THE AMINO ACID SEQUENCE OF A DNA BINDING PROTEIN, THE GENE 5 PRODUCT OF fd FILAMENTOUS BACTERIOPHAGE

Yasutsugu NAKASHIMA, A.K. DUNKER, D.A. MARVIN and William KONIGSBERG
Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06510, USA

Received 25 January 1974

1. Introduction

The F-specific filamentous bacterial viruses such as fd, f1 and M13 synthesize single-stranded circular viral DNA on a duplex DNA intermediate [1]. Synthesis of this single-stranded DNA requires the products of viral gene 2 [2, 3] and gene 5 [4] as well as rifampicin-sensitive funtion [5].

The gene 5 product (5p), a protein having a molecular weight of 9830 daltons, forms a complex with intracellular single-stranded DNA preventing the synthesis of a complementary DNA strand [6]. During maturation of the virus, the 5p protein is completely displaced from the DNA by the coat protein associated with the bacterial membrane [7]. Because of the importance of 5p in DNA replication and phage maturation, we have determined the primary structure of 5p, as a necessary step prior to studying the molecular details of the interaction of 5p with DNA in solution and in crystals.

2. Materials and methods

2.1. *Growth and isolation of 5p*

E. coli DM48 (derived from strain S26 provided by A. Garen (Hohn et al. [8]) grown to a density of 10^8 cells/ml was infected at a multiplicity of 10 with wild-type fd bacteriophage [9]. The cells were harvested 6 hr after infection, disrupted by stirring with glass beads in a Waring blender, and the 5p protein purified according to Alberts et al. [10]. About 80 mg of 5p could be obtained from 100 g of cells. The purity of 5p was about 95% as shown by polyacrylamide gel electrophoresis [11].

2.2. *Sequence analysis of 5p*

Oxidation, reduction, radioalkylation, cleavage at cysteine and CNBr cleavage was performed by standard methods [12–15]. Tryptic and chymotryptic digestion was carried out as previously described [16, 17]. The resulting peptides were separated by ion-exchange chromatography, gel-filtration and by paper electrophoresis. Automatic sequencing of 5p was performed on a JOEL-JAS-47K sequenator using a protein program [18]. The dansyl-Edman method [19] was used to sequence some of the smaller tryptic and chromotryptic peptides. Amides were determined by thin layer chromatography of the released PTH amino acids and by paper electrophoresis of the smallest peptide containing an Asx or Glx residue [20].

3. Results and discussion

The gene 5 protein contains 88 amino acid residues. The results we have obtained up to now permit us to propose the primary structure shown in fig. 1. The evidence for the proposed sequence was obtained as described below.

The 5p was oxidized with performic acid to increase its solubility in aqueous solution, thus making it possible to carry out enzymatic digestions and

 North-Holland Publishing Company – Amsterdam

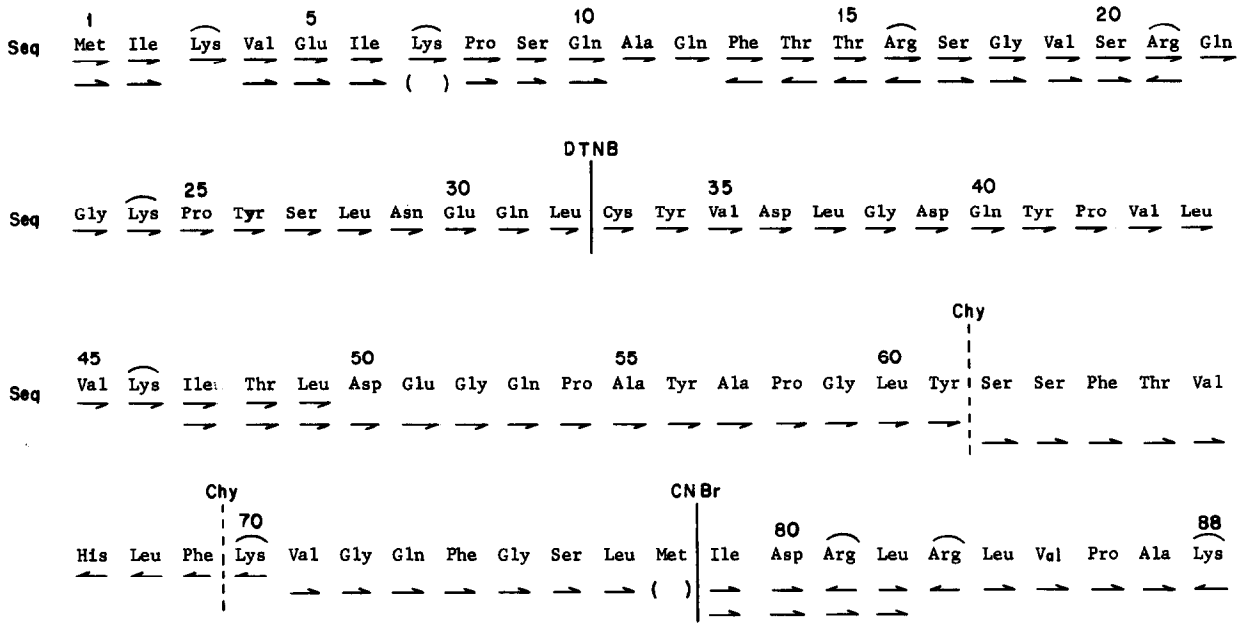

Fig. 1. Amino acid sequence of the gene 5 protein. Seq ——→; sequences established by the JEOL-JAS-47K sequenator: ——→ sequences established by dansyl-Edman degradation: ←——, sequences determined by carboxypeptidase A and B digestion: ( ); dansyl amino acid could not be positively identified: DTNB; site of cleavage by 5,5′-dithiobis 2-nitrobenzoate and KCN: Chy; sites of cleavage by chymotrypsin: CNBr; site of cleavage by cyanogen bromide.

Edman degradations without difficulty. The oxidized 5p (20 mg) was first submitted to automatic Edman degradation using the JOEL-JAS-47K sequenator. The sequence of 49 residues from the NH_2-terminus was established in this way. After digestion of oxidized 5p with trypsin, eight peptides were isolated and the sequence of the shorter peptides (10 residues of less) were determined by the dansyl-Edman degradation and carboxypeptidase A and B digestion. For the tryptic peptides that include residues 1–21, these results confirmed the data already obtained from the sequenator. For other peptides, the sequences obtained by these methods were necessary, since the automatic Edman degradation on oxidized 5p gave information only through residue 49. Thus residues 71–88 were assigned by using the dansyl-Edman and carboxypeptidase methods. The sequenator was also used on a large tryptic peptide (residues 47–70) and the results obtained extended the sequence of 5p to residue 61. The remaining portion, from 62–70, that was not determined by direct sequencing of 5p or by dansyl-Edman degradation of the tryptic peptides, was established using the dansyl-Edman method and carboxypeptidase A digestion on a chymotryptic peptide that included residues 62–69.

The order of the first six tryptic peptides was established by automatic Edman degradation of 5p and the order of the remaining four was determined by isolation and sequencing of the COOH-terminal CNBr peptide from residues 79–88. Chymotryptic peptides derived from other regions of 5p were also isolated. Their compositions and sequences were consistent with the proposed structure. There is one cysteine residue in 5p and treatment with 5,5′-dithiobis 2-nitrobenzoate (DTNB) followed by potassium cyanide [14] produced two fragments, one from residue 1–32 and the other from residues 33–88, a result entirely in accord with the proposed structure of 5p. It is not known what role, if any this sulfhydryl group plays in the function of 5p.

Inspection of the sequence shows that the basic residues, except for lysine 46 are clustered near the NH_2- and COOH-termini whereas the middle of the sequence has a considerable number of hydrophobic residues. It has been shown in several laboratories that 5p binds four nucleotides per monomer [10, 21,

22]. Data on the spectral titration of a 5p-DNA complex show that tyrosine residues are probably involved in the interaction of 5p with DNA [22]. On the basis of circular dichroism studies, it has been suggested that less than 10% of the protein is in the alpha helix. The distribution of the six proline residues at positions 8, 25, 42, 54, 58 and 86 are in accord with this concept. We have isolated several temperature-sensitive mutants of fd that map in gene 5. The 5p isolated from these mutants has an altered affinity for DNA. We are trying how to identify the type and location of the amino acid substitutions in an attempt to get a better understanding of the 5p-DNA interaction.

Acknowledgements

We wish to thank Mr. E.A. Anderson for help with preparation of the gene 5 protein and Mr. Edward Prawdzik for some of the amino acid analyses. This work was supported in part by U.S. Public Health Service Research Grants No. CA-13094 and GM-12607.

References

[1] Ray, D.S. (1969) J. Mol. Biol. 43, 631.
[2] Lin, N.S.-C. and Pratt, S. (1972) J. Mol. Biol. 77, 37.
[3] Tseng, B.Y. and Marvin, D.A. (1972) J. Virology 10, 384.
[4] Pratt, D. and Erdahl, W.S. (1968) J. Mol. Biol. 37, 181.
[5] Staudenbauer, Q.L. and Hofschneider, P.H. (1972) Eur. J. Biochem. 30, 403.
[6] Salstrom, J.S. and Pratt, D. (1971) J. Mol. Biol. 61, 489.
[7] Smilowitz, H., Carson, J. and Robbins, P.W. (1972) J. Supramol. Struct. 1, 8.
[8] Hohn, B., Lechner, H. and Marvin, D.A. (1971) J. Mol. Biol. 56, 143.
[9] Marvin, D.A. and Hoffman-Berling (1963) Nature 197, 517.
[10] Alberts, B.M., Frey, L. and Delius, H. (1972) J. Mol. Biol. 68, 139.
[11] Davis, B.J. (1964) Ann. N.Y. Acad. Sci., 121, 404.
[12] Hirs, C.H.W. (1956) J. Biol. Chem., 219, 611.
[13] Hirs, C.H.W. (1967) in: Methods in Enzymology (Hirs, C.H.W., ed.) Vol. XI, p. 199, Academic Press, New York.
[14] Vanaman, T.C., and Stark, G. (1970) J. Biol. Chem. 245, 3565.
[15] Gross, E. and Witkop, B. (1962) J. Biol. Chem. 237, 1856.
[16] Konigsberg, W., Weber, K., Notani, G. and Zinder, N.D. (1966) J. Biol. Chem. 241, 2579.
[17] Weber, K. and Konigsberg, W. (1967) J. Biol. Chem. 242, 3563.
[18] Hermodsen, M.A., Ericsson, L.H., Neurath, H. and Walsh, K.A. (1973) Biochem. 12, 3146.
[19] Gray, W.R. (1972) in: Methods in Enzymology (Hirs, C.H.W. and Timasheff, S.N. ed.) Vol. XXV, part B, p. 121, Academic Press, New York.
[20] Offord, R.E. (1966) Nature 211, 591.
[21] Oey, J.L. and Knippers, R. (1972) J. Mol. Biol. 37, 181.
[22] Day, L. (1973) Biochem. 12, 5329.