Processing of the Tail Lysozyme (gp5) of Bacteriophage T4

Nanzhang Ye^{1*} and Naoki Nemoto²

Department of Life Science, Faculty of Bioscience and Biotechnology, Tokyo Institute of Technology, Midoriku, Yokohama,¹ and Department of Molecular Biology, Tokyo University of Pharmacy and Life Science, Horinouchi, Hachioji,² Japan

Received 9 October 2003/Accepted 14 June 2004

The processing site of gp5 has been determined to be between residues Val-390 and His-391, instead of Ser-351 and Ala-352 as previously reported (H. Kanamaru, N. C. Gassner, N. Ye, S. Takeda, and F. Arisaka, *J. Bacteriol.* 181:2739–2744). Moreover, the maturation of gp5 is abolished by null mutations in other hub genes, indicating that cleavage requires the interactions of several baseplate proteins.

gp5 is an essential structural component in the hub of T4 baseplates (10). Expression of a cloned gene 5 resulted in the synthesis of a protein with a mass of about 63 kDa, consistent with the size predicted from the gene sequence (14) but about 20 kDa greater than the size of gp5 present in the hub. This led to the suggestion that a 20-kDa fragment was proteolytically removed from the gp5 precursor during assembly of the baseplate and that at least one phage protein was required for processing (14). A recent report (9) contradicting earlier results (14) suggested that the 75-kDa precursor was synthesized from a cloned gene 5 and that it was cleaved spontaneously or by a host enzyme at residue 351. To resolve this issue, N-terminal and C-terminal sequence analysis was performed on mature gp5 isolated from T4 tube baseplates, and polypeptides synthesized from several cloned truncated derivatives of gene 5 were compared with the mature gp5.

N-terminal sequence analysis of the gene products detached from tube baseplates and C-terminal sequence analysis of mature gp5. All the media and buffers used in this study were described previously (1, 2, 19). Phage tails and tube baseplates were prepared according to published procedures (15, 21). Purified tube baseplates were treated with 6 M urea at 37°C for 15 min or sodium dodecyl sulfate (SDS) sample buffer at 70 and 90°C for 3 min each and were subjected to SDS-polyacrylamide gel electrophoresis (PAGE) as described by Laemmli (13). Urea treatment overcame the problem of gp5 and gp48 comigrating and being difficult to resolve (15). As shown in Fig. 1, when the tube baseplates were treated with 6 M urea, gp5 as well as gp29, gp27, gp26, gp53, gp3, and gp25 was released from the tube baseplates but gp48 was not (Fig. 1, lane A). Because all hub proteins except gp28, the inside wedge proteins gp53 and gp25 (which are close to the hub), and the tail tip protein gp3 were released from baseplates, the results

showed that the centers of the baseplates including the hubs were dissociated during urea treatment. Since gp28 was not detected in this experiment, the results indicated that gp28 is not a hub structural protein, as reported previously (10, 11). When the tube baseplates were treated with SDS buffer at 70°C, two bands appeared in the 24-kDa region (Fig. 1, lane B), the upper band being gp11 and the lower band being gp26 (confirmed by N-terminal sequence analysis) (Table 1). When the tube baseplates were treated at 90°C, the gp26 band comigrated with gp11 (Fig. 1, lane C). All the observed molecular weights of the gene products are

FIG. 1. SDS-PAGE analysis of gene products from tube baseplates. The polyacrylamide concentration was 12.5%. Tube baseplates were treated with 6 M urea at 37°C for 15 min (lane A), with SDS buffer at 70°C for 3 min (lane B), or with SDS buffer at 90°C for 3 min (lane C). Lane D, protein molecular weight standards.

* Corresponding author. Present address: Nansho Kano, Department of Ecology, Malcom Co., Ltd., 15-10 Honmachi 4-chome, Shibuya-ku, Tokyo 151-0071, Japan. Phone: 81 3 3320 5611. Fax: 81 3 3320 5615. E-mail: nyeco@hotmail.com.

TABLE 1. N-terminal sequence analysis of some gene products detached from tube baseplates

Gene product	Location	Source of sequence sample	Sequence from N terminus	Comments
gp29	Hub	Fig. 3, lane A, 77-kDa region	M-K-K-P-Q-E-M-	Met-1 retained
gp27	Hub	Fig. 3, lane A, 49-kDa region	S-M-L-Q-R-P-G-	Met-1 removed
gp5	Hub	Fig. 3, lane A, 42-kDa region	M-E-M-I-S-N-	Met-1 retained
gp11	Wedge pin	Fig. 3, lane B, 24-kDa region	S-L-L-N-N-K-A-G-V-	Met-1 removed
gp26	Hub	Fig. 3, lane A, 24-kDa region	F-D-V-R-V-G-S-K-I-I-N-	N-terminal 5 residues removed
gp53	Wedge	Fig. 3, lane A, 23-kDa region	M-L-F-T-F-F-	Met-1 retained
gp3	Tube tip	Fig. 3, lane A, 19-kDa region	S-Q-A-L-Q-Q-I-	Met-1 removed
gp25	Wedge	Fig. 3, lane A, 15-kDa region	A-N-I-N-K-L-Y-S-D-I-	Met-1 removed

close to the values calculated except for gp5 because of its processing. After electrophoresis, the proteins were transferred from the gel to a polyvinylidene difluoride (PVDF) membrane (20), and the bands of interest were subjected to automatic Edman degradation for N-terminal sequence analysis. The N-terminal sequences are almost identical to the DNA-deduced sequences except that the five N-terminal amino acid residues (MYEYK) are removed from gp26 (Table 1). Met-1 is retained or removed, consistent with reference 3. It was reported that the eight N-terminal residues of gp11 were cleaved (5) and gp26 was not located in the baseplate (10), but the present results showed that only Met-1 of gp11 is removed and gp26 is a component of the tail. One reason for the discrepancy may be interference from gp26, which comigrates with gp11 (Fig. 1, lane C).

For C-terminal sequence analysis of mature gp5, the gels were stained after electrophoresis (22) and gp5 was extracted from the excised bands (43 kDa) with a 50 mM (NH₄)HCO₃ solution containing 1% SDS at 4°C for 12 h

three times. Approximately 680 pmol (quantitative analysis as described in reference 17) of gp5 was extracted from the gels and incubated with 1.2 U of carboxypeptidase (Sigma) at 25°C. The incubation was stopped (7) at 0, 15, 30, and 45 min. One cycle of Edman degradation was carried out for each sample. The results showed that carboxypeptidase digestion produced Val, Tyr, Pro, Tyr, and Glu in sequence, and all PTH amino acid peaks showed equivalent molar yields. There was only one sequence, Glu-Tyr-Pro-Tyr-Val, at positions 386 to 390, consistent with that deduced from the DNA sequence. None of the peaks corresponding to amino acids adjacent to the C-terminal Ser-351 reported in reference 9 were found. These findings indicated that the mature gp5 has a C-terminal Val-390, and the cleavage occurs between Val-390 and His-391, not between Ser-351 and Ala-352. The calculated molecular weight of 42,900 is consistent with the observed value, 42,000 to 43,000. The calculated isoelectric point of the mature gp5 was close to 8, similar to the reported value (15).

FIG. 2. Recombinant gene 5-expressed protein products detected by Western blot analysis. After electrophoresis on SDS-12.5% PAGE, proteins were transferred to PVDF membranes. Lane A, Prestained protein molecular weight standards; lane B, gp5 separated from tube baseplates; lane C, sample from induced *E. coli* BL21(DE3) cells containing recombinant gene 5 with the codon for C-terminal Ser-360; lane D, sample from uninduced *E. coli* BL21(DE3) cells containing recombinant gene 5 with the codon for C-terminal Ser-360; lane E, sample from induced *E. coli* BL21(DE3) cells containing recombinant gene 5 with the codon for C-terminal Val-380; lane F, sample from uninduced *E. coli* BL21(DE3) cells containing recombinant gene 5 with the codon for C-terminal Val-380; lane G, sample from induced *E. coli* BL21(DE3) cells containing recombinant gene 5 with the codon for C-terminal Val-390; lane H, sample from uninduced *E. coli* BL21(DE3) cells containing recombinant gene 5 with the codon for C-terminal Val-390.

FIG. 3. Western blot analysis of processed gp5 from lysates. After electrophoresis on SDS-8% PAGE, proteins were transferred to PVDF membranes. Lane A, 25am N67 lysate; lane B, 26am N131 lysate; lane C, 27am N120 lysate; lane D, 28am A452 lysate; lane E, 29am B7 lysate; lane F, tube baseplates; lane G, 5am B256 lysate; lane H, 48am NO22X lysate; lane I, 51am S29 lysate.

Expression and detection of gp5 polypeptides. Three truncated derivatives of gene 5 were cloned (16). After induction, they produced gp5 derivatives with C-terminal Ser-360, Val-380, or Val-390 (18). Because the pET-17b vector used in this study expresses an N-terminal fusion protein, the ribosomal binding site and other unnecessary sequences were removed from the vector by restriction with XbaI and EcoRI. The XbaI site and ribosome binding site sequences were included in the 5' primer upstream of the gene 5 start codon. The stop codon and the EcoRI site were included in three distinct 3' primers for PCR amplification of the three gene 5-truncated derivatives that expressed 360, 380, or 390 amino acid polypeptides. T4 dcDNA (Takara, Kyoto, Japan) was used as a template for PCR. The primers were synthesized by Bio/Can Scientific or Applied Biosystems Japan Inc. The three gp5-derived polypeptides were examined by Western blot analysis (6, 20) by using antiserum against gp5 (maintained and used in this laboratory). As shown in Fig. 2, the gp5 C-terminal Val-390 derivative comigrated with mature gp5 obtained from tube baseplates, consistent with the assignment of this cleavage site in gp5.

Examination of the processed gp5 from amber lysates. To identify the T4 gene products responsible for gp5 processing,

T4 amber mutants with mutations in the baseplate genes were used to infect *Escherichia coli* BE (*sup*⁰) for preparation of amber lysates (1, 15), and mature gp5 was detected by Western blot analysis. We first interpreted the upper band in the blots (Fig. 3 and 4) as the uncleaved gp5 precursor or, perhaps, a cross-reacting *E. coli* protein. The results shown in Fig. 4 reveal that this band is gp23 because it has the correct size, 49 kDa (4), and it did not appear in the 23⁻ lane (Fig. 4, lane G). This indicated that gp23 has some amino acid sequences similar to gp5, because gp23 cross-reacted, as shown in Table 2 and Fig. 5. Since only the 26am N131 lysate gave rise to mature gp5, these results suggested that gp5 processing is a result of the interactions of several baseplate gene products and does not involve an *E. coli* enzyme or autolysis. If the processing occurred either via an *E. coli* enzyme or autolytically, the mature gp5 should be detected in all lysates (except for gene 5 amber lysate); however, this was not the case. On the other hand, because the gene 26 amber mutant is unable to form hubs (10) and the mature gp5 was detected in this lysate, gp5 processing may occur before hub formation. This suggests that gp5 enters the hub pathway after processing.

FIG. 4. Western blot analysis of processed gp5 from lysates. After electrophoresis on SDS-8% PAGE, proteins were transferred to PVDF membranes. Lane A, prestained protein molecular mass standards (Bio-Rad Laboratories); lane B, 25am N67 lysate; lane C, 26am N131 lysate; lane D, 28am A452 lysate; lane E, 29am B7 lysate; lane F, 5am N135 lysate; lane G, 23am H11 lysate (tails); lane H, uninfected *E. coli* BE (*sup*⁰); lane I, 51am S29 lysate.

```

1'  MEMISNNLNWFVGVVEDRMDPLKLGRVVRVVRVGLHPPQRA
46" TAPEYKDEKIAQAFGSFLTETAEIGGDHGYNATNIAAGQTS

41'  QGDVMGIPTEKLPWMSVIQPI TSAAMSGIGGSVTGPVEGT
86"  GAVTQIGPAVMGMVRRRAIPNLIAFDICGVQPMNSPTGQVF

81'  RVYGHFLDKWKTNGLIVLGTYG GIVREKPNRLEGFS DPTGQ
126" ALRAVYGKD PVAAGAKEAFHPMYGPDAMFS GQGAAKKFPA

121' YPRRLGNDTNVLNQGGGEVGYDSSSNVIQDSNLDTA INPDD
166" LAASTQT TVGD IYTHFFQETGT VYLQASVQVTIDAGATDA

161' RPLSEIPTDDNPNMSMAEMLRRDEGLRLKVYWDTEGYPTI
206" AKLDAEIKKQMEAGALVEIAEGMATSI AELEQEGFN GSTDN

201' GIGHLIMKQPVRAMAQINKVLSKQVGREITGNPGSITMEE
246" PWNEMGFRI DKQVIEAKSRQLKAAYSIELAQDLRAVHGM D

241' ATTLFERDLADMQRDIKSHSKVGPVWQAVNRSRQMALENM
286" ADAELSGILATEIMLEINREVVDWINYSAAQVGKSGMTLT P

281' AFQMGVGGVAKFNTMLTAML AGDWEKAYKAGRDSLWYQQT
326" GSKAGVDFDQDPIDIRGARWAGE SFKALLFQIDKEAVEIA

321' KGRASRVTMII LTGNLESYGVEVKTPARSLSAMAATVAKS
366" RQTGRGEGNFIIASRNVVNVLASVDTGISIYAAQGLATGFS

↓
361' SDPADPPIPNDSRILFKEPVSSSYKGEYPPYV
406" TDTTKSVFAGV LGGKYRVYIDQYAKQDYFT

```

FIG. 5. Similar regions of the amino acid sequence between gp5 (') and gp23 ("). The ratio of identical amino acid pairs in whole gp5/gp23 (1'/46" to 390'/435") is 8.2%, much higher than 5.0% in two common proteins. When the number of conservative changes is included, the total similarity ratio of amino acid pairs between gp5 and gp23 reaches 30.0%. In the main region from 155'/200" to 362'/407" (208 amino acid pairs, the region indicated by arrows), the ratio of identical amino acid pairs in gp5 and gp23 is 11.5% and the total similarity ratio including conservative changes is 35.1%. Colons indicate identical amino acids. Periods indicate conservative changes.

Western blot analysis of gp26. The assignment of gp26 as a band in SDS-PAGE had some ambiguity because gp26 comigrated with gp11 (Fig. 1). To identify gp26 on SDS-PAGE, an antiserum against a gp26 C-terminal peptide was prepared according to Iwai et al. (8), and Western blot analysis was carried out. As shown in Fig. 6, the 24-kDa band was found to be gp26, which was confirmed by N-terminal sequence analysis (Table 1). The molecular mass of gp26 was similar to that deduced from the nucleotide sequence, not 41 kDa as reported (12), and it is unlikely that the peptide cleaved from precursor gp5 was transferred to gp26 (14).

In conclusion, these results demonstrate that gp5 is processed proteolytically, with a cleavage between Val-390 and His-391, resulting in the loss of 20 kDa and the conversion of the 63-kDa precursor gp5 to the 43-kDa mature gp5 found in the baseplate, and support the earlier suggestion (14) that processing of gp5 depends on other T4 genes.

FIG. 6. Western blot analysis of gp26. The tube baseplate proteins were transferred from SDS-12.5% PAGE gels to PVDF membranes. Lane A, gp26 band detected with antiserum against the C-terminal peptide of gp26; lane B, gp5 band detected with antiserum against gp5; lane C, gp26 and gp5 bands detected with a mixture of the two antisera.

TABLE 2. Amino acid fragments with similar sequences in gp5 and gp23

Fragment (amino acid positions in gp5 ['] and gp23 ['])	No. of amino acid pairs	No. (%) of identical amino acid pairs
355'/329" to 377'/351"	23	7 (30.4)
341'/300" to 360'/319"	20	5 (25.0)
340'/19" to 359'/38"	20	5 (25.0)
330'/376" to 356'/402"	27	7 (25.9)
288'/330" to 310'/352"	23	6 (26.1)
285'/330" to 313'/358"	29	8 (27.6)
268'/361" to 289'/382"	22	6 (27.3)
265'/306" to 284'/325"	20	5 (25.0)
255'/176" to 274'/195"	20	5 (25.0)
231'/238" to 250'/257"	20	5 (25.0)
221'/266" to 362'/407"	142	19 (13.4)
215'/210" to 234'/229"	20	5 (25.0)
213'/206" to 283'/276"	71	10 (14.1)
209'/83" to 229'/103"	21	7 (33.3)
190'/188" to 209'/207"	20	5 (25.0)
186'/190" to 327'/331"	142	16 (11.3)
136'/373" to 156'/393"	21	6 (28.6)
117'/120" to 136'/139"	20	5 (25.0)
102'/97" to 121'/116"	20	6 (30.0)
95'/91" to 114'/110"	20	5 (25.0)
87'/10" to 106'/29"	20	5 (25.0)
84'/60" to 103'/79"	20	5 (25.0)
62'/84" to 88'/110"	27	9 (33.3)
61'/55" to 80'/74"	20	5 (25.0)
48'/246" to 67'/263"	20	6 (30.0)
33'/73" to 53'/93"	21	6 (28.6)
32'/389" to 51'/408"	20	6 (30.0)

^a A portion of the fragments of gp5 and gp23 which contain over 20 amino acid residues and reach over 25.0% identical amino acid residues in the sequences or contain over 50 amino acid residues and reach over 10.0% identical amino acid in the sequences are listed.

REFERENCES

1. Arisaka, F., T. Nakako, H. Takahashi, and S.-I. Ishii. 1988. Nucleotide sequence of the tail sheath gene of bacteriophage T4 and amino acid sequence of its product. *J. Virol.* **62**:1186–1193.
2. Arisaka, F., S. Takeda, K. Funane, N. Nishijima, and S. Ishii. 1990. Structural studies of the contractile tail sheath protein of bacteriophage T4. 2. Structural analysis of the tail sheath protein, Gp18, by limited proteolysis, immunoblotting, and immunoelectron microscopy. *Biochemistry* **29**:5057–5062.
3. Ben-Bassat, A., and K. Bauer. 1987. Amino-terminal processing of proteins. *Nature* **369**:315.
4. Black, L. W., M. K. Showe, and A. C. Steven. 1994. Morphogenesis of the T4 head, p. 218–258. *In* J. D. Karam (ed.), *Molecular biology of bacteriophage T4*. ASM Press, Washington, D.C.
5. Coombs, D. H., and F. Arisaka. 1994. T4 tail structure and function, p. 259–281. *In* J. D. Karam (ed.), *Molecular biology of bacteriophage T4*. ASM Press, Washington, D.C.
6. Duda, R. L., M. Gingery, and F. A. Eiserling. 1986. Potential length determiner and DNA injection protein is extruded from bacteriophage T4 tail tubes in vitro. *Virology* **151**:296–314.
7. Hayashi, R. 1983. Carboxypeptidase Y protein. *Nucleic Enzyme Exp. Lecture* **28**:1421–1431. (In Japanese.)
8. Iwai, K., S.-I. Fukuoka, T. Fushiki, K. Kido, Y. Sengoku, and T. Semba. 1988. Preparation of a verifiable peptide-protein immunogen: direction-controlled conjugation of a synthetic fragment of the monitor peptide with myoglobin and application for sequence analysis. *Anal. Biochem.* **171**:277–282.
9. Kanamaru, H., N. C. Gassner, N. Ye, S. Takeda, and F. Arisaka. 1999. The C-terminal fragment of the precursor tail lysozyme of bacteriophage T4 stays as a structural component of the baseplate after cleavage. *J. Bacteriol.* **181**:2739–2744.
10. Kikuchi, Y., and J. King. 1975. Genetic control of bacteriophage T4 baseplate morphogenesis. II. Mutants unable to form the central part of the baseplate. *J. Mol. Biol.* **99**:673–694.
11. Kikuchi, Y., and J. King. 1975. Genetic control of bacteriophage T4 baseplate morphogenesis. III. Formation of the central plug and overall assembly pathway. *Mol. Gen. Genet.* **99**:695–716.
12. Kozloff, L. M., and M. Lute. 1984. Identification of bacteriophage T4D gene products 26 and 51 as baseplate hub structural components. *J. Virol.* **52**:344–349.
13. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**:680–685.
14. Mosig, G., G. W. Lin, J. Franklin, and W.-H. Fan. 1989. Functional relationships and structural determinants of two bacteriophage T4 lysozymes: a soluble (gene e) and a baseplate-associated (gene 5) protein. *New Biol.* **1**:171–179.
15. Nakagawa, H., F. Arisaka, and S.-I. Ishii. 1985. Isolation and characterization of the bacteriophage T4 tail-associated lysozyme. *J. Virol.* **54**:460–466.
16. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
17. Smith, P. K., R. I. Krohn, G. T. Hermanson, A. K. Mallia, F. H. Gartner, M. D. Provenzano, E. K. Fujimoto, N. M. Goeke, B. J. Olson, and D. C. Klenk. 1985. Measurement of protein using bicinchoninic acid. *Anal. Biochem.* **150**:76–85.
18. Tabor, S., and C. C. Richardson. 1985. A bacteriophage T7 RNA polymerase / promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci. USA* **82**:1074–1078.
19. Takeda, S., F. Arisaka, S.-I. Ishii, and Y. Kyogoku. 1990. Structural studies of the contractile tail sheath protein of bacteriophage T4. 1. Conformational change of the tail sheath upon contraction as probed by differential chemical modification. *Biochem.* **29**:5050–5056.
20. Towbin, H., T. Staehelin, and J. Gordon. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. USA* **76**:4350–4354.
21. Tshopp, J., F. Arisaka, R. V. Driel, and J. Engel. 1979. Purification, characterization and reassembly of the bacteriophage T4D tail sheath protein P18. *J. Mol. Biol.* **128**:247–258.
22. Wyckoff, M., D. Rodbard, and A. Chrambach. 1977. Polyacrylamide gel electrophoresis in sodium dodecyl sulfate-containing buffers using multiphasic buffer systems: properties of the stack, valid R_F -measurement, and optimized procedure. *Anal. Biochem.* **78**:459–482.