THE AMINO ACID SEQUENCE OF GENE 5 PROTEIN OF BACTERIOPHAGE M 13

Theo Cuypers[+], Frans J. van der Ouderaa[*] and Wilfried W. de Jong[*]

[*]Department of Biochemistry and [+]Laboratory of Molecular Biology,
University of Nijmegen, Nijmegen, The Netherlands

Received May 2,1974

SUMMARY

The amino acid sequence of gene 5 protein of bacteriophage M 13 was deduced by analysis of peptides derived from the protein by cleavage with cyanogen bromide and by digestion with trypsin and chymotrypsin. A combination of sequential degradation procedures has been applied: the automatic solid-phase technique and the manual method with direct identification of the phenylthiohydantoin derivatives, and the dansyl-Edman technique. The polypeptide chain of gene 5 protein contains 87 residues, accounting for a molecular weight of 9688.

Bacteriophage M 13 belongs to the F-specific filamentous coliphages, which have a single-stranded, circular DNA genome and a replication cycle which closely resembles that of the coliphages φx174 and S13 (1). Synthesis of progeny viral DNA occurs on a double-stranded circular DNA intermediate (RF-DNA) and requires hostcell functions (2) as well as the proteins encoded by the viral gene 2 (3, 4) and gene 5 (5-7).
The gene 5 protein enhances the asymmetric displacement synthesis of viral DNA (8) and acts as a repressor by preventing the conversion of single-stranded DNA to double-stranded RF molecules (5-7), most likely by complex formation with single-stranded DNA (6, 7, 9). In the infected cell nearly all progeny viral strands are complexed with gene 5 protein (mol.wt 10,000) to an extent of about 1300 monomers per DNA strand (7, 9, 10). During phage assembly these monomers are eventually displaced by the coat proteins present in the bacterial inner membrane.
For a better understanding of the structural role of gene 5 protein in

phage maturation and DNA-replication, knowledge of the primary structure of this DNA-binding protein is a prerequisite.

MATERIALS AND METHODS

A 50 l culture of E. coli K 38, grown to a density of $2x10^8$ cells/ml, was infected with bacteriophage M 13 wild-type (obtained from Dr. P.H. Hofschneider) at a multiplicity of 10-15. After 2 h at 37°C the infected cells were harvested, yielding 60-75 g wet cells, and the cell extract was prepared essentially as described by Alberts et al (9).

Gene 5 protein was isolated from the cell extract by chromatography on a column of denatured calf thymus DNA-cellulose (9), eluted with stepwise increasing concentrations of NaCl. The 0.6 M NaCl fraction contained the pure protein. The 0.4 M and 2.0 M NaCl fractions also contained a considerable quantity of gene 5 protein and were purified on a Sephadex G-50 fine column in 7 M urea - 20% formic acid. The 0.6 M fraction and the Sephadex G-50 purified fractions were desalted over Sephadex G-25 in 0.1 M NH_3 or 20% formic acid, respectively. The final yield was 50 - 70 mg gene 5 protein from 60 g wet cells.

Up to 50 mg gene 5 protein was reduced and aminoethylated (11), and desalted over a Sephadex G-25 column in 20% formic acid.

After cyanogen bromide cleavage in 70% formic acid (12) the resulting fragments were separated on a Sephadex G-50 fine column (120x1.5 cm) eluted with 20% formic acid. The small cyanogen bromide fragment CB2 could easily be isolated by paper electrophoresis at pH 6.5.

Special precautions had to be taken to achieve satisfactory enzymic digestions. Lyophilized gene 5 protein does not dissolve under usual digestion conditions, resulting in poor yields of enzymic peptides. Tryptic, chymotryptic and carboxypeptidase digestions of native gene 5 protein were carried out in 0.1 M NH_3, brought to pH 8.9 with solid CO_2, at a

protein concentration of 1 mg/ml. Tryptic digestion of aminoethylated gene 5 protein was performed after dissolving the lyophilized protein in 20% formic acid, at a concentration of 80 mg/ml, diluting this solution 20 times and bringing it to pH 8.9 with 25% NH_3. Tryptic and chymotryptic digestions were carried out at 37°C for 2 h and 1 h, respectively, using an enzyme/substrate ratio of 2% (w/w). The tryptic peptide T6 was digested with thermolysin (1 mg/μmole) in 0.1 M NH_4HCO_3, pH 8.9, for 1 h at 37°C. Analytical peptide mapping and preparative isolation of peptides by means of gel filtration, high-voltage paper electrophoresis at pH 6.5 and paper chromatography were carried out as described previously (13). Performic acid oxidation of peptides was carried out on the electrophoresis paper (14).

Manual direct Edman degradation was performed essentially as described by Niall and Potts (15). Lysyl and S-aminoethylcysteinyl peptides were reacted with 4-sulfophenylisothiocyanate (16) prior to the first step (13). Solid phase sequential degradation was carried out according to Laursen (17) in an automatic apparatus, obtained from LKB, built by Sequemat Inc. Peptide T5 was coupled to aminopolystyrene resin (18) and T6 was attached to the phenylenediisothiocyanate derivative of aminopropyl glass (19). Phenylthiohydantoin derivatives of amino acids were identified by thin-layer and gas chromatography (13). The dansyl-Edman procedure was carried out as described by Gray and Smith (20). Identification of the dansyl amino acids was established by thin-layer chromatography (21).

Amide groups were assigned by direct identification of the phenylthiohydantoins of glutamine and asparagine, and confirmed by the electrophoretic mobilities of peptides at pH 6.5 (22).

RESULTS AND DISCUSSION

The amino acid compositions of gene 5 protein and its two CNBr fragments

Table 1. Amino acid compositions of gene 5 protein, its cyanogen bromide fragments and its tryptic and chymotryptic peptides. Purification procedures were: gel filtration on Sephadex G-50 fine (S); electrophoresis at pH 6.5 (E); paper chromatography (C) (13). Electrophoretic mobility was taken at pH 6.5, relative to arginine (=100).

	Gene 5 Pr.		CNBr Fragm.		Chymotryptic Peptides								
	comp.	seq.	CB1	CB2	C1	C2	C3	C4	C5	C6	C7	C8	C9
Asp	5.1	5	4.1	0.9			1.1	3.2				1.1	
Thr	4.1	4	3.5			1.7		0.9	0.7				
Ser	6.9	7	6.3		1.0	1.7	1.0		1.9		1.0		
Glu	10.6	10	10.0		3.1	1.0	2.2	3.2		1.0			
Pro	6.2	6	5.1	1.0	1.1	0.9		2.8					1.0
Gly	7.2	7	6.8			1.9		3.0		1.0	0.9		
Ala	4.1	4	3.2	1.1	1.2			2.0					1.0
Cys	0.8	1	+				1.0[b]						
Val	8.0	8	7.2	1.0	1.1	1.0		2.8	1.1	1.0			1.0
Met	2.0	2	+a								1.1		
Ile	3.9	4	3.2	0.9	1.7			1.0				0.9	
Leu	9.8	10	8.5	1.9			1.9	4.1	1.1		1.1	1.1	1.0
Tyr	4.6	5	4.7			0.9	0.8	2.7					
Phe	3.0	3	2.9		1.0				1.0	1.0			
Lys	5.8	6	5.2	1.0	1.8	0.9		1.2		0.9			1.0
His	1.1	1	1.2						0.9				
Arg	3.9	4	2.4	2.2		2.1						0.9	1.0
Sum	87	87	76	10	12	13	8	27	7	5	4	4	6
Mobil.				63	13	82	-30	-30	33	55	0	0	92
Purif.			S	E	EC	EC	SEC	SEC	SEC	SEC	SEC	SEC	SEC

	Tryptic Peptides										
	T1	T2	T3	T4	T4a	T4b	T5	T6	T7	T8	T9
Asp				1.0		1.0	1.9	1.0	1.1		
Thr		2.0						1.8			
Ser		1.0	1.7	1.0		1.0		1.8	1.0		
Glu		3.1		3.1	1.0	2.2	1.1	2.2	1.2		
Pro		1.0		1.0	1.2		1.1	2.0			1.0
Gly			1.1	1.1	0.9		1.1	2.1	2.0		
Ala		1.1						2.1			1.0
Cys				+[c]		+[c]					
Val		1.0	1.1				3.0	1.0	0.9		1.1
Met	0.8								0.7		
Ile	1.0	1.0						0.9	1.0		
Leu				2.0		1.8	2.0	2.8	1.1	0.9	0.9
Tyr				1.0	0.9		1.9	1.8			
Phe		1.0						1.0	1.0		
Lys	1.0	1.0		1.0	0.9		1.0	1.1			1.0
His								1.0			
Arg		1.1	1.0						0.9	1.1	
Sum	3	13	5	12	5	7	13	23	11	2	5
Mobil.	67	25	58	27	48	0	-16[d]	0[d]	0	75	56
Purif.	SEC	SEC	SEC	SEC	SEC	SEC	S	S	SEC	SEC	SEC

[a] determined as homoserine
[b] determined as cysteic acid
[c] determined as aminoethyl-cysteine
[d] gel filtration in 7% formic acid

are given in Table 1. As already known from the litterature, gene 5 protein does not contain tryptophan (7). The analyses of the tryptic and chymotryptic peptides, also given in Table 1, accounted for all 87 residues present in gene 5 protein. Peptide T4 was apparently partially split by chymotrypsin-like activity in the trypsin preparation, resulting in peptides T4a and T4b.

The compiled evidence for the sequence of gene 5 protein is summarized in Fig. 1. The isolated tryptic peptides were sequenced using several methods, as indicated in the figure. The sequence of the large peptide T6 was established by manual Edman degradation up to residue 17; by performing a run in the automatic solid-phase sequenator, residue 22 could still be identified. The sequence of chymotryptic peptide C5 completed the structure of T6.

The order of the first three tryptic peptides was established by direct Edman degradation of gene 5 protein up to residue 19. The overlap T3-T4 was evident from the composition of C2. T5 must follow T4 because of the sequence of C3 and in agreement with the observation that both T4 and T5 are absent on the peptide map of the non-aminoethylated protein. C4 comprised the C-terminal end of T5 and the N-terminal part of T6, whereas C6 provided the connection between T6 and T7. Peptides T8 and T9, together with the C-terminal tripeptide Ile-Asp-Arg of T7 were present in the tryptic digest of CB2. CB2 must represent the C-terminal part of the protein, considering the absence of homoserine in its composition. C8 and C9 together account for the sequence of CB2 and establish the order T7-T8-T9. This places T9 in the C-terminal position of the total sequence, which is confirmed by the fact that carboxypeptidase degradation releases only one lysine residue from gene 5 protein. The presence of proline in position 85 apparently prevents the release of the penultimate alanine residue.

While our work was in progress the amino acid sequence of gene 5 protein of bacteriophage fd was independently determined (23). The finding that

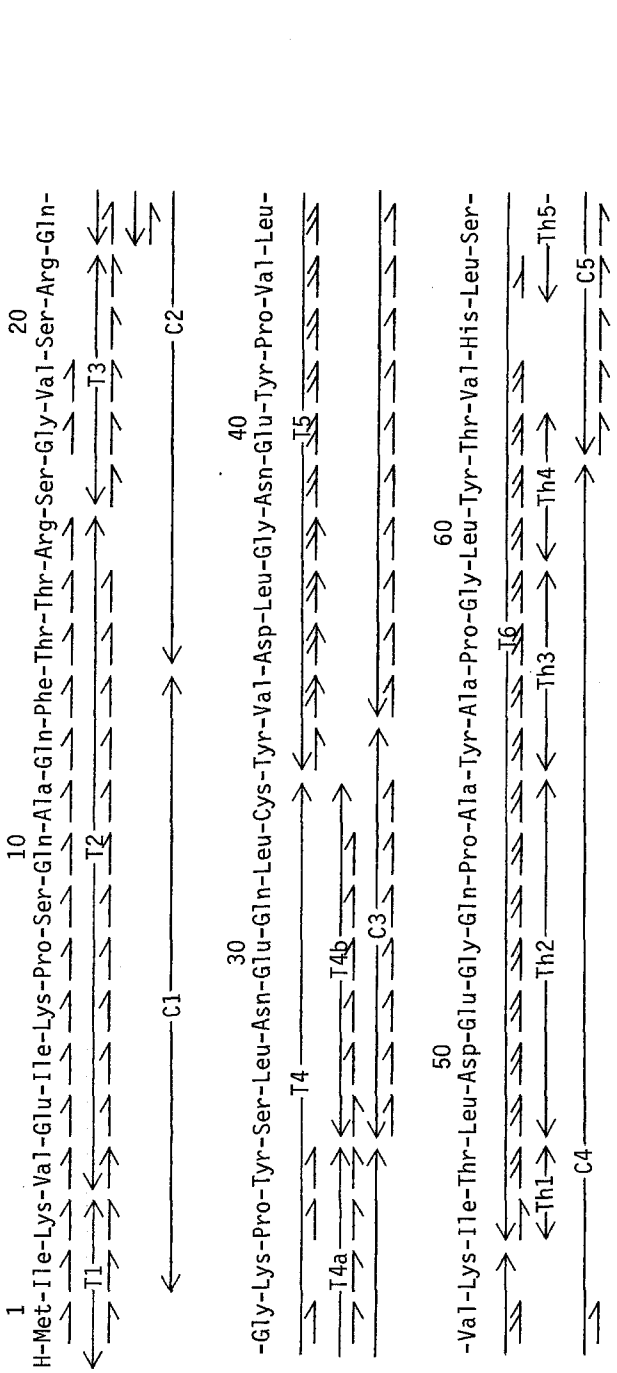

Fig. 1. Proposed sequence of gene 5 protein of bacteriophage M13. The order of the tryptic peptides (T) is determined by the chymotryptic peptides (C). Peptide T6 was further digested with thermolysin (Th). Sequences were established by Edman degradation using the automatic solid-phase (⟶) or manual (⟶) phenylthiohydantoin methods, by dansyl-Edman analysis (⟶) or by a combination of methods. The C-terminal residue of the chain was determined by carboxypeptidase B degradation (⟵).

both sequences are completely identical is of interest from a comparative point of view. Both phages belong to the group of the F-specific filamentous DNA coliphages, which also includes f1 and ZJ-2. The B-proteins (M ± 5000) of the coat of phages fd and ZJ-2 differ in two positions (24, 25). The elucidation of the sequences of more homologous proteins from these and other groups of bacteriophages might shed light on the relationships among these organisms.

ACKNOWLEDGEMENTS

We thank dr. J.G.G. Schoenmakers for stimulating discussions and cooperation during the isolation of the protein, and Miss M. Versteeg and Miss A. Hilderink for technical assistance.

REFERENCES

1. Marvin, D.A. and Hohn, B. (1969) Bacteriol. Rev. 33, 172-209.
2. Staudenbauwer, W.L. and Hofschneider, P.H. (1972) Eur. J. Biochem. 30, 403-412.
3. Lin, N.S.-C. and Pratt, D. (1972) J. Mol. Biol. 72, 37-49.
4. Tseng, B.Y. and Marvin, D.A. (1972) J. Virol. 10, 384-391.
5. Pratt, D. and Erdahl, W.S. (1968) J. Mol. Biol. 37, 181-200.
6. Salstrom, J.S. and Pratt, D. (1971) J. Mol. Biol. 61, 489-501.
7. Oey, J.L. and Knippers, R. (1972) J. Mol. Biol. 68, 125-138.
8. Staudenbauwer, W.L. and Hofschneider, P.H. (1973) Eur. J. Biochem. 34, 569-576.
9. Alberts, B., Frey, L. and Delius, H. (1972) J. Mol. Biol. 68, 139-152.
10. Pratt, D., Laws, P. and Griffith, J. (1974) J. Mol. Biol. 82, 425-439.
11. Raftery, M.A. and Cole, R.D. (1966) J. Biol. Chem. 241, 3457-3461.
12. Gross, E. (1967) Methods Enzymol. 11, 238-254.
13. Van der Ouderaa, F.J., de Jong, W.W. and Bloemendal, H. (1973) Eur. J. Biochem. 39, 207-222.
14. Brown, J.R. and Hartley, B.S. (1963) Biochem. J. 89, 59p-60p.
15. Niall, H.D. and Potts, J.T. in Peptides, Chemistry and Biochemistry (Weinstein, B. and Lande, S. eds.) p. 215-241, Dekker, New York (1970).
16. Braunitzer, G., Schrank, B. and Ruhfus, A. (1970) Hoppe Seyler's Z. Physiol. Chem. 351, 1589-1590.
17. Laursen, R.A. (1971) Eur. J. Biochem. 20, 89-102.
18. Laursen, R.A., Horn, M.J. and Bonner, A.G. (1972) FEBS Letters 21, 67-70.
19. Wachter, E., Machleidt, W., Hofner, H. and Otto, J. (1973) FEBS Letters 35, 97-102.
20. Gray, W.R. and Smith, J.F. (1970) Anal. Biochem. 33, 36-42.
21. Woods, K.R. and Wang, K.T. (1967) Biochim. Biophys. Acta 133, 369-370.
22. Offord, R.E. (1966) Nature (Lond.) 211, 591-593.
23. Nakashima, Y., Dunker, A.K., Marvin, D.A. and Konigsberg, W. (1974) FEBS Letters 40, 290-292.
24. Asbeck, F., Beyreuther, K., Köhler, H., von Wettstein, G. and Braunitzer, G. (1969) Hoppe-Seyler's Z. Physiol. Chem. 350, 1047-1066.
25. Snell, D.T. and Offord, R.E. (1972) Biochem. J. 127, 167-178.