THE NUCLEOTIDE SEQUENCE OF GENE 1 OF T7 PHAGE DNA WHICH CODES FOR THE PHAGE-SPECIFIC DNA-DEPENDENT RNA-POLYMERASE

M. A. GRACHEV and A. G. PLETNEV

Novosibirsk Institute of Organic Chemistry, Siberian Division of the USSR Academy of Sciences, Novosibirsk 90, USSR

Received 29 February 1981

1. Introduction

Gene 1 of T7 phage codes for the phage-specific DNA-dependent RNA-polymerase [1,2] ~100 000 M_r. This enzyme interacts selectively with promotors present in T7 DNA [3] and synthesizes 'late' mRNAs at twice the rate exhibited by the RNA-polymerase of *Escherichia coli;* it also stops the synthesis at selective terminators. Hence, T7 RNA-polymerase, in spite of its relative simplicity, fulfils successfully the three major functions of a template-dependent enzyme, and its chemical study for this reason is of great interest. However, not knowing the amino acid sequence of the protein, it is difficult to interpret on a molecular level the data of chemical methods of investigation, especially of such a powerful method as affinity modification.

This communication gives an account of investigation [4,5] aimed at the elucidation of the primary structure of T7 DNA in the region of gene 1; from this sequence, it is possible to deduce the amino acid sequence.

2. Materials and methods

In order to elucidate the structure within gene 1, T7 DNA was cleaved by restriction endonucleases followed by analysis [6] as modified in [7,8]. The strategy of sequencing is shown in fig.1. T7 DNA was cleaved in different experiments by endonucleases *Bsp*I, *Msp*I, *Hin*dII, *Hpa*I; the digest were fractionated by gel-electrophoresis, and the fragments mapped [4,9–11] within gene 1 and in its vicinity were labelled terminally according to [6] using either [γ-^{32}P]ATP and polynucleotide kinase, or d[α-^{32}P]-NTPs and DNA-polymerase. The sequencing experiments were run either after chains separation, or after treatment by another endonuclease. A considerable part of the sequence was read-off from 2 complementary strands. If only 1 of the 2 complementary strands was sequenced, the experiments were run in triplicate. Details of the experiments will be published elsewhere.

3. Results and discussion

The sequence of nucleotides found, which is situated between 7–15% of the T7 genome, is presented in fig.2. T7 DNA has been sequenced independently over the regions 110–490 and 2771–3175 [11–14]. Our sequence over these regions is practically the same as [11–14] except for a few pyrimidine–pyrimidine substitutions. The restrictase sites suggested by the primary structure are in complete accordance with the data of mapping [4,9–11].

According to [11], the adenosine residue of the initiating triplet for T7 RNA-polymerase resides at position 316. The sequence between this residue and the residue 2965, where a terminating triplet is present, is translatable (but in the frame given) into an uninterrupted protein sequence. The molecular mass of the protein presented in fig.2 is 98 000, in good accord with the above cited data from physical studies [1]. The amino acid composition is in reasonable agreement with the data of direct amino acid analysis [2].

Acknowledgements

The authors are thankful to Dr E. F. Zaychikov and to Mrs T. G. Maksimova for participation in the first stages of the present investigation, to Drs V. G. Korobko, S. A. Grachev and S. V. Netesov for the gifts of restriction endonucleases and to Professor D. G. Knorre for encouragement and critical discussions.

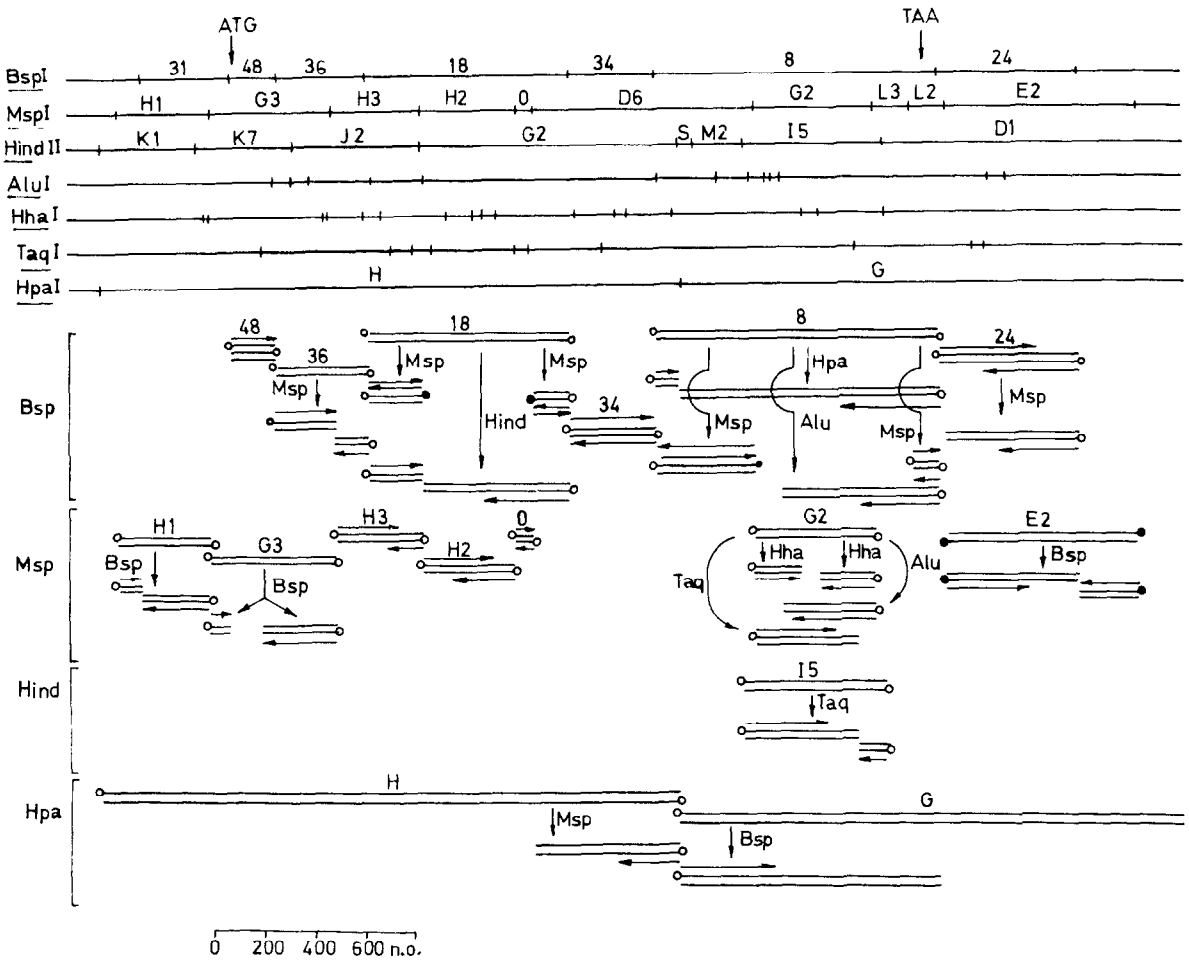

Fig.1. Strategy of the elucidation of the primary structure of T7 DNA in the region of gene 1: upper part, restrictase maps according to [4,9–11]; lower part, scheme of the terminal labelling ((○) at the 5′-end; (●) at the 3′-end) and of sequencing (——→).

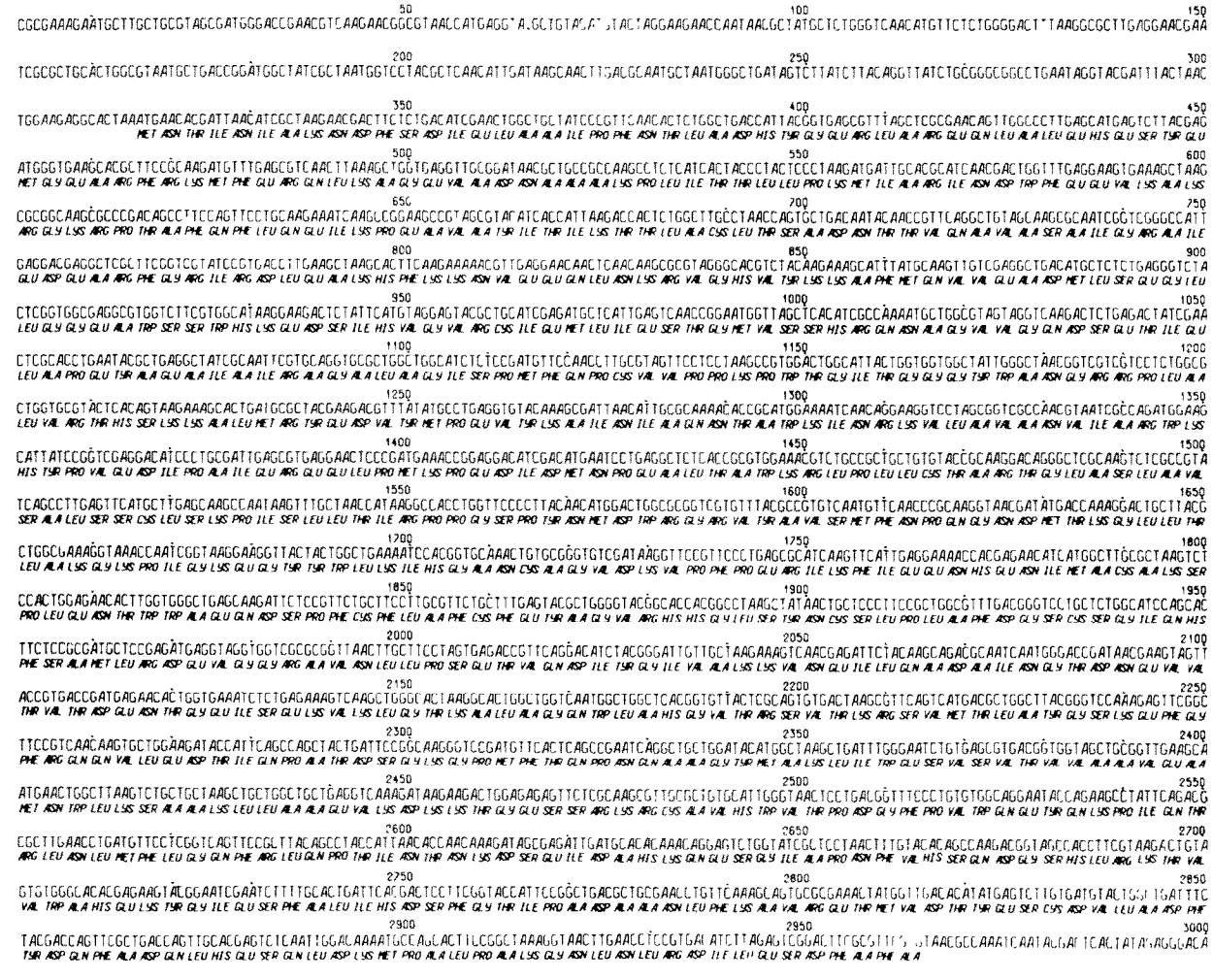

Fig.2. The sequence of nucleotides in the region of gene 1 of T7 DNA and the sequence of amino acids deduced.

References

[1] Chamberlin, M., McGrath, J. and Weskell, L. (1970) Nature 228, 227–231.
[2] Niles, E. G., Conlon, S. W. and Summers, W. C. (1974) Biochemistry 13, 3904–3912.
[3] Golomb, M. A. and Chamberlin, M. (1974) J. Biol. Chem. 249, 2858–2863.
[4] Grachev, M. A., Zaychikov, E. F., Maksimova, T. G. and Pletnev, A. G. (1979) Bioorg. Khim. 5, 1587–1590.
[5] Grachev, M. A. and Pletnev, A. G. (1980) Bioorg. Khim. 6, 1737–1739.
[6] Maxam, A. M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560–564.
[7] Korobko, V. G. and Grachev, S. A. (1977) Bioorg. Khim. 3, 1420–1422.
[8] Korobko, V. G., Grachev, S. A. and Kolosov, M. N. (1978) Bioorg. Khim. 3, 1281–1283.
[9] Studier, F. W., Rosenberg, A. H., Simon, M. N. and Dunn, J. J. (1979) J. Mol. Biol. 135, 907–915.
[10] Panayotatos, N. and Wells, R. D. (1979) J. Mol. Biol. 135, 91–109.
[11] McConnell, D. J. (1979) Nucleic Acids Res. 6, 3491–3503.
[12] Oakley, J. L., Strothkamp, R. E., Sarris, A. H. and Coleman, J. F. (1979) Biochemistry 18, 528–537.
[13] Panayotatos, N. and Wells, R. D. (1979) Nature 280, 35–39.
[14] Saito, H., Tabor, S., Tamanoi, F. and Richardson, C. C. (1980) Proc. Nat. Acad. Sci. USA 77, 3917–3921.