ISSN 0006-2979, Biochemistry (Moscow), 2008, Vol. 73, No. 9, pp. 995-999. © Pleiades Publishing, Ltd., 2008.
Published in Russian in Biokhimiya, 2008, Vol. 73, No. 9, pp. 1236-1242.
Originally published in Biochemistry (Moscow) On-Line Papers in Press, as Manuscript BM08-022, April 13, 2008.

Role of the C-Terminus in Folding and Oligomerization of Bacteriophage T4 Gene Product 9

L. P. Kurochkina*, A. Yu. Vishnevskiy, and V. V. Mesyanzhinov

Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, ul. Miklukho-Maklaya 16/10, 117997 Moscow, Russia; fax: (495) 336-6022; E-mail: lpk@ibch.ru

Received January 24, 2008
Revision received February 6, 2008

Abstract—Bacteriophage T4 gene product 9 (gp9) is a structural protein of baseplate that plays a key role at the beginning of the infection process. Biologically active gp9 is a trimer that consists of three domains. It is a convenient model to study folding and oligomerization mechanisms of complex multidomain proteins. The influence of deletions and mutations of several amino acid residues in the C-terminal part of molecule on protein folding, oligomerization, and functional activity has been studied. It was determined that gp9 trimerization occurs post-translationally. It was shown that Gln282 and Ile284 are essential for gp9 trimer stabilization. The disruption of hydrogen bonds formed by Gln282 with Leu203 and Thr205 of neighboring chain has effect not only on interaction between monomers within trimer but also on folding of the polypeptide chain. *Tsf* (temperature sensitive for folding) and *su* (suppressor) mutations in the C-terminal region of the polypeptide chain affecting protein folding have been found.

DOI: 10.1134/S000629790809006X

Key words: bacteriophage T4, base plate, gene product 9, protein folding, protein oligomerization, *tsf* mutation, *su* mutation

The baseplate (BP) of *Escherichia coli* bacteriophage T4 is a multiprotein molecular complex that is responsible for recognition, adsorption, and DNA injection into the host cell [1]. Gene product 9 (gp9) is a structural BP protein that connects the long tail fibers to it [2]. At the beginning of the infection process the long tail fibers are attached to receptors of the host cell, then a signal from fibers is transmitted through gp9 to the BP. Gp9 initiates its structural transition with subsequent tail sheath contraction and DNA injection into the *E. coli* cell. Interest in gp9 study is due not only to a key role of the protein in the phage infection process. Gp9 is of significant interest as a convenient experimental model for studying folding and oligomerization of complex multidomain proteins. Clarification of principles polypeptide chain folding into native spatial structure is one of the most important problems of molecular biology and has important practical value for biotechnology and medicine.

The spatial structure of gp9 (288 amino acid residues (aa)) at 2.3 Å resolution had been determined by us earlier [3]. This allows carrying out directed mutagenesis of the polypeptide chain by genetic engineering methods and studying the influence of various mutations on protein folding, oligomerization, and functional activity. The wild-type protein is a SDS-resistant trimer, so it can be revealed by electrophoresis under denaturing conditions without preliminary heating of the sample (during heating the trimer dissociates to monomers). This test can be used for defining the degree of oligomerization of mutant proteins. Finally, functional activity of mutant proteins can be estimated by *in vitro* complementation assay, considering the property of wild-type protein to incorporate into the 9^--defective particles and convert them into infectious phages.

We previously found that the C-terminal part of the polypeptide chain is responsible for gp9 trimerization [4]. Studying properties of various gp9 deletion variants, we have concluded that the protein integrates into the phage BP via the N-terminal domain [5]. This confirms the three-dimensional model of BP determined by cryo-electron microscopy and image reconstruction [6]. The gp9 C-domain connects long tail fibers and causes their mobility. We have established that the C-terminal part of the polypeptide chain is essential for proper protein folding and trimerization.

The objective of this study was to establish how various mutations and truncations within the C-terminal domain influence gp9 folding and oligomerization.

Abbreviations: aa) amino acid residues; BP) baseplate; gp9) gene product 9.

* To whom correspondence should be addressed.

MATERIALS AND METHODS

***Escherichia coli* bacterial strains.** The Top10 strain (Invitrogen, USA) was used for screening of recombinant clones and for the plasmid DNA preparation. The BL21(DE3) strain (Invitrogen) was used for expression of genes cloned into the plasmid vectors. The CR63 strain was used as the permissive host, and $B^e/1$ was used as the restrictive host for the phage T4 gene 9 amber mutant.

Cloning of gene 9 fragments. Gene 9 fragments truncated at the 3′-end were amplified by PCR using oligonucleotide primers containing point base substitutions to generate suitable cloning sites. We used 5′-GAGCCCATGGGCTCCCT-3′ (*Nco*I restriction site is underlined) as forward primer, and 5′-TTATGGATCCACCTAGATTTTCTG-3′, 5′-TTGAGGGATCCCTTATTTCTGCGTA-3′ and 5′-TTGAGGGATCCCGATTTACTGCGT-3′ (*Bam*HI restriction site is underlined) as reverse primers for amplification of gene 9 fragments coding proteins with deletions of 4, 5, and 6 C-terminal amino acid residues (CΔ4, CΔ5, and CΔ6), respectively. DNA fragments truncated at the 3′-end of gene 9 were cloned into the *Nco*I-*Bam*HI sites of pET-23d(+) vector (Novagene, USA).

Construction of plasmid vectors for expression of gp9 mutants with Gln282 substitution. Mutated genes coding gp9* with Gln282 substitution were amplified by PCR using oligonucleotide primers containing random substitutions and the point base substitutions to generate suitable cloning sites. We used oligonucleotide 5′-GAAACCCATATGTTCATTCAAGAACCA-3′ (*Nde*I restriction site is underlined) as forward primer, and 5′-ATATTTGGATCCATTATTGAGCCACCCCGATTTTnnnCGTAGCGATAAC-3′, where n is A, C, G, or T (*Bam*HI restriction site is underlined), as reverse primer. The DNA fragment was cloned into the *Nde*I-*Bam*HI restriction sites of pET-23b(+) vector (Novagene).

Expression of recombinant proteins in *E. coli* BL21(DE3). Expression of genes cloned under control of phage T7 promoter was carried out as described earlier [7]. Competent BL21(DE3) cells were transformed by the plasmids, plated on 1.5% agar containing ampicillin (100 μg/ml), and incubated at 37°C overnight. The transformants were inoculated in 2× TY medium containing ampicillin (200 μg/ml) and grown at 37°C until $A_{600} = 0.5$. Isopropyl β-D-thiogalactoside was added for induction of the expression to a final concentration of 1 mM, and the culture was incubated additionally for 3 h. The cells were pelleted by centrifuging at 3500 rpm for 20 min (Megafuge 2.0 R; Heraeus Instruments, Germany).

Immunoblotting was carried out as described earlier [8].

Preparation of suspension of phage T4 9^--particles. A phage T4 gene 9 amber mutant (amC215) plaque in the *E. coli* CR63 cells from top agar was transferred into 0.5 ml of physiological solution containing 20 mM $MgCl_2$ and incubated for 1 h at room temperature. This suspension was used for infecting 5 ml of liquid culture of *E. coli* CR63 in an exponential growth phase with $2 \cdot 10^8$ cells per ml with following incubation for 16 h at 30°C. The cells were lysed by adding 2-3 drops of chloroform. The debris was removed by centrifuging at 12,000 rpm for 10 min. The titer of infective particles was about $5 \cdot 10^{10}$ plaque forming units per ml.

Preparation of cell extracts containing recombinant proteins. A pellet from 0.5-ml cell culture was resuspended in 100 μl of lysis buffer (Na_2HPO_4, 0.6 g/liter; KH_2PO_4, 0.3 g/liter; NaCl, 0.05 g/liter; NH_4Cl, 0.1 g/liter; 20 mM $MgCl_2$, pH 7.5) containing 1 mM phenylmethylsulfonyl fluoride and sonicated for 2-3 min (15-sec pulses with 15-30-sec pauses) using a Virsonic 100 disintegrator (Virtis, USA). Cell debris was removed by centrifugation at 12,000 rpm (Eppendorf, Germany) for 2 min. Concentration of recombinant protein in the supernatant was estimated by electrophoresis. Freshly prepared cell extracts were used in the experiments.

Preparation of phage T4 9^--particles. Cell extract containing phage T4 9^--particles was prepared as described earlier [9] using gene 9 amber mutant C215. The *E. coli* $B^e/1$ cells were grown in 50 ml 2× TY medium containing 5 mM $MgCl_2$ at 37°C to a density of $2 \cdot 10^8$ cells per ml and twice (with 10-min interval) infected by the amC215 with multiplicity of 5. After the second infection, the concentration of $MgCl_2$ was increased to 20 mM and the culture was additionally incubated for 40 min. The cells were pelleted at 4000 rpm for 10 min and resuspended in 2 ml of lysis buffer containing 0.05 mg/ml DNase I. Aliquots were stored at −70°C.

Complementation assay *in vitro*. A 20-μl sample of phage T4 9^--particle extract was mixed with a 20-μl sample of cell lysate containing recombinant protein. The mixture was incubated for 2 h at 30°C. During incubation, 5-μl aliquots were 1000-fold diluted in physiological solution and titrated using *E. coli* CR63 cells.

Protein electrophoresis in polyacrylamide gel. Electrophoresis was carried out as described earlier [10]. Gels were stained with 0.3% Coomassie brilliant blue R (Sigma, USA) in acetic acid–ethanol–water (1 : 3 : 6 v/v) solution and destained by 50% ethanol and 7% acetic acid solution. Samples were prepared by mixing of 100 μl cell culture after expression and 30 μl of 4× sample buffer (0.25 M Tris-HCl, pH 6.8, 8% SDS, 40% (v/v) glycerol, 20% (v/v) 2-mercaptoethanol, 0.02% (w/v) bromophenol blue) and heated at 95°C for 5 min. Samples (10-20 μl) were loaded on the gel.

Buffer solutions without SDS and 2-mercaptoethanol were used for protein electrophoresis under nondenaturing conditions.

RESULTS AND DISCUSSION

Co-expression of full-length gp9 with its deletion variants (NΔ54, NΔ167, and CΔ113). Functionally active gp9

(288 aa) is a trimer and consists of three domains: N-terminal, middle, and C-terminal. Spatial structure of gp9 solved by X-ray crystallography to 2.3 Å resolution [3] is shown in Fig. 1.

Earlier we determined that N-terminal deletions of polypeptide chain do not affect protein trimerization. Long deletions up to complete removal of the N-terminal and middle domains do not affect the assembly of SDS-resistant trimers, but none of them are active in *in vitro* complementation assay [4]. Co-expression of full-length gp9 with its N-terminally truncated fragment shortened by 20 aa results in the assembly of homotrimers, as well as heterotrimers, which also are SDS-resistant [5]. We have carried out experiments on co-expression of full-length gp9 with its NΔ54 and NΔ167 fragments shortened by 54 and 167 aa, accordingly, from the N-terminus, and CΔ113 fragment shortened by 113 aa from the C-terminus that form N-, middle, and C-domains, respectively. Products of co-expression were analyzed by electrophoresis under native conditions (Fig. 2a).

Additional protein bands have been observed on co-expression of gp9 with NΔ54 and NΔ167 fragments (Fig. 2a, lanes IV and V). These bands were excised from the gel, heated in buffer containing SDS, and resolved by electrophoresis under denaturing conditions (Fig. 2b). As shown in Fig. 2b, additional bands are heterotrimers of full-length gp9 and NΔ54 truncated fragment (data not shown for heterotrimers of gp9 and NΔ167). One of the heterotrimers is a product of the assembly of one full-length gp9 chain and two truncated chains (Fig. 2b, lane *3*), while another is a product of the assembly of two full-length chains and one truncated chain (Fig. 2b, lane *4*). Co-expression of gp9 and its CΔ113 fragment shortened by 113 aa from the C-terminus did not result in the assembly of heterotrimers (Fig. 2a, lane I). Only the band corresponding to the trimer of wild-type protein was observed in this case. This indicates that the C-terminal part of the gp9 molecule is essential for trimerization, and oligomerization apparently occurs post-translationally.

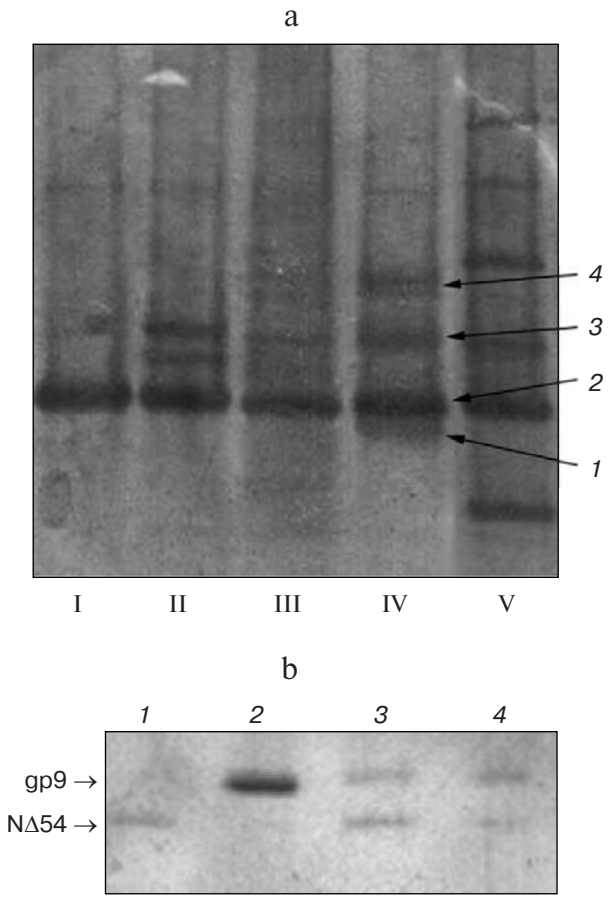

Fig. 2. Native PAGE (8%) analysis (a) of gp9 (III) and products of its co-expression with CΔ113 (I), NΔ20 (II), NΔ54 (IV), and NΔ167 (V). Arrows show protein bands that were excised from lane IV, heated in SDS-sample buffer, and loaded on 12% SDS-polyacrylamide gel (b). Lane numbers on gel (b) are correspondent to band numbers on gel (a). Arrows show positions of monomer forms of gp9 and its NΔ54 deletion mutant.

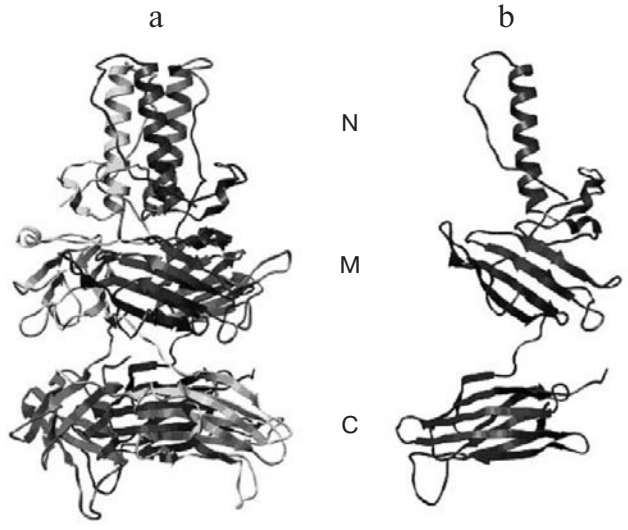

Fig. 1. Crystal structure of gp9 of bacteriophage T4 determined to 2.3 Å resolution: a) trimer; b) monomer in trimer. N-, C-, and middle domains are designated by letters N, C, and M, respectively [3].

Design of gp9 C-truncated fragments. We previously found that expression of gene 9 with C226 amber mutation resulted in two recombinant proteins. One of them is a fragment (CΔ7) shortened by 7 aa from the C-terminus, and another is full-length protein (gp9*) with Gln282→Ala single point substitution. It was shown that deletion in the CΔ7 mutant completely blocks proper folding of polypeptide chain, and its monomer is not able to oligomerize. The Gln282→Ala single point substitution does not prevent protein trimerization [8].

We carried out more detailed mutagenesis of the C-terminus to establish how deletions of different length

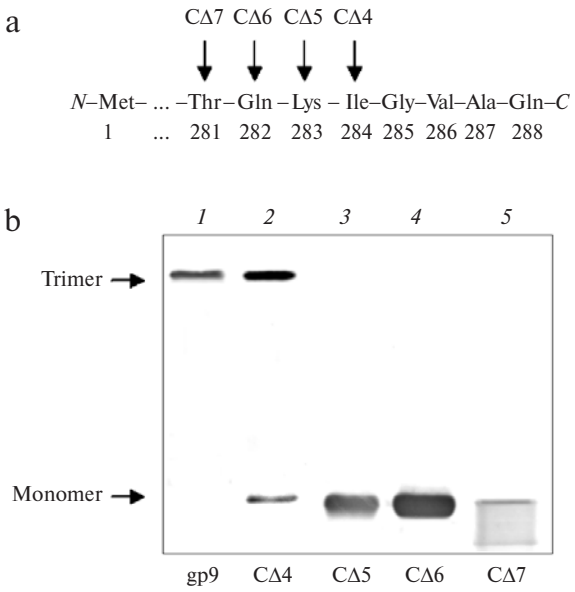

Fig. 3. a) The C-terminal part of the gp9 polypeptide chain. Arrows show the C-terminal amino acid residues of deletion fragments. b) SDS-PAGE (12%) analysis of non-heated samples of full-length gp9 and its C-truncated fragments. Arrows show positions of monomer and trimer forms of proteins.

influence the protein folding, oligomerization, and functional activity. We have constructed plasmid vectors for expression in *E. coli* cells of gp9 CΔ4, CΔ5, and CΔ6 deletion mutants shortened by 4, 5, and 6 aa from the C-terminus of the polypeptide chain, respectively (Fig. 3a).

All vectors constructed are characterized by high level expression. Sensitivity of mutant trimers to SDS (Fig. 3b) and their biological activity (table) have been studied. It was found that deletion of 4 aa has no affect on trimer stability (Fig. 3b, lane *2*) and biological activity (table).

Consecutive removal of 5 and 6 aa from the C-terminus does not affect the functional properties of the proteins. CΔ5 and CΔ6 mutants increase titer of 9$^-$-defective particles in *in vitro* complementation assay (table), but their trimers lose resistance to SDS (Fig. 3b, lanes *3* and *4*). The different degree of stability of trimers depending on deletion length is apparently determined by interactions between monomers within a trimer. According to X-ray data, Gly285 (the fourth amino acid residue from the C-terminus) makes a hydrogen bond with Thr178 of a neighboring monomer, Ile284 (the fifth amino acid residue from the C-terminus) has hydrophobic interaction with Thr205 of a neighboring monomer, and Gln282 (the seventh amino acid residue from the C-terminus) makes two hydrogen bonds with Leu203 and Thr205 of a neighboring monomer in the wild-type protein [3]. In the case of the CΔ4 mutant, disruption of one hydrogen bond (between Gly285 and Thr178 of the adjacent monomer) with preservation of other interactions between monomers is not essential for assembly of stable trimer. In the CΔ5 and CΔ6 mutants, additional disruption of hydrophobic interaction between Ile284 and Ile205 leads to loss of resistance of their trimers to SDS. Deletion of the last seven C-terminal residues, up to Gln282, results in the formation of monomer that completely loses the ability to trimerize (Fig. 3b, lane *5*). Therefore, the C-terminal Gln282-Gly285 segment, which is a part of β-strand [3], is essential for assembly of stable, SDS-resistant gp9 trimer.

Titer of T4 infection particles in *in vitro* complementation of phage 9$^-$-particles with full-length gp9 and its C-truncated fragments

Recombinant protein	Titer of phage particles (plaque forming units/ml)
– (negative control)	$(1.5 \pm 0.5) \cdot 10^8$
Full-length gp9	$(6 \pm 0.4) \cdot 10^{10}$
CΔ4	$(5.5 \pm 0.5) \cdot 10^{10}$
CΔ5	$(5.0 \pm 0.3) \cdot 10^{10}$
CΔ6	$(4.3 \pm 0.2) \cdot 10^{10}$

Mutagenesis of Gln282 in the gp9 molecule. We have carried out 3′-end mutagenesis of gene 9 by PCR using primers containing random nucleotide substitutions. Plasmid vectors for expression of mutants with substitution of Gln282 to other amino acid residues have been constructed. Amino acid residue substitutions were identified by DNA sequencing of mutated gene 9. Single point mutants were tested in *in vitro* complementation assay and analyzed by electrophoresis under denaturing conditions to study SDS-sensitivity of their trimers. It was found that substitutions of Gln282 to Ser, Cys, Ile, Leu, and Ala did not affect the assembly of functionally active trimers (data not shown). Since trimers of all these mutant proteins were SDS-sensitive, it is possible to conclude that Gln282 together with other C-terminal amino acid residues (Ile284 and Gly285) influences the interaction of monomers within a trimer.

Along with soluble point mutants described above, two insoluble mutants have been revealed. Proteins with substitution of Gln282 to Pro or Arg were expressed in the form of inclusion bodies under the same conditions (37°C). The picture of fractionation of *E. coli* cells producing mutant proteins with Gln282→Pro and Gln282→Arg substitutions at 37°C is present in Fig. 4.

As shown in Fig. 4, the proteins are in the pellet fraction. Decreasing expression temperature to 25°C resulted in production by *E. coli* cells of soluble proteins. Similar to point mutants listed above, they assembled into functionally active, but SDS-sensitive trimers. Thus, Gln282 apparently influences not only degree of association of polypeptide chains within trimer, but also protein folding.

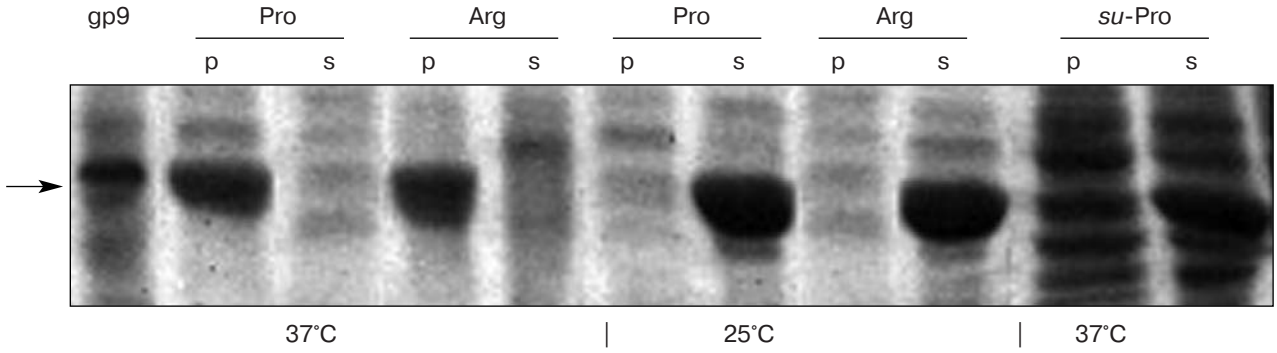

Fig. 4. SDS-PAGE (12%) analysis of wild-type gp9 and its mutants with Gln282→Pro and Gln170→Pro; Gln282→Pro; Gln282→Arg substitutions designated as *su*-Pro, Pro, and Arg, respectively, after expression at 37 and 25°C (temperature of expression is indicated under lanes), and fractionation of cell extract to supernatant (s) and pellet (p). The arrow shows the gp9 monomer.

As mentioned above, Gln282 makes hydrogen bonds with Leu203 and Thr205 of a neighboring monomer within the trimer. This suggests that substitution of Gln282 to positively charged Arg or Pro results in change of chain direction, whereupon it cannot fold quickly and properly at 37°C and thus aggregates into inclusion bodies. On decreasing temperature to 25°C polypeptide chains with these mutations obviously have time to fold properly and assemble into functionally active trimer. We have assigned Gln282→Pro and Gln282→Arg mutations to the so-called *tsf*-mutations (temperature sensitive for folding).

Protein, which has the additional Gln170→Pro mutation together with Gln282→Pro substitution, has been revealed during analysis of single point mutants. Different from single point mutant with Gln282→Pro substitution, double (Gln282→Pro, Gln170→Pro) mutant was produced by *E. coli* cells at 37°C in soluble form (Fig. 4). Since an additional mutation eliminates the effect of *tsf*-mutation, we have related it to *su*-mutation (suppressor) suppressing the temperature sensitive effect of Gln282→Pro substitution. We have constructed a vector for expression of gene 9 coding mutant with Gln170→Pro substitution. It was found that Gln170→Pro substitution in itself does not affect assembly of stable, SDS-resistant trimer. However, in combination with *tsf*-mutation it efficiently recovers productive protein folding at 37°C and results in assembly of soluble product. Testing of *su*-mutant in *in vitro* complementation assay has shown that the protein is a functionally active trimer, but it is SDS-sensitive similar to single point mutants described above. According to structure, Gln170 makes hydrogen bonds with Glu176 and Lys277 of a neighboring monomer. It might be supposed that substitution of Gln170 to positively charged Arg, probably as in the case of Gln282→Pro substitution, redirects a polypeptide chain and in that way prevents from its self-association leading to formation of inclusion bodies. It is probable that *tsf*-mutations observed by us destabilize thermolabile intermediates of gp9 folding and assembly, and *su*-mutations on the contrary stabilize them. Similar type (*tsf* and *su*) mutations have been identified and characterized for tail spike protein of phage P22 [11, 12], folding of which has been intensively studied by various methods within the two last decades.

Summarizing the data described above, we conclude that gp9 trimerization occurs post-translationally and the C-terminal part is essential for proper folding of the polypeptide chain and following oligomerization resulting in the formation of stable SDS-resistant trimer.

REFERENCES

1. Leiman, P. G., Chipman, P. R., Kostyuchenko, V. A., Mesyanzhinov, V. V., and Rossmann, M. G. (2004) *Cell*, **118**, 419-429.
2. Crowther, R. A., Lenk, E. V., Kikuchi, Y., and King, J. (1977) *J. Mol. Biol.*, **116**, 489-523.
3. Kostyuchenko, V. A., Navruzbekov, G. A., Kurochkina, L. P., Strelkov, S. V., Mesyazhinov, V. V., and Rossmann, M. G. (1999) *Structure Folding Des.*, **7**, 1213-1222.
4. Navruzbekov, G. A., Kurochkina, L. P., Kostyuchenko, V. A., Zurabishvili, T. G., Venyaminov, S. Y., and Mesyanzhinov, V. V. (1999) *Biochemistry (Moscow)*, **64**, 1266-1272.
5. Kurochkina, L. P., Vishnevskiy, A. Y., Zhemaeva, L. V., Sykilinda, N. N., Strelkov, S. V., and Mesyanzhinov, V. V. (2006) *J. Struct. Biol.*, **154**, 122-129.
6. Kostyuchenko, V. A., Leiman, P. G., Chipman, P. R., Kanamaru, S., van Raaij, M. J., Arisaka, F., Mesyanzhinov, V. V., and Rossmann, M. G. (2003) *Nat. Struct. Biol.*, **10**, 688-693.
7. Studier, F. W., and Moffat, B. A. (1986) *J. Mol. Biol.*, **189**, 113-130.
8. Zhemaeva, L. V., Sykilinda, N. N., Navruzbekov, G. A., Kurochkina, L. P., and Mesyanzhinov, V. V. (2000) *Biochemistry (Moscow)*, **65**, 1068-1074.
9. Edgar, R. S., and Wood, W. B. (1966) *Proc. Natl. Acad. Sci. USA*, **55**, 498-505.
10. Laemmli, U. K. (1970) *Nature*, **227**, 680-685.
11. Yu, M.-H., and King, J. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 6584-6588.
12. Yu, M.-H., and King, J. (1988) *J. Biol. Chem.*, **263**, 1424-1431.