IDENTIFICATION OF TWO NEW CAPSID PROTEINS IN BACTERIOPHAGE M13

Guus F. M. SIMONS, Ruud N. H. KONINGS and John G. G. SCHOENMAKERS
Laboratory of Molecular Biology, University of Nijmegen, Nijmegen, The Netherlands

Received 29 June 1979

1. Introduction

M13 is a male-specific filamentous coliphage, closely related to the phages fd and f1 (reviewed [1,2]). The virion is composed of a circular single-stranded DNA (mol. wt 2×10^6) encapsulated in a tubular protein coat of about 850 nm long and 6 nm wide. Up to now it has generally been assumed that the virion coat is composed of two proteins only. The major coat protein (mol. wt 5200) encoded by gene VIII is present in about 2400 copies and accounts for virtually all of the protein mass of the virion coat. The minor coat protein (mol. wt 42 600) which is encoded by gene III is present in 4–5 copies only which are located at one tip of the filament.

Since there have been reports [3–5] which suggest that the filamentous phages may contain an additional protein component, we have investigated the protein constituents of M13 virions in more detail. Here we demonstrate that the M13 virion contains besides the gene III and gene VIII encoded products two additional capsid proteins. These proteins, designated C and D protein, have app. mol. wt 3500 and 11 500, respectively. We further could demonstrate that the smallest structural protein is not the product of one of the eight well-characterized M13 genes but is encoded by the hitherto unidentified gene IX [6,7].

2. Materials and methods

2.1. *Materials*

All materials were of reagent grade purchased from sources given in [8]. The [^{3}H]amino acids arginine, histidine, lysine and proline were from New England Nuclear. All other radioactive amino acids were from Radiochemical Center, Amersham. In all cases the radioactive compounds with the highest specific activity were used.

2.2. *Growth and purification of radioactive phages*

To prepare ^{3}H-labeled phages, *Escherichia coli* K38 was grown in M9 minimal medium [9] supplemented with glucose (20 mM) and the common 20 amino acids (1 mM) minus the amino acid used for labeling. At 1×10^8 cells/ml the culture was infected with M13 phage (multiplicity of infection of 20) and at 15, 30, 45 and 60 min after infection 2.5 μCi/ml of [^{3}H]amino acid was added as in [10]. After 90 min infection the cells were spun down and the virus was precipitated from the supernatant as in [11]. The recovered phages were resuspended in 15 ml 0.01 M Tris–HCl (pH 7.6), 1 mM EDTA, 0.1% Sarkosyl and kept at room temperature for 30 min. After reprecipitation of the phages with 5% polyethylene glycol 6000 and 0.5 M NaCl for 2 h at 4°C, the precipitate was spun down and resuspended in 9 ml STE buffer (0.1 M NaCl, 0.01 M Tris–HCl (pH 7.6), 1 mM EDTA). The phages were further purified by CsCl density gradient centrifugation for 20 h at 15°C at 42 000 rev./min in an IEC-B60 rotor. The fractions of the gradient containing the radioactive phages were collected, dialysed against STE buffer and the phages were pelleted for 4 h at 360 000 $\times$ *g* and finally resuspended in a small volume of STE buffer.

Labeling with a [^{14}C]amino acid mixture was in 25 ml cultures of *E. coli* K38 grown in M9 minimal medium plus 20 mM glucose but no amino acids added. At 1×10^8 cells, M13 phages were added (2×10^9/ml) and at 30, 60, 90 and 120 min after

infection quarter parts of the 150 μCi ^{14}C-labeled algal hydrolysate supplemented with 10 μCi each of [^{14}C]methionine, histidine, asparagine and glutamine [12], respectively, were added to the culture. After a 5 h infection period the cells were spun down and further used for the isolation of ^{14}C-labeled gene V protein. The phages were recovered from the supernatant and further purified as described above.

2.3. *Polyacrylamide gel electrophoresis and recovery of phage proteins*

Fractionation of M13 capsid proteins was performed on polyacrylamide slab gels (14 × 12 × 0.12 cm) containing 15% acrylamide, 0.4% bisacrylamide, 8 M urea and 0.1% SDS in 0.5 M Tris–HCl (pH 8.9). Samples of M13 phages were diluted with cracking buffer [8] and heated for 5 min at 100°C prior to electrophoresis which was run for 5 h at 200 V. The running buffer was 6 M urea, 77 mM glycine, 10 mM Tris–HCl (pH 8.9), 0.1% SDS. After electrophoresis, the gels containing ^{3}H-labeled samples were fixed in 15% trichloroacetic acid and the radioactive bands were vizualized by means of fluorography [13]. Gels containing ^{14}C-labeled material were fixed and stained with Coomassie brilliant blue (0.1%), destained, dried and autoradiographed [8]. Electrophoretic elution of radioactive proteins from gel-segments was performed in 2 ml 25 mM Tris, 192 mM glycine, 0.1% SDS, 0.5% 2-mercaptoethanol in the presence of 200 μg bovine serum albumin using an ISCO sample concentrator. After dialysis of the eluate, the proteins were precipitated, washed and lyophilized according to [12].

The isolation of ^{14}C-labeled gene V protein was performed essentially as in [14]. The radioactive protein was further purified by electrophoresis on polyacrylamide gels as described above.

3. Results and discussion

Electrophoretic analysis of the proteins present in purified M13 phages reveals, as expected, the presence of the major and minor capsid protein after staining the gel with Coomassie brilliant blue R (fig.1A). The staining intensities reflect their occurrence in the M13 virion. In contrast, when the

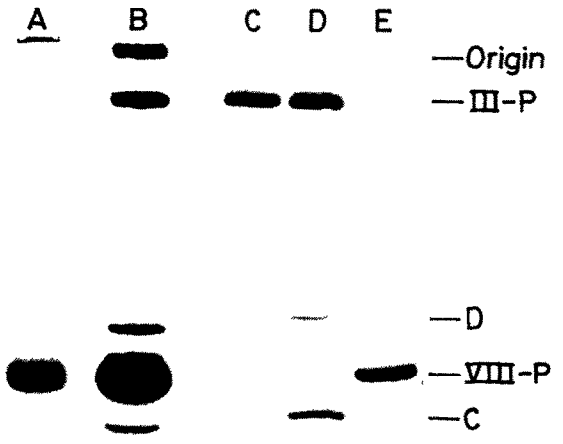

Fig.1. SDS–polyacrylamide gels of the capsid proteins in purified M13 phages. Conditions for electrophoresis and autoradiography are described in section 2.3. (A) M13 capsid proteins after staining the gel with Coomassie blue. (B) ^{14}C-Labeled M13 capsid proteins. (C) [^{3}H]Histidine-labeled capsid proteins. (D) [^{3}H]Arginine-labeled capsid proteins. (E) Purifed [^{3}H]leucine gene VIII protein. III-P and VIII-P denotes the mature proteins of genes III and VIII, respectively.

virion proteins present in ^{14}C-labeled M13 phages were electrophoretically analysed, two additional polypeptides were clearly visible on the autoradiogram (fig.1B). These polypeptides, designated C and D protein, have app. mol. wt 3500 and 11 500, respectively. Repeated CsCl density gradient and/or sucrose gradient centrifugation of the ^{14}C-labeled M13 phages did not change the electrophoretic pattern of these proteins. Hence, we assume that these polypeptides are not impurities contributed by the host cell but represent structural proteins of the virion.

To characterize these structural components and to determine their genetic origin we attempted to isolate these proteins on a preparative scale by conventional ion-exchange and exclusion chromatographic procedures. These attempts, however, were unsuccessful mainly because of the great abundancy of gene VIII protein as compared to C and D protein, their small differences in molecular weights and the tendency of the phage structural proteins to aggregate even under strongly denaturing conditions. For this reason we attempted to characterize these proteins on the basis of their capability to incorporate several

Table 1
Incorporation of [^{3}H]amino acids into phage M13 capsid proteins

Amino acid	III-P	VIII-P	C	D
[^{3}H]Histidine	+	–	–	–
[^{3}H]Arginine	+	–	+	+
[^{3}H]Proline	+	+	–	+
[^{3}H]Lysine	+	+	–	+
[^{3}H]Tryptophan	+	+	+	+
[^{3}H]Tyrosine	+	+	+	+
[^{3}H]Leucine	+	+	+	+
[^{35}S]Methionine	+	+	+	+
[^{3}H]Valine	+	+	+	+

(+) indicates incorporation; (–) no incorporation detected by radioautography

representative [^{3}H]amino acids. Since we recently have sequenced the entire M13 DNA genome [15], the established nucleotide sequence of the eight M13 genes could successfully be used as a guide for these labeling experiments. An example of such labeling studies is shown in fig.1 and the results obtained with several amino acids is summarized in table 1.

To verify that the radioactive labeling of the proteins was solely due to incorporation of the [^{3}H]amino acid originally added to the growing culture and not by a metabolised product, the radioactive protein hydrolysates were analysed by thin-layer chromatography [24]. In each case a single spot was observed on the autoradiogram which could not be discriminated from the radioactive precursor used to label the virion proteins (data not shown).

Given these data, several possibilities for the occurrence of these proteins can be excluded. It is well known that the gene III- and gene VIII-encoded capsid proteins are synthesized in the infected cell in a precursor form [16,17]. Prior to packaging of these proteins into the mature virions 18 and 23 amino acid residues from their N-terminal end are cleaved off at the host cell membrane. From the nucleotide sequence we know that both precursor parts contain proline and lysine but they lack arginine, while histidine is only present in the precursor part of gene III protein [6,7,15]. These data, therefore, completely rule out the possibility that C protein is one of these N-terminal cleavage products and that D protein is in some way a larger N-terminal part of the gene III protein precursor. Gene VIII protein itself lacks histidine but also arginine (see fig.1) [18]. Since C and D protein are labeled with [^{3}H]-arginine (fig.1), it is also excluded that C protein is a degradation product of gene VIII protein and that D protein, for instance, represents a dimer of the major capsid protein.

In order to establish a gene–product relationship for these two new capsid proteins, we inspected the entire nucleotide sequence of the M13 genome in search for possible coding regions, irrespective of the reading frame, which could agree with our amino acid incorporation data and the apparent molecular weights of these proteins.

As far as C protein is concerned two regions were found which approximated the above requirements, namely gene VII (its encoded product is 33 amino acids long and lacks His, Pro and Lys) and a region of 94 nucleotides between genes VII and VIII, the coding function of which has never been established by genetic data so far and which therefore is designated the hypothetical gene IX [6,7]. In contrast to gene VII, the hypothetical gene IX contains a TGG (tryptophan) triplet which led us to suppose that C protein is most probably the product of gene IX (table 1). With respect to D protein an unambiguous gene–product relationship could not be proposed although a good candidate would be gene VI protein. The latter protein does not contain histidine and has mol. wt 12 260 as deduced from the nucleotide sequence data [7,20].

In order to establish unambiguously the genetic origin of these proteins we have performed amino acid analysis of uniformly ^{14}C-labeled C and D protein, and of uniformly ^{14}C-labeled gene VIII protein and gene V protein. As the amino acid sequence of the latter proteins have been determined [18,19], the number of each [^{14}C]amino acid residue could be calculated using these proteins as external standards. The data in table 2 demonstrate that with each standard an identical amino acid composition of C protein was found. It is also evident that C protein does not contain the amino acid residues Pro, Lys, His and Asp. Furthermore, the number of residues estimated for the other amino acids correspond with the numbers calculated for the protein encoded by gene IX. Hence, we feel confident to conclude that gene IX is a real M13 gene which codes for a new

Table 2
Amino acid composition of C-protein

Amino acid	Number of residues for			
	C protein a	b	IX-p[g] c	VII-p[h] c
Asp	0.2	0.2	0	2
Thr	2.7	3.8	3	1
Ser	5.3	4.9	6	1
Glu	1.5	1.4	1	5
Pro	0.1	0.1	0	0
Gly	2.3	2.1	2	3
Ala[f]	3.7	4.7	1	4
Val	2.7	2.7	3	3
Met	2.0	1.8	2	2
Ile[d]	2.1	2.4	1	5
Leu[d]	3.0	2.9	4	2
Tyr	1.7	1.7	2	1
Phe	3.1	2.9	3	2
His	–[e]	0.3	0	0
Lys	0.3	0.3	0	0
Arg	–[e]	1.9	2	1
Cys	n.d.	0.6	1	1

[a,b] These columns represent the amino acid composition estimated using uniformly ^{14}C-labeled gene VIII protein and gene V protein, respectively
[c] The calculated number of residues as deduced from the nucleotide sequence of gene IX and VII
[d] Note that the sum of Ile + Leu equals the sum of these residues in gene IX
[e] These values cannot be estimated as these residues are absent in gene VIII protein
[f] Values are less reliable due to buffer change
[g] Mature protein of gene IX
[h] Mature protein of gene VII

Amino acid analysis was performed on radioactive protein samples which were hydrolyzed in 6 N HCl in sealed evaculated tubes at 110°C for 22 h and analysed on a Chromaspek amino analyser. Fractions of 30 s were collected and the radioactivity determined

minor coat protein, i.e., C protein. The amino acid composition of D protein (data not shown) varied slightly from one batch to the other, indicating that still capsid protein aggregates are present among D protein. That D protein contains nevertheless a substantial amount of gene VI-encoded products is inferred from our analysis data which consistently showed the absence of histidine, low numbers of arginine and methionine and relatively high numbers of isoleucine and particularly leucine (cf. [20]). Recent N-terminal sequence analysis, which will be detailed elsewhere, have substantiated this conclusion.

As evidence now has been provided that the M13 virion is composed of four instead of two capsid proteins, further studies are needed to determine the orientation of C and D protein within the tubular phage coat in order to understand their specific role in the process of phage assembly at the host cell membrane. One of the minor capsid proteins, namely gene III protein, is located on one tip of the filament [4,5,21] and is required for adsorption of the phage to the F-pilus of the host bacterium [22]. Although any evidence is lacking it is attractive to postulate that one of the other minor components is located on the opposite tip of the filament. In a model on phage assembly at the host cell membrane [23] it was suggested that gene III protein forms the end of the virion that is extruded last. It is also known that under nonpermissive conditions cells infected with gene III or gene VI amber mutants produce an increased number of multiple genome-length defective particles [22]. This suggests that packaging of the viral DNA into the viral protein tube is primarily determined by the length of the DNA but that in this packaging process both gene III and gene VI protein are functioning as cut-off agents for further tubular growth. Our data are still ambiguous but if gene VI codes for D protein, then it is very likely that both gene III protein and D protein are located on the same tip of the filament. Further investigations with recombinant DNA-M13 phages of variable lengths, currently in progress, are needed to elucidate the location of these capsid proteins.

Acknowledgements

We thank Miss Titia Peters for her expert technical assistance in this work. We also acknowledge the collaboration of Feitse Jan Duisterwinkel and Barend Kraal, who performed the amino acid analysis cited. This work was supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO).

References

[1] Denhardt, D. T. (1975) CRC Crit. Rev. Microbiol. 4, 161–223.

[2] Ray, D. S. (1977) in: Comprehensive Virology (Fraenkel-Conrat, H. and Wagner, R. R. eds) vol. 7, pp. 105–178, Plenum, New York.
[3] Henry, T. J. and Pratt, D. (1969) Proc. Natl. Acad. Sci. USA 62, 800–807.
[4] Goldsmith, M. E. and Konigsberg, W. H. (1977) Biochemistry 16, 2686–2693.
[5] Woolford, J. L., Steinman, H. M. and Webster, R. E. (1977) Biochemistry 16, 2694–2700.
[6] Hulsebos, T. and Schoenmakers, J. G. G. (1978) Nucleic Acids Res. 5, 4677–4698.
[7] Beck, E., Sommer, R., Auerswald, E. A., Kurz, Ch., Zink, B., Osterburg, G., Schaller, H., Sugimoto, K., Sugisaki, H., Okamoto, T. and Takanami, M. (1978) Nucleic Acids Res. 5, 4495–4503.
[8] Konings, R. N. H. (1979) Methods Enzymol. 65/I, in press.
[9] Miller, J. H. (1972) in: Experiments in Molecular Genetics, p. 431, Cold Spring Harbor Laboratory, NY.
[10] Makino, S., Woolford, J. L., Tanford, C. and Webster, R. E. (1975) J. Biol. Chem. 250, 4327–4332.
[11] Yamamoto, K. R., Alberts, B. M., Benzinger, R., Lawhorne, L. and Treiber, G. (1970) Virology 40, 734–744.
[12] Shaw, D. C., Walker, J. E., Northrop, F. D., Barrell, B., Godson, G. N. and Fiddes, J. C. (1978) Nature 272, 510–515.
[13] Bonner, W. M. and Laskey, R. A. (1974) Eur. J. Biochem. 46, 83–88.
[14] Garssen, G. J., Hilbers, C. W., Schoenmakers, J. G. G. and Van Boom, J. H. (1977) Eur. J. Biochem. 81, 453–463.
[15] Van Wezenbeek, P., Hulsebos, T. and Schoenmakers, J. G. G. (1979) submitted.
[16] Konings, R. N. H., Hulsebos, T. and Van den Hondel, C. A. (1975) J. Virol. 15, 570–584.
[17] Chang, C. N., Blobel, G. and Model, P. (1978) Proc. Natl. Acad. Sci. USA 75, 361–365.
[18] Asbeck, F., Beyreuther, K., Köhler, H., Von Wettstein, G. and Braunitzer, G. (1969) Hoppe Seyler's Z. Physiol. Chemie 350, 1047–1066.
[19] Cuypers, T., Van der Ouderaa, F. J. and De Jong, W. W. (1974) Biochem. Biophys. Res. Commun. 59, 557–563.
[20] Van Wezenbeek, P. and Schoenmakers, J. G. G. (1979) Nucleic Acids Res. 6, 2799–2818.
[21] Marco, R. (1975) Virology 68, 280–282.
[22] Pratt, D., Tzagoloff, H. and Beaudoin, J. (1969) Virology 39, 42–53.
[23] Chang, C. N., Model, P. and Blobel, G. (1979) Proc. Natl. Acad. Sci. USA 76, 1251–1255.
[24] Pataki, G. (1966) Dünnschichtchromatographie in der Aminosaüre- und Peptidchemie, 1st edn pp. 91–98, De Gruyter, Berlin.