Biochemistry (Moscow), Vol. 67, No. 12, 2002, pp. 1366-1370. Translated from Biokhimiya, Vol. 67, No. 12, 2002, pp. 1655-1660.

Engineering of Bacteriophage T4 Tail Sheath Protein[1]

A. V. Efimov, L. P. Kurochkina*, and V. V. Mesyanzhinov

Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, ul. Miklukho-Maklaya 16/10, Moscow, 117997 Russia; fax: (095) 336-6022; E-mail: lpk@mail.ibch.ru

Received May 22, 2002
Revision received June 21, 2002

Abstract—Gene product 18 (gp18, 659 amino acids) forms bacteriophage T4 contractile tail sheath. Recombinant protein assembles into different length polysheaths during expression in the cell, which complicates the preparation of protein crystals for its spatial structure determination. To design soluble monomeric gp18 mutants unable to form polysheaths and useful for crystallization, we have used *Bal*31 nuclease for generation deletions inside gene *18* encoding the Ile507-Gly530 region. Small deletions in the region of Ile507-Ile522 do not affect the protein assembly into polysheaths. Protein synthesis termination occurs because of reading frame failure in the location of deletions. Some fragments of gp18 containing short pseudo-accidental sequence in the C-terminal, while being soluble, have lost the ability for polysheath assembly. For the first time we succeeded in obtaining crystals of a soluble gp18 fragment containing 510 amino acids which, according to trypsin resistance, is similar to native protein monomer.

Key words: bacteriophage T4, contractile tail sheath, gene product 18, polymerization, protein folding

Bacteriophage T4 tail sheath, essential for viral DNA injection into the cell, is a simple metastable contractile system. Phage sheath consists of 144 subunits of gene product 18 packing into 24 annuli [1]. Gp18 subunits are arranged around the tail tube (gp19) which is the tunnel for viral DNA passage during infection [2]. In the absence of the tube gp18 assembles into polysheaths of different length, and their structure is probably similar to that in the contracted state [3].

The interaction of phage T4 with the receptors on the cell surface initiates the structural changes of base plate located on the tail-distal end that results in subsequent irreversible sheath contraction. During the contraction sheath length decreases from 98 to 36 nm and its outer diameter increases from 21 to 27 nm [4]. Moody [5] first observed that during contraction transformation of the surface lattice occurs: a screw axis translation decreases from 4.1 to 1.5 nm, and rotation between subunits increases from 17 to 32°. Using electron microscopy data, Amos and Klug [6], and later Smith et al. [7] developed the image reconstruction of extended and contracted sheaths. The comparison of three-dimensional models thus obtained has shown that sheath contraction is accompanied by changes in the orientation of subunits as well as small changes in their shape. The protein secondary structure may change too. According to circular dichroism data, the α-helices in the protein decrease by about 14% and the β-structure increases by about the same [8, 9]. Comparative analysis of monomer and different structural forms of gp18 was studied by limited proteolysis and chemical modification methods [10, 11]. It was established that the protein is resistant to proteolysis in the assembled form of polysheaths. Hydrolysis of monomer results in a trypsin-resistant fragment Ala82-Lys316 with a molecular mass of 27 kD that according to immunoelectron microscopy data constitutes the protruding part of the sheath surface. The mechanism and details of sheath contraction have not yet been clarified.

Earlier we constructed an expression system and developed a method for purification of recombinant gp18 [12]. We also constructed a set of gp18 fragments with deletions in the N- and C-terminal regions of the molecule. Deletions of the N-terminal region resulted in formation of insoluble aggregates in *E. coli* cells [12]. Deletions of the C-terminal region did not disturb protein assembly into polysheaths, and moreover, in addition to native polysheaths, filaments with thinner diameter called "noncontracted polysheaths" (NCP) were observed [13]. The structure of the NCP filaments significantly differs from both extended and contracted polysheaths, and probably corresponds to transitional helices

[1] The authors dedicate this paper to the memory of Boris Fedorovich Poglazov.

* To whom correspondence should be addressed.

0006-2979/02/6712-1366$27.00 ©2002 MAIK "Nauka/Interperiodica"

[5]. In studying the properties of gp18 deletion fragments we have identified a few important regions of the protein molecule: the Phe373-Arg411 region might be responsible for polymerization of subunits, and the Ile507-Gly530 region is probably important for conformational transitions of the protein during the sheath contraction [13].

In the present work, to design gp18 soluble mutants lacking the ability for polysheath assembly and suitable for crystallization, we have generated deletions inside the Ile507-Gly530 region of the polypeptide chain. Short deletions inside this region of the molecule did not disturb the ability of the protein to assembly into polysheaths of two types. Reading frame failure during deletions resulted in the appearance of comparatively short pseudo-accidental sequences in the C-terminal. We have observed that some of the gp18 mutants with reading frame failure did not form polysheaths. These proteins are soluble and have structure similar to gp18 monomer. These gp18 mutants are potential candidates for crystallization and subsequent spatial structure determination of protein fragments.

MATERIALS AND METHODS

Bacterial strains. Bacterial strain *E. coli* Top10 from Invitrogene (USA) was used for production of plasmid DNA. Strains *E. coli* BL21(DE3) and JM109(DE3) from Novagen (USA) were used for expression of genes cloned into plasmids under control of the phage T7 promoter.

Design of deletion fragments. Deletion mutants were designed on the basis of pGT-18 expression vector [12]. Deletions in gene *18* were generated by controlled digestion with nuclease *Bal*31 [14]. Plasmid DNA was restricted at one unique restriction site (*Esp*I or *Mlu*I) and linear DNA with protruding ends was used as a substrate for subsequent digestion with *Bal*31. Incubation was carried out at 25°C in buffer for *Bal*31. Aliquots (20 μl) were withdrawn from reaction mixture (100 μl) at 1 min intervals. The reaction was stopped by addition of 10 μl mixture of 0.1 M EGTA and 0.25 M EDTA. DNA 5′-ends were repaired with Klenow's fragment DNA polymerase I and ligated. Deletion length in gene *18* was determined by DNA sequencing.

Gene expression in *E. coli* BL21(DE3) cells. Gene *18* cloned under phage T7 promoter was expressed as described earlier [12]. To induce expression, isopropyl-1-thio-β-D-galactoside (IPTG) was added to the final concentration of 1 mM, and incubation was continued overnight at 25°C. The cells were centrifuged at 2000*g* for 15 min.

Protein solubility analysis. Cell pellet after protein expression was suspended in TE-buffer, pH 8, containing lysozyme from Serva (Germany) (0.3 mg/ml), and sonicated on ice for 2-3 min (15 sec pulses with 1 min pauses). Cellular debris was removed by centrifugation at 12,000*g* for 10 min. Supernatant and pellet were analyzed by SDS-PAGE.

Electrophoretic analysis of recombinant proteins was carried out in 10 or 12.5% polyacrylamide gels [15]. A standard set of LMW proteins from Pharmacia and Biotech (Sweden) containing α-lactalbumin (14.4 kD), soybean trypsin inhibitor (20.1 kD), carboanhydrase (30 kD), ovalbumin (43 kD), BSA (67 kD), and phosphorylase *b* (94 kD) was used as the molecular weight marker.

Limited proteolysis of recombinant proteins with trypsin. Trypsin was added into the protein solution (1.5 mg/ml) in ratio 1 : 50 and incubated at 23°C. Aliquots (30 μl) were withdrawn from the reaction mixture in 5, 10, and 30 min and mixed with 10 μl of 4× sample buffer and boiled for 3 min. The samples were analyzed by electrophoresis in 12.5% gels.

Electron microscopy. The samples on the grids were contrasted with 0.75% uranyl formate or 1% uranyl acetate. Microphotographs were obtained using a Hitachi 7000 electron microscope (Japan).

RESULTS

To design soluble gp18 mutants suitable for crystallization, we generated deletions in the gene region encoding the Ile507-Gly530 sequence. For this pGT-18 plasmid vector containing gene *18* [12] was linearized at one of the sites *Esp*I (1535 bp) or *Mlu*I (1730 bp) of the 3′-end (Fig. 1a). Then two-directed deletions were generated by digestion with *Bal*31 nuclease. The plasmids that expressed recombinant proteins with molecular weights approximating full-length gp18 were selected after termination of the reaction and subsequent ligation. Deletion length was determined by DNA sequencing of gene *18*.

The mutants obtained were divided into three groups. The first contains proteins with deletions inside the molecule that did not disturb the reading frame and did not have internal termination codons. The proteins with pseudo-accidental amino acid sequence resulted from reading frame failure in deletion location belong to the second group. The third group contains truncated gp18 fragments without change in amino acid sequence (Fig. 1b).

Group I mutants. We selected 9 clones producing proteins with deletion length varying from 8 to 42 amino acids (Fig. 1b). Protein solubility analysis showed that some of them with deletions inside the investigated polypeptide chain region are soluble (mutants E13, E37, E53, and E69), while the other form inclusion bodies (E34, E40, E51, E56, and E98). As an example, the electrophoretic picture of cell lysates is presented in Fig. 2; soluble proteins are in supernatant, insoluble in pellet fraction. Electron microscopy showed that soluble fragments form two types of polymeric structures (Fig. 3a) described by us earlier [13]: NCP and polysheaths. Thus, deletions inside the Ile507-Ile522 region do not affect the

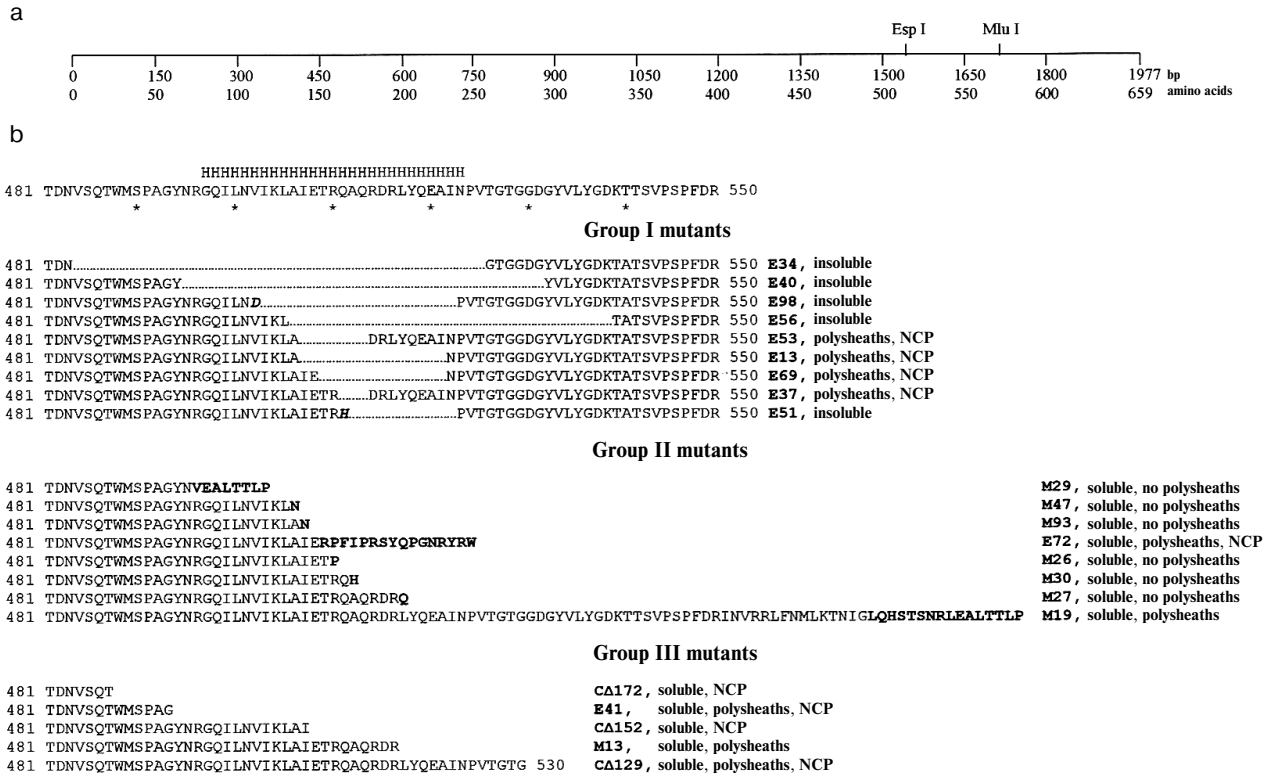

Fig. 1. a) Location in gene *18* of unique restriction sites used for deletion production with *Bal*31 nuclease. b) Primary structure of gp18 mutants in deletion region. Dotted line, deleted amino acids; italic, amino acid substitutions; bold, pseudo-accidental sequences; H, α-helix.

ability of the protein for polysheath assembly. Deletion shift to the C-terminal of the molecule probably disturbs gp18 folding. All insoluble proteins have deletions including the Asn523 residue.

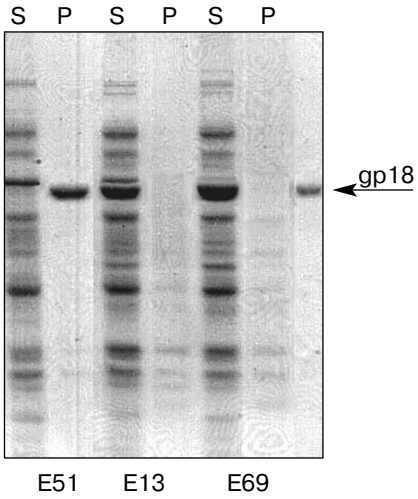

Fig. 2. Extracts of *E. coli* cells expressing different gp18 mutants after induction with IPTG. Samples were analyzed in 10% SDS-PAGE. S, supernatant; P, pellet. Full-length gp18 has been used as a marker.

Group II mutants. We studied the properties of 8 mutants containing pseudo-accidental amino acid sequence with the length from 1 to 16 amino acids in the C-terminal because of reading frame failure upon deletion (Fig. 1b). All proteins of this group were soluble even at high concentration. Electron microscopy showed that gp18 fragments containing substitution of the last C-terminal amino acid residue do not form polysheaths. For example, earlier designed gp18-CΔ152 with Ile507 residue in the C-terminal forms regular polymeric structures, NCP [13], while deletion fragment M93 with substitution of Ile507 to Asn has completely lost the ability to polymerize. Fragment M26 with substitution of Arg510 to Pro is soluble and does not form polysheaths too. However, this property is not a rule, and in some cases gp18 mutants with longer pseudo-accidental amino acid sequences, for example M19 and E72 form polysheaths of two types (Fig. 3b).

Group III mutants. Deletion variants of gp18 truncated from the C-terminal and not containing pseudo-accidental sequences belong to the third group. E41 and M13 fragments, containing 493 and 516 amino acids, respectively, are soluble and form polysheaths of two types like earlier designed CΔ152 and CΔ129 deletion fragments. Electron microphotographs of polysheaths formed by E41 protein are presented in Fig. 3c.

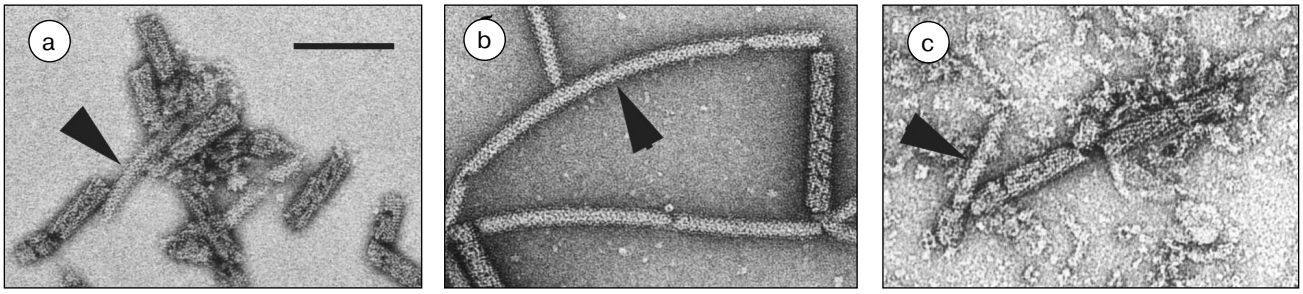

Fig. 3. Electron microphotographs of polysheaths formed by gp18 mutants: E37 (a), E72 (b), and E41 (c). Scale bar represents 100 nm. Arrow shows NCP.

To compare the structure of fragments and native gp18, we studied the resistance of recombinant proteins to trypsin. Earlier it has been shown that full-length gp18 is resistant to proteolysis in the assembled form of polysheaths, but is cleaved by proteases in the monomeric form [11]. The electrophoretic picture of proteolytic products for M26 protein, which does not form polymeric structure, for M19 mutant, which forms polysheaths, and for native gp18 in polysheath state is shown in Fig. 4. According to the figure, gp18 showed no significant degradation during incubation for 150 min, while M19 protein is completely degraded in this time. The M26 mutant is more unstable and is cleaved 5 min later after enzyme addition. The proteolysis of both mutant proteins results in formation of the same intermediates, and the final digestion product (~31 kD) is similar to that of native gp18 (Fig. 4). These data confirm that mutants M19 and M26 and native gp18 have apparently similar structure. However, the association of subunits in polysheaths formed by M19 mutant is weaker than in native polysheaths, and the equilibrium between monomer and polysheaths is shifted into monomer formation.

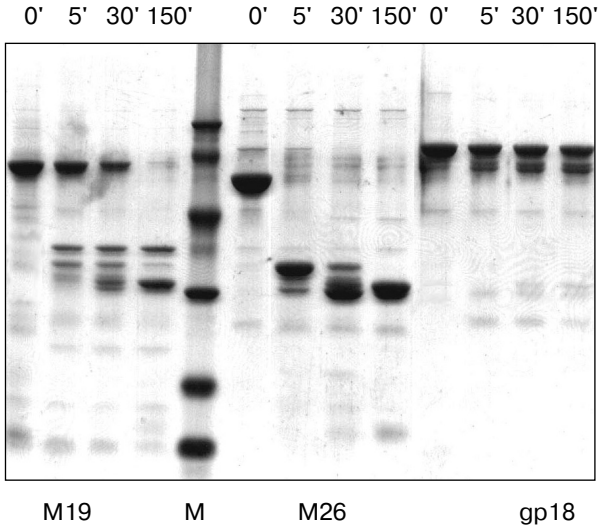

Fig. 4. Analysis of limited trypsinolysis products of M19 and M26 mutants and native gp18 in 12.5% SDS-PAGE. Hydrolysis time (min) is shown above the lanes; M, standard proteins.

DISCUSSION

As previous study has shown, the gp18 Ile507-Gly530 sequence is probably important for conformational reorganization of the protein during sheath contraction. In this work we generated short deletions directly in this region of the molecule and studied the ability of the protein for polysheath assembly. The results directly show that besides the assistance in gp18 conformational transition during sheath contraction, the Ile507-Gly530 region is also necessary for directed assembly of the protein into polysheaths.

The gp18 secondary structure prediction using known algorithms [16] showed that the Gly497-Ile522 region of the polypeptide chain might be responsible for α-helix formation (Fig. 1b). Study of properties of mutants obtained shows that comparatively longitude deletions not including Asn523 do not affect the ability of the protein to form polysheath. However, deletion of even short polypeptide chain fragments including Asn523 disturbs protein folding and results in its nonspecific aggregation in the cell. For example, the E51 mutant with deletion of 12 amino acids including Asn523 forms inclusion bodies. Apparently, the Asn523 residue located in the terminal of a predicted α-helix forms C-cap [17].

Group II mutants and particularly those which lost the ability for polysheath formation because of point substitution in the C-terminal are of greatest interest for further work. These proteins are synthesized in the cell in large amounts, and they are soluble and stable. It is important that their structure, according to the proteolysis test, is similar to native gp18. These mutants, in contradistinction to full-length protein and its other deletion variants forming polysheaths, are potential candidates for crystallization and further determination of their struc-

ture by X-ray analysis. M26 mutant was produced in preparative amounts and purified. We succeeded in obtaining protein crystals with well-define facets and size of 0.1-0.2 mm that polarized light (A. V. Efimov, L. V. Zhemaeva, unpublished) which will be used for spatial structure determination.

This study was supported by the Russian Foundation for Basic Research (projects No. 00-04-48372 and No. 02-04-06854) and a grant from the Howard Hughes Medical Institute (No. 55000324).

REFERENCES

1. Arisaka, F., Nakao, T., Takahashi, H., and Ishii, S.-I. (1988) *J. Virol.*, **62**, 1186-1189.
2. King, J., and Mikolajewycz, N. (1973) *J. Mol. Biol.*, **75**, 339-358.
3. Moody, M. F. (1967) *J. Mol. Biol.*, **25**, 167-200.
4. Simon, L. D., and Anderson, T. F. (1967) *Virology*, **32**, 279-297.
5. Moody, M. F. (1973) *J. Mol. Biol.*, **80**, 613-636.
6. Amos, L. A., and Klug, A. (1975) *J. Mol. Biol.*, **99**, 51-73.
7. Smith, P. R., Aebi, U., Josephs, R., and Kessel, M. (1976) *J. Mol. Biol.*, **106**, 243-275.
8. Venyaminov, S. Yu., Rodikova, L. P., Metlina, A. L., and Poglazov, B. F. (1975) *J. Mol. Biol.*, **98**, 657-664.
9. Serysheva, I. I., Venyaminov, S. Yu., Selivanov, N. A., Rodikova, L. P., and Poglazov, B. F. (1990) *Mol. Biol. (Moscow)*, **24**, 541-547.
10. Takeda, S., Arisaka, F., Ishii, S., and Kyogoku, Y. (1990) *Biochemistry*, **29**, 5050-5056.
11. Arisaka, F., Takeda, S., Funane, K., Nishijima, N., and Ishii, S. (1990) *Biochemistry*, **29**, 5057-5062.
12. Kuznetsova, T. A., Efimov, A. V., Aijrich, L. G., Kireeva, I. Yu., Marusich, E. I., Cappuccinelli, P., Fiori, P., Rappelli, P., Kurochkina, L. P., Poglazov, B. F., and Mesyanzhinov, V. V. (1998) *Biochemistry (Moscow)*, **63**, 702-709.
13. Poglazov, B. F., Efimov, A. V., Marco, S., Carrascosa, J., Kuznetsova, T. A., Aijrich, L. G., Kurochkina, L. P., and Mesyanzhinov, V. V. (1999) *J. Struct. Biol.*, **127**, 224-230.
14. Henriquez, V., and Gennaro, M. L. (1990) *Nucleic Acids Res.*, **22**, 6735-6736.
15. Laemmli, U. K. (1970) *Nature*, **227**, 680-685.
16. Ptitsyn, O. B., and Finkelstein, A. V. (1983) *Biopolymers*, **22**, 15-25.
17. Richardson, J. S., and Richardson, D. C. (1988) *Science*, **240**, 1548-1652.