the plasmids as well as partial determination of the primary structure showed that cloned T5 DNA corresponds to the *Pst*I-J fragment but contains an IS1 element inserted into p602 and p627 plasmids (fig.1).

The results of the recombinational analysis of these plasmids by the marker rescue method were generally the same as in [9], the only difference being that we showed additionally the presence of the D13 gene in the fragment and arrived at a different conclusion as to the size of gene D14. According to our results, a protein of 70 kDa, found in the maxi-cell system, corresponds to the D13 gene product and not to D14 as expected [10] (details of this part of our work will be published separately).

The left part of the cloned fragment was subcloned in the form of *Xba*I, *Bgl*II, *Hpa*I and *Bsp*RI subfragments in pUC9 and pUC19 plasmids, sequenced according to the Maxam-Gilbert method and analysed additionally by a marker rescue technique.

The nucleotide sequence of the *Bal*I (*Bsp*RI)-*Pst*I fragment is shown in fig.2. Analysis of this structure for the open reading frame showed that the major one is found only in the r-strand of DNA and corresponds to the protein consisting of 291 amino acid residues (33.4 kDa). This agrees well with 35 kDa reported by Moyer and Rothe [2] after the partial purification of 5′-exonuclease. The Shine-Dalgarno sequence AGGA is located 12 nucleotides upstream from the initiation codon ATG. There are no sequences typical of the consensus-promoter structure of *E. coli* or early promoters of bacteriophage T5 [11]. Moreover, in addition to gene D15, the sequence shown in fig.1 contains the open reading frames upstream and downstream from the D15 gene. The open reading frame, preceding the gene of 5′-exonuclease, corresponds to the C-terminal end of the D14 gene product, since it was shown that it is plasmid pBB3, not pXX12, which is efficient in the analysis for the D14 am H6b marker rescue. Therefore, the D14 am H5b amber mutation of bacteriophage T5 is located within the *Bal*I-*Xba*I fragment, 135 bp in length. Another open reading frame, adjacent to the C-terminal end of the D15 gene has the Shine-Dalgarno sequence and corresponds to the N-terminal end of the DX gene not identified hitherto, since no amber mutations have been described for the region in the bacteriophage T5 genetic map between the D15 and D17 genes.

All 3 genes are organized in such a way that a termination codon of the preceding gene overlaps an initiation codon of the following one. Similar

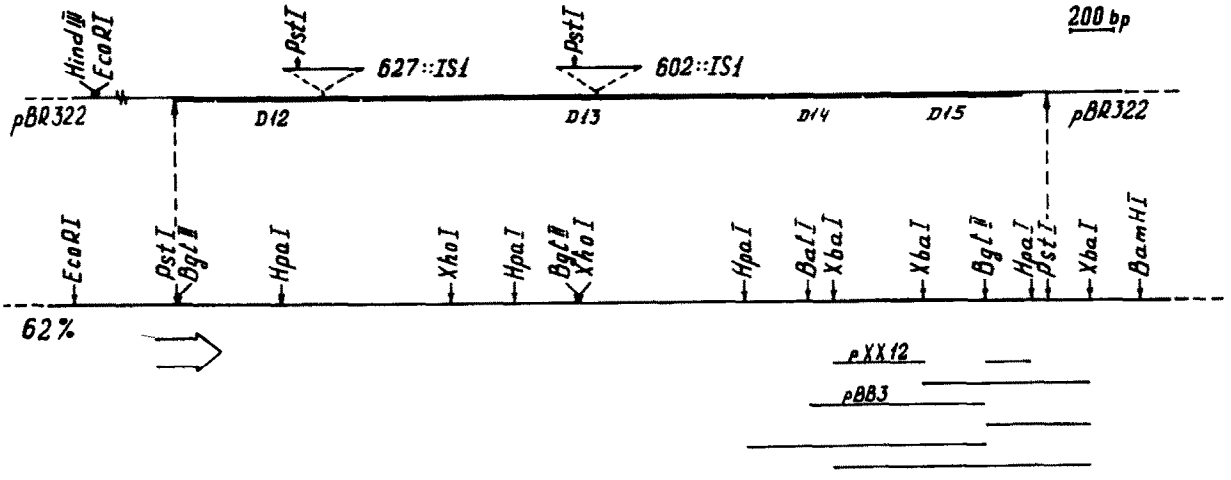

Fig.1. Restriction map of phage T5 *Pst*I-J fragment. In the upper part, the orientation of the cloned fragment in plasmid pBR322, the location of D12–D15 genes on the fragment and sites of IS1 element insertion in plasmids 627 and 602 are shown. In the lower part there are fragments which have been subcloned into plasmids pUC9 and pUC19 for the following determination of the nucleotide sequence. The arrow denotes the direction of gene transcription of this part of the genome.

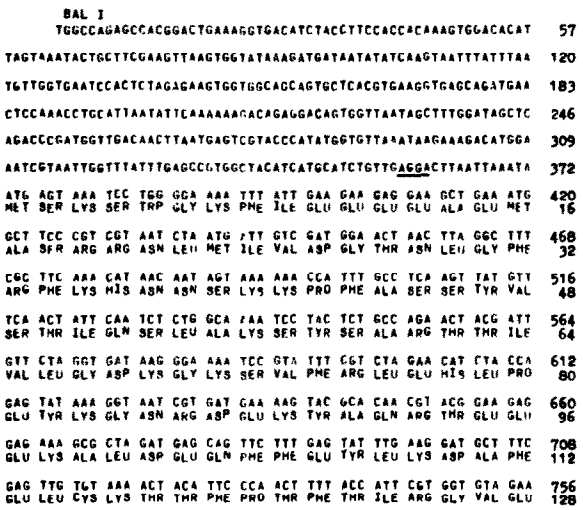

Fig.2. DNA structure of the *Bal*I-*Pst*I fragment and the derived amino acid sequence of 5′-exonuclease of phage T5. Underlined is the Shine-Dalgarno sequence.

structures were reported for the genes of bacteriophage T7 [12], bacteriophage λ [13] and *E. coli* [14]. The availability of open reading frames throughout the sequence is indicative of the fact that the promoter of this group of genes is located outside the sequenced region. Preliminary results show that the DX gene is the last in this cluster of genes and a strong potential terminator occurs in the region of *Bam*HI site.

The study of codon usage in a number of genes [15] revealed a non-random patern of their distribution, the sequence of 5′-exonuclease of phage T5 being no exception in this respect (table 1). There is an obvious preference for the use of codons ending with T and A residues, which might reflect a high AT content of the phage T5 DNA. On the whole, 71.4% codons of the D15 gene end in T and A residues, with the greatest preference for codons ending in T residues (43.6%). Similar data obtained on prokaryotic phages on the whole and ribosomal genes of *E. coli* amount to 39 and 30%, respectively [16].

Table 1

Codon usage in the 5′-exonuclease gene of bacteriophage T5

	T	C	A	G	
T	Phe 10	Ser 5	Tyr 9	Cys 2	T
	Phe 6	Ser 5	Tyr 3	Cys 0	C
	Leu 5	Ser 2	stop	stop	A
	Leu 4	Ser 1	stop	Trp 3	G
C	Leu 4	Pro 2	His 6	Arg 12	T
	Leu 1	Pro 0	His 1	Arg 2	C
	Leu 7	Pro 4	Gln 4	Arg 1	A
	Leu 7	Pro 0	Gln 5	Arg 0	G
A	Ile 16	Thr 6	Asn 7	Ser 3	T
	Ile 2	Thr 2	Asn 5	Ser 0	C
	Ile 3	Thr 5	Lys 15	Arg 1	A
	Met 6	Thr 3	Lys 6	Arg 0	G
G	Val 9	Ala 7	Asp 22	Gly 7	T
	Val 1	Ala 2	Asp 2	Gly 1	C
	Val 4	Ala 8	Glu 13	Gly 9	A
	Val 2	Ala 1	Glu 11	Gly 1	G

An important factor for codon usage and translation efficiency is the availability of no less than 20 tRNA genes in phage T5 [17,18]. Comparison between codons used in the gene D15 structure and anticodons of phage T5 tRNAs sequenced previously [19] shows that tRNAs evidently do not play an important role in the increase in efficiency of this gene translation. It is noteworthy that half of the phage T5 tRNAs studied have G in the first position anticodons, though C is the least occurring residue in the third position codons (11.3% of all the codons). One can suppose that a greater conformity between codons and anticodons coded for by phage T5 should be expected in the case of late genes since tRNA genes belong, similarly to gene D15, to early genes and their active expression coincides in time.

Cell-free extracts obtained from *E. coli* cells containing initial plasmids and their deletion derivatives exhibit increased 5′-exonuclease activity. Moreover, from *E. coli* W 3350 (p627) cells the enzyme was isolated identical to exonuclease from T5-infected cells. Therefore, early genes of phage T5 can also be transcribed by non-modified RNA polymerase. However the level of expression of the D15 gene cloned in pBR322 was not sufficiently high to promote the production of large quantities of homogenous protein necessary for a comprehensive study of its structure and function. We believe that knowledge of the D15 gene structure will help to construct a plasmid which can provide a high and regulated expression of the 5′-exonuclease gene of phage T5.

ACKNOWLEDGEMENTS

Thanks are due to Dr D.J. McCorquodale for T5 mutants, Dr V.N. Ksenzenko for help in constructing recombinant plasmids and discussions of the work, Mrs N. Lunina for help in the use of computer programs and Dr V.V. Velkov for reading of the manuscript and helpful remarks.

REFERENCES

[1] Frenkel, G.D. and Richardson, C.C. (1971) J. Biol. Chem. 246, 4839–4847.
[2] Moyer, R.W. and Rothe, C.T. (1977) J. Virol. 24, 177–193.
[3] Frenkel, G.D. and Richardson, C.C. (1971) J. Biol. Chem. 246, 4848–4853.
[4] McCorquodale, D.J. (1975) Crit. Rev. Microbiol. 4, 213–234.
[5] Ficht, T.A. and Moyer, R.W. (1980) J. Biol. Chem. 255, 7040–7048.
[6] Krutilina, A.I., Ksenzenko, V.N., Kamynina, T.P., Kryukov, V.M. and Bayev, A.A. (1982) Mol. Biol. (USSR) 16, 1253–1262.
[7] Vieira, J. and Messing, J. (1982) Gene 19, 259–268.
[8] Maxam, A.M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560–564.
[9] Davison, J., Brunel, F., Merchez, M. and Ha-Thi, V. (1981) Gene 16, 99–106.
[10] Brunel, F., Davison, J., Ha-Thi, V. and Reeve, J. (1981) Gene 16, 107–118.
[11] Hawley, D.K. and McClure, W.R. (1983) Nucleic Acids Res. 11, 2237–2255.
[12] Dunn, J.J. and Studier, F.W. (1981) J. Mol. Biol. 148, 303–330.
[13] Sanger, F., Coulson, A.R., Hong, G.F., Hill, D.F. and Petersen, G.B. (1982) J. Mol. Biol. 162, 729–773.
[14] Platt, T. and Yanofsky, C. (1975) Proc. Natl. Acad. Sci. USA 72, 2399–2404.
[15] Crantham, R., Gautier, C., Gouy, M., Jacobzone, M. and Mercier, R. (1981) Nucleic Acids Res. 9, 43–74.
[16] Kohli, J. and Grosjean, H. (1981) Mol. Gen. Genet. 182, 430–439.
[17] Hunt, C., Desai, S.M., Vaughan, J. and Weiss, S.B. (1980) J. Biol. Chem. 255, 3164–3173.
[18] Kazantsev, S.I., Tchernov, A.P., Shlyapnikov, M.G., Kryukov, V.M. and Bayev, A.A. (1979) Dokl. Akad. Nauk USSR 247, 744–748.
[19] Sprinzl, M. and Causs, D.H. (1984) Nucleic Acids Res. 12, suppl., r1–r57.