COMPLETE NUCLEOTIDE SEQUENCE OF THE BACTERIOPHAGE λ DNA REGION CONTAINING GENE *Q* AND PROMOTER p_R'

N. A. PETROV, V. A. KARGINOV, N. N. MIKRIUKOV, O. I. SERPINSKI and V. V. KRAVCHENKO*
All-Union Research Institute of Molecular Biology, Novosibirsk 90, USSR

Received 25 August 1981

1. Introduction

It is well known now that the main regulation stages of bacteriophage λ development occur on the level of transcription. The regulation of transcription does not occur only at the initiation stage but also at the termination stage. It has been shown that at the early steps of phage λ development the regulation of termination is controlled by phage-specific protein, product of gene *N*. This protein provides anti-termination of RNA originating from promoters p_L and p_R [1–5].

At the late stage of the phage λ development, promoter p_R' is used for the effective transcription of the lysis and morphology genes [5–8]. A product of λ gene *Q* is necessary for efficient synthesis of late messenger RNA origination from p_R' (reviews [5,8]). Protein *Q* seems not to be an activator of the promoter p_R', but to provide elongation of the short 6 S RNA promoted by p_R' [6,7,9]. It has been proposed [5,6] that this transcript is a 'leader sequence' for late gene expression, and that protein *Q*, like protein *N*, is an anti-terminator protein.

Here, we have determined the primary structure of the λ DNA fragment between 90.8% and 93.1% of the λ genome length. This fragment contains gene *Q*, promoter p_R' and gene 6 S RNA with its terminator.

2. Materials and methods

Restriction endonucleases *Eco*RI, *Bam*HI and *Bsu*I were isolated as in [10–12]. Restriction endonucleases *Msp*I, *Hin*dIII and *Tag*I were isolated according to [13]. DNA-ligase and polynucleotide kinase were kindly provided by Dr Yu. S. Nechaev and alkaline phosphatase and DNA polymerase I (Klenov's fragment) by Dr V. G. Korobko.

λcI857 DNA was isolated as in [9]. A plasmid DNA was isolated according to [14].

[γ-^{32}P]ATP, [α-^{32}P]dATP and [α-^{32}P]dGTP with spec. act. 2000–3000 Ci/mmol were purchased from Amersham Radiochemical Centre (England).

DNA hydrolysis by restriction endonucleases was performed under the conditions in [15]. Electrophoresis of the DNA fragments in 1% agarose or 4% polyacrylamide gels and isolation of the fragments from the gels were performed as in [9].

2.1. *Preparation of plasmid* pQp_R'

The mixture (final vol. 20 μl) containing 0.2 μg large *Eco*RI + *Bam*HI-fragment of pBR322 DNA, 0.2 μg *Eco*RI + *Bam*HI-fragment of λcI857 DNA (see fig.1) and 2 units of DNA-ligase was incubated during 4 h at 37°C in a buffer: 20 mM Tris–HCl (pH 7.6), 10 mM $MgCl_2$, 1 mM dithiothreitol and 1 mM ATP. $CaCl_2$-treated *Escherichia coli* C600 cells were transformed by this mixture and plated on 1.5% LB-agar containing 20 μg ampicillin/ml. The sensitivity of the transformants to tetracycline was checked.

2.2. *Sequencing of DNA fragments*

5′-Ends of DNA fragments (fig.1) were labelled with polynucleotide kinase in the presence of [γ-^{32}P]-ATP [16]. 3′-ends were labelled with DNA polymerase I from *E. coli* (Klenov's fragment) [17]. The primary structure of the 5′- or 3′-labelled subfragments was determined by the Maxam-Gilbert method [16].

* To whom reprint requests should be addressed

Published by Elsevier/North-Holland Biomedical Press
00145793/81/0000–0000/$02.75 © 1981 Federation of European Biochemical Societies

3. Results and discussion

3.1. *Preparation and characterisation of a plasmid containing gene* Q *and promoter* p_R'

According to genetic and physical maps of λ DNA [18] gene *Q* and promoter p_R' are located at the right end of the DNA molecule between 90.8% and 92.9% of the λ genome length. To simplify the sequencing of this region, *Eco*RI + *Bam*HI-fragment of λcI857 DNA containing gene *Q* and 6 S RNA transcription was preliminarily transformed into the plasmid pBR322 (fig.1a,b). The new plasmid was called pQp_R'. Its analysis was performed with restriction endonuclease *Bsu*I. Fig.2 represents the results of gel-electrophoresis of pBR322, λcI857 and pQp_R' DNAs digested by *Bsu*I. It is seen that pQp_R' DNA hydrolysate lacks fragments *Bsu*-192 and *Bsu*-104 (the numbers show the length of the fragments in base pairs) when compared with pBR322 DNA hydrolysate; the intensity of fragment *Bsu*-123 is decreased and 4 new additional fragments appear (1250, 900, 740 and 300, see fig.2).

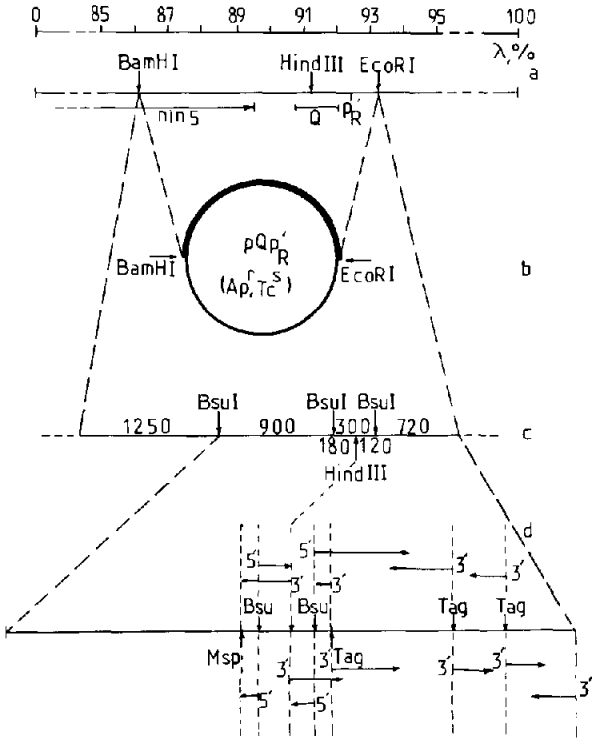

Fig.1. Preparation of plasmid pQp_R' containing gene *Q* and promoter p_R' and strategy of sequencing of gene *Q* and the 6 S RNA transcription: (a) Physical–genetic map of λ DNA in the region of gene *Q* and promoter p_R' according to [18]; (b) physical map of plasmid pQp_R' constructed in vitro on the basis of pBR322 DNA (*Eco*RI + *Bam*HI-fragment from λcI857 DNA is shown by thick line); (c) position of *Bsu*I sites on *Eco*RI + *Bam*HI-fragment containing gene *Q* and promoter p_R'; (d) 5'- and 3'-labelled ^{32}P subfragments used for the sequencing.

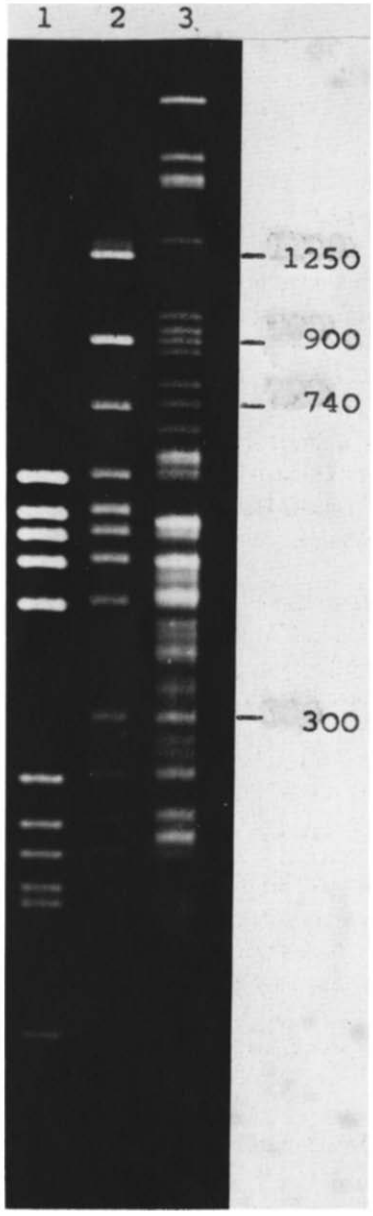

Fig.2. Gel-electrophoresis of plasmid and phage DNAs digested by *Bsu*I: (1) pBR322; (2) pQp_R': (3) λcI857. Numbers 1250, 900, 740 and 300 show the lengths (in basepairs) of the λcI857 DNA fragments contained in plasmid pQp_R'.

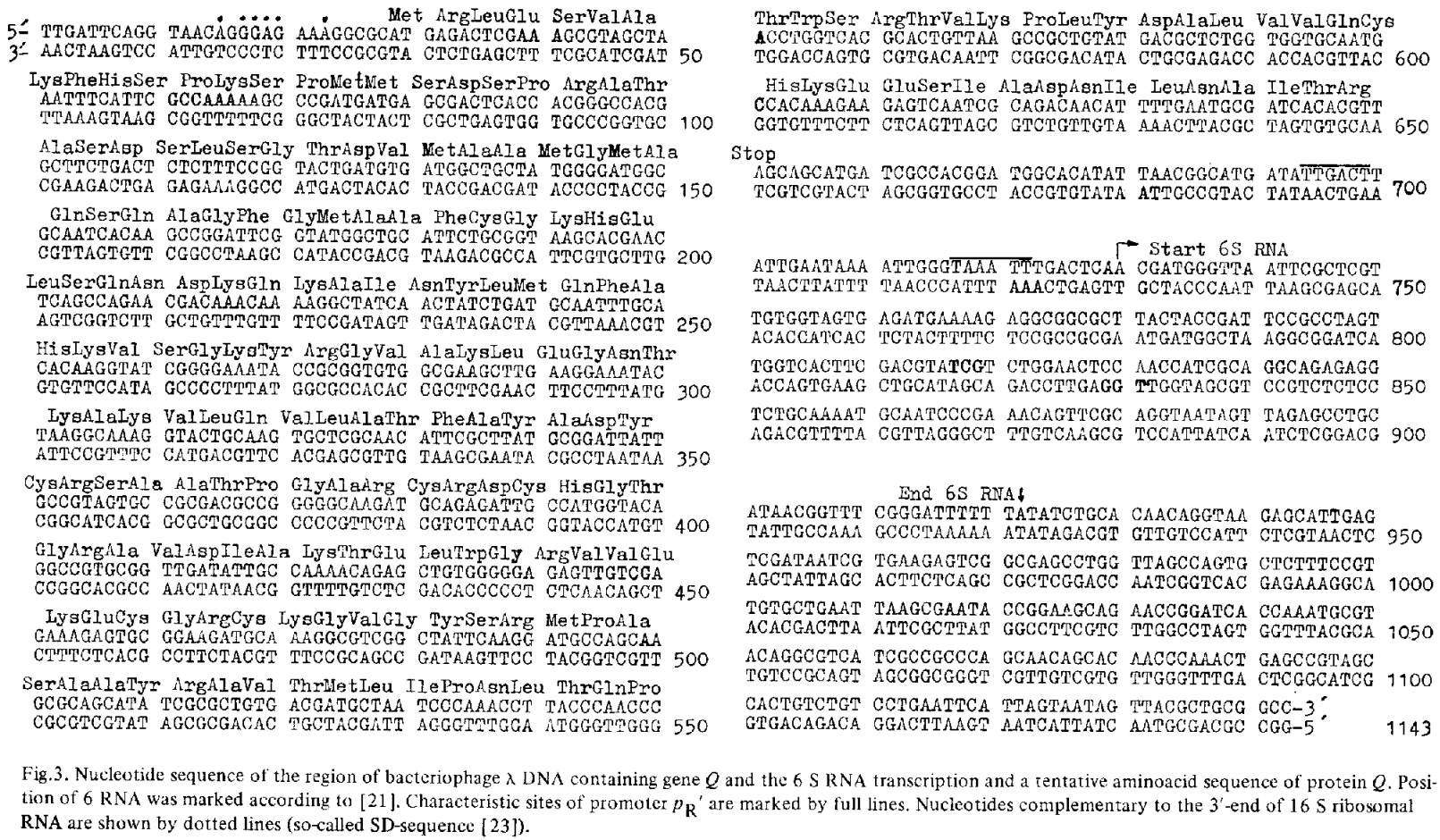

Fig.3. Nucleotide sequence of the region of bacteriophage λ DNA containing gene *Q* and the 6 S RNA transcription and a tentative aminoacid sequence of protein *Q*. Position of 6 RNA was marked according to [21]. Characteristic sites of promoter p_R' are marked by full lines. Nucleotides complementary to the 3′-end of 16 S ribosomal RNA are shown by dotted lines (so-called SD-sequence [23]).

Fragments 1250, 740 and 300 contain *Bam*HI-, *Eco*RI- and *Hin*dIII-sites, respectively (not shown).

Fragments *Bsu*-900, *Bsu*-740 and *Bsu*-300 have analogues among *Bsu*I fragments of λcI857 DNA (compare lines 2 and 3, fig.2). Fragment *Bsu*-740 contains the promoter p_R' [9]. Fragment *Bsu*-900 contains the *nin*5 deletion region since it is absent from the *Bsu*I digest of λ*gt* -λC DNA [19] possessing this deletion [20].

All these data allow us to find out the order of the *Bsu*I-fragments 1250, 900, 740 and 300 on the physical map of the pQp_R' plasmid (fig.1c). Besides, it can be concluded from the restriction analysis that pQp_R' DNA contains the λcI857 DNA fragment with the *Q* gene and 6 S RNA transcripton – promoter p_R' and the gene for 6 S RNA with its terminator.

3.2. *Sequencing of gene Q and 6 S RNA transcripton*

After comparison of the physical -genetic map of λcI857 DNA and the *Bsu*I-physical map of pQp_R' plasmid one can propose that the 6 S RNA transcription is located in the *Bsu*I + *Eco*RI-720 fragment and the sequence coding for the *Q* gene is divided into 3 fragments. The beginning and the end of the gene are located within the fragments *Bsu*-900 and *Bsu*I + *Eco*RI-720, respectively and the central part of it appears in the fragment *Bsu*-300 (fig.1).

The strategy of sequencing of these 3 fragments is shown in fig.1d. The determined primary structure is shown in fig.3. It contains the complete sequence of 6 S RNA, transcription of which is initiated from the promoter p_R' [6,7,9] (compare the sequence of the region 730–922 with data in [21]). From the left side of 6 S RNA initiation site there are two characteristic clusters (see fig.3, regions 694–699 and 717–722). They have a very high extent of homology with the recognition sites of the promoter for RNA polymerase from *E. coli* [22]. Hence, it can be concluded that the region 690–922 is a 6 S RNA transcripton containing promoter p_R' and 6 S RNA sequence.

According to the genetic map of λ DNA gene *Q* is located from the right side of the promoter p_R' [18]. In fact one can see adenosine base of the initiating AUG triplet in position 28 (fig.3). Before this adenosine there is a sequence complementary to 3′-end of 16 S ribosomal RNA (fig.3). A sequence from 28 adenosine could be translated permanently up till the terminator in position 650–652, and a polypeptide of 207 amino acids could be formed. The amino acid sequence of the polypeptide is shown in fig.3, its calculated M_r is 23 114, in accordance with the M_r of the protein *Q* (23 000 [18]). These data lead to the conclusion that the region between 28–649 nucleotides codes the protein *Q*.

Hence, we have determined a complete nucleotide sequence of gene *Q* and 6 S RNA transcription (promoter p_R' and 6 S RNA gene) as a target of the protein *Q* action. The report of the primary structure of protein *N* [24] allows one to compare these two proteins with similar functions. Such theoretical comparison is not the subject of this work but it seems to be very interesting in order to understand some details of the mechanism of the anti-termination effected by *N* and *Q* proteins.

Acknowledgements

The authors are thankful to Dr Yu. S. Nechaev for DNA-ligase and polynucleotide kinase, and to Dr V. G. Korobko for alkaline phosphatase and DNA polymerase I.

References

[1] Roberts, J. W. (1969) Nature 224, 1168–1174.
[2] Echols, H. (1971) in: The Bacteriophage Lambda (Hershey, A. D. ed) pp. 247–270, Cold Spring Harbor Lab. NY.
[3] Thomas, R. (1971) in: The Bacteriophage Lambda (Hershey, A. D. ed) pp. 211–220, Cold Spring Harbor Lab. NY.
[4] Couturier, M., Dambly, C. and Thomas, R. (1973) Mol. Gen. Genet. 120, 231–252.
[5] Herskowitz, I. (1974) Annu. Rev. Genet. 7, 289–324.
[6] Roberts, J. W. (1975) Proc. Natl. Acad. Sci. USA 72, 3300–3304.
[7] Sklar, J., Yot, P. and Weissman, S. M. (1975) Proc. Natl. Acad. Sci. USA 72, 1817–1821.
[8] Szybalski, W. (1977) in: Regulatory Biology (Copeland, J. C. and Marzluff, G. A. eds) pp. 3–45, Ohio State University Press, Columbus OH.
[9] Kravchenko, V. V., Vassilenko, S. K., Gileva, I. P. and Zaychev, B. N. (1980) Bioorg. Khim. 6, 217–222.
[10] Thomas, M. and Davis, R. (1974) J. Mol. Biol. 91, 315–328.
[11] Wilson, G. A. and Young, F. E. (1975) J. Mol. Biol. 97, 123–125.
[12] Bron, S., Murray, K. and Trautner, T. A. (1975) Mol. Gen. Genet. 143, 13–23.
[13] Green, P. J., Heyneker, H. L., Bolivar, F., Rodriguez, R. L., Betlach, M. C., Covarrubias, A. A., Backman, K., Russel, D. J., Tait, R. and Boyer, H. W. (1978) Nucleic Acids Res 8, 2373–2380.

[14] Guerry, P., Le Blanc, D. J. and Falkow, S. (1973) J. Bacteriol. 116, 1064–1066.
[15] Kravchenko, V. V., Vassilenko, S. K. and Grachev, M. A. (1979) Gene, 7, 181–195.
[16] Maxam, A. M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560–564.
[17] Petrov, N. A., Serpinski, O. I., Mikriukov, N. N., Karginov, V. A., Kravchenko, V. V. and Vassilenko, S. K. (1979) Bioorg. Khim. 5, 1873–1875.
[18] Szybalski, E. H. and Szybalski, W. (1979) Gene 7, 217–270.
[19] Vassilenko, S. K., Grachev, M. A., Zaychikov, E. F. and Kravchenko, V. V. (1978) Dokl. Akad. Nauk SSSR 241, 710–713.
[20] Thomas, M., Cameron, J. and Davis, R. (1974) Proc. Natl. Acad. Sci. USA 71, 4579–4583.
[21] Sklar, J. L. (1977) PhD. Thesis, Yale University, New Haven CT.
[22] Siebenlist, U., Simpson, R. B. and Gilbert, W. (1980) Cell 20, 269–281.
[23] Shine, J. and Dalgarno, L. (1975) Nature 254, 34–38.
[24] Franklin, N. C. and Bennett, G. N. (1980) Gene 8, 107–109.