Biol. Chem., Vol. 382, pp. 1049–1055, July 2001 · Copyright © by Walter de Gruyter · Berlin · New York

Identification and Crystallisation of a Heat- and Protease-Stable Fragment of the Bacteriophage T4 Short Tail Fibre

Mark J. van Raaij[1,2,*], Guy Schoehn[1], Michel Jaquinod[3], Keith Ashman[4,a], Martin R. Burda[5] and Stefan Miller[5]

[1] European Molecular Biology Laboratory, Grenoble Outstation, BP 181, F-38042 Grenoble, France
[2] Gorleaus Laboratoria, Leiden University, P.O. Box 9502, NL-2300 RA Leiden, The Netherlands
[3] Institut de Biologie Structurale, CEA-CNRS, 41 rue Jules Horowitz, F-38027 Grenoble cedex 1, France
[4] European Molecular Biology Laboratory, Meyerhofstr. 1, D-69117 Heidelberg, Germany
[5] Institut für Biophysik und physikalische Biochemie, Universitätsstraße 31, Universität Regensburg, D-93040 Regensburg, Germany

* Corresponding author

Irreversible binding of T-even bacteriophages to *Escherichia coli* is mediated by the short tail fibres, which serve as inextensible stays during DNA injection. Short tail fibres are exceptionally stable elongated trimers of gene product 12 (gp12), a 56 kDa protein. The N-terminal region of gp12 is important for phage attachment, the central region forms a long shaft, while a C-terminal globular region is implicated in binding to the bacterial lipopolysaccharide core.

When gp12 was treated with stoichiometric amounts of trypsin or chymotrypsin at 37 °C, an N-terminally shortened fragment of 52 kDa resulted. If the protein was incubated at 56 °C before trypsin treatment at 37 °C, we obtained a stable trimeric fragment of 3 × 33 kDa lacking residues from both the N- and C-termini. Apparently, the protein unfolds partially at 56 °C, thereby exposing protease-sensitive sites in the C-terminal region and extra sites in the N-terminal region. Well-diffracting crystals of this fragment could be grown.

Our results indicate that gp12 carries a stable central region, consisting of the C-terminal part of the shaft and the attached N-terminal half of the globular region. Implications for structure determination of the gp12 protein and its folding are discussed.

Key words: Bacteriophage T4 adhesin / Crystallisation / Electron microscopy / Heat and protease stability / Short tail fibre / X-ray diffraction.

[a] Present address: Samuel Lunenfeld Institute, 600 University Avenue, Toronto, Ontario M5G 1X5, Canada

Introduction

Adsorption of T4 bacteriophage to the *Escherichia coli* (*E. coli*) host cell is mediated by six long and six short tail fibres. The long tail fibres are composed of the gene products gp34, gp35, gp36 and gp37 (in this order from the phage attachment end to the receptor-binding end; King and Laemmli, 1971; Bishop *et al*., 1974). Except for the monomeric gp35, all of these proteins are trimeric (Cerritelli *et al*., 1996). The long tail fibres recognise the outer membrane protein (OmpC) or lipopolysaccharide (LPS) of *E. coli* B and are responsible for the initial, reversible, attachment of the virion. After at least three long tail fibres have bound, the short tail fibres, which are normally incorporated in a compact conformation in the base plate, extend. They bind irreversibly to the host cell LPS core region (Kells and Haselkorn, 1974), serving as inextensible stays during contraction of the tail-sheath and penetration of the cell envelope by the tail tube (Makhov *et al*., 1993). The short tail fibre does not contribute to host specificity and is interchangable between T-even bacteriophages (Riede, 1987).

Short tail fibres are composed of a single protein, gp12 (Mason and Haselkorn, 1972), a homotrimer of 526 residues per subunit. They are attached to the baseplate *via* their N-termini, while the C-termini are thought to bind to the LPS core (Makhov *et al*., 1993). The overall structure of the protein as seen by electron microscopy is very similar to the adenovirus fibre and contains a thin N-terminal virus attachment sequence, a long central shaft region and a more globular albeit elongated C-terminal receptor-attachment domain (head domain). Like adenovirus fibre (Mitraki *et al*., 1999), gp12 resists dissociation by sodium dodecylsulphate (SDS) at ambient temperature and monomerisation requires heating of the polypeptide chains (King and Laemmli, 1971).

The structure of the adenovirus central shaft has been suggested to be a triple cross-beta helix (Stouten *et al*., 1992). Cerritelli *et al*. (1996) proposed the shafts of gp12 and gp37 to contain this same fold, based on scanning transmission electron microscopy data (*i. e*. the estimated similar overall dimensions), together with the presence of repetitive sequences. Circular dichroism spectra of gp12 (Burda *et al*., 2000) were very similar to those of adenovirus fibre (Green *et al*., 1983). Subsequently, van Raaij *et al*. (1999) showed the adenovirus fibre shaft to contain a different fold, a triple beta-spiral, in which a conserved glycine or proline residue and hydrophobic residues are important. Gp12 also contains repeats with

a conserved glycine or proline residue and hydrophobic residues, but their relative positions are not the same and they also contain conserved threonines. The adenovirus repeats can be summarised as XXφXφXφGXGφXφXX or XXφXφXφXXPφXφXX (X is any amino acid, φ is hydrophobic), while protein sequences of gp12 contain repeats that can be noted as follows: XTXXXXGφφXφXTXXXφ and XGXXXXXφφTPXXφXXX. Therefore, the gp12 sequence may not be compatible with folding into a triple beta-spiral identical to that in the adenovirus fibre and may have a different, albeit beta-sheet containing, fold.

Assembly of the short fibre depends on the bacteriophage chaperone gp57, an oligomer of a 79 residue polypeptide, which is not present in the mature phage. In mutants missing gp57 (Hashemolhosseini *et al*., 1996), gp12 fails to trimerise and assemble into the resulting, defective, phage particles (King and Laemmli, 1971; Ward and Dickson, 1971). Gp57 has been shown to act as a molecular chaperone in the folding pathway of gp12 (Burda and Miller, 1999), but at present it is not known which region of gp12 needs the chaperone for correct folding and therefore the mechanism of gp57 action is not understood in molecular detail.

In order to study the folding and structure of gp12, we have performed partial unfolding and proteolysis experiments. At 37 °C, we observed only N-terminal deletions by treatment with trypsin or chymotrypsin. However, if we incubated the protein at 56 °C, proteolysis was also observed at the C-terminus. The resulting, N-terminally and C-terminally shortened 33 kDa fragment remained trimeric and could be crystallised.

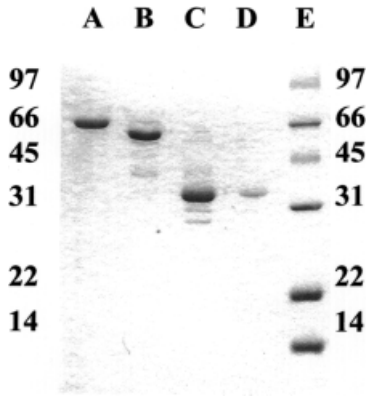

Fig. 1 Polyacrylamide Gel Electrophoresis of Small-Scale Proteolysis Experiments.
In lane A, full-length protein was electrophoresed, lanes B, C and D contain the results of the proteolysis experiments. Full-length gp12 was incubated for 30 min at 37 °C (lane B), 56 °C (lane C) or 75 °C (lane D) and subsequently incubated with stoichiometric amounts of trypsin as described in the Materials and Methods section. A distinct band is not visible for trypsin, presumably due to auto-cleavage of the protease. Lane E contains a set of size marker proteins, of which the sizes in kDa are shown on the left and right. Samples were heated to 95 °C for 5 min to disrupt the trimeric association and subsequently electrophoresed on a 15% (w/v) polyacrylamide gel in the presence of 0.1% (w/v) SDS. Staining was done with Coomassie Brilliant Blue.

Results

The Full-Length Protein

The gp12 protein was co-expressed with the bacteriophage T4 chaperone gp57 (Burda and Miller, 1999) and purified as described in the Materials and Methods section, using hydrophobic interaction and gel filtration chromatographic techniques (Figure 1, lane A). Although the protein can be concentrated to high concentrations by ammonium sulphate precipitation, we could not obtain crystals, even though we investigated a wide range of different conditions. The protein precipitates at low salt (less than 0.1 M ammonium sulphate) and high salt concentrations (greater than 0.7 M ammonium sulphate) at room temperature. Even at intermediate salt concentrations, the protein appears to aggregate aspecifically, as it was difficult to concentrate by ultrafiltration (unpublished observations). This aggregation tendency may explain why crystals of the full-length protein could not be obtained.

Proteolysis Experiments

Initial proteolysis experiments, performed with both trypsin and chymotrypsin, showed that the protein is very resistant to digestion at 37 °C, requiring stoichiometric amounts of protease to generate a 52 kDa sub-fragment of the 56 kDa protein (Figure 1, lane B). The heating and proteolysis experiments were done at gp12 concentrations of around 1 mg/ml. N-terminal sequence analysis of the trypsinised protein showed that the protein now starts at residue Ala-38 or Gly-43 (Table 1), indicating that a small portion of the N-terminus has been removed. The N-terminal fragment removed may correspond to the proposed base-plate binding region, which, in the absence of the base-plate, may be suceptible to digestion. The mass of the resulting protein was 51.9 kDa, as determined by mass spectrometry (Table 1). This suggests that the C-terminus of the protein is still intact. The resulting fragment (the '37 °C fragment') is singly underlined in Figure 2. SDS-polyacrylamide gel electrophoresis (SDS-PAGE) of the protein without prior heating showed this fragment was still trimeric (not shown). We noted that the N-terminally shortened protein had a less tendency to precipitate and was easier to concentrate by ultrafiltration. Nevertheless, crystallisation trials of this 52 kDa fragment were also not successful.

We subsequently tried to partially unfold the protein by heating it to various temperatures before proteolysis at 37 °C. The samples were then analysed by SDS-PAGE, N-terminal sequencing and mass spectrometry. Heating to 56 °C before proteolysis led to the appearance of a fragment of around 33 kDa (the '56 °C fragment'). N-terminal sequence analysis showed the protein now starts at residue Leu-84 (Table 1), indicating that an extra portion of the N-terminus has been removed. The mass of

Table 1 N-Terminal Sequences and Determined Masses of Fragments Generated.

Protein	Molecular mass in kDa as determined by mass spectrometry (minor fragments in brackets)	N-terminal sequence as determined by Edman degradation (minor fragment in brackets)	Start and end sequence of proposed fragments	Molecular mass in kDa (calculated)
Full-length	not determined	SNNTYQHVSN	(a) SNNTY - IKVKE	(a) 56.083
Fragment 37 °C	51.9 (55.3)	AISPAGVNGV (GVNGVPDASS)	(a) AISPA - IIKVK (b) GVNGV - IIKVK	(a) 51.893 (b) 51.453
Fragment 56 °C	32.8 (24.9, 26.5)	LSYPNATEAV	(a) LSYPN - LPDMR (b) LSYPN - GLFVR	(a) 32.682 (b) 33.254

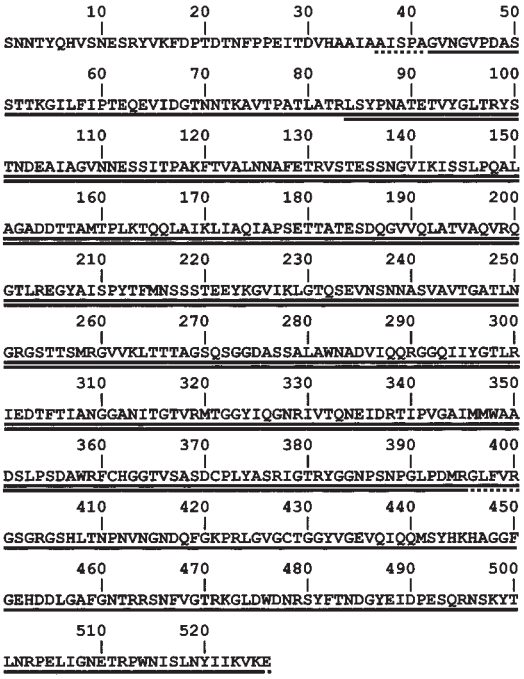

Fig. 2 Protein Sequence of gp12.
The proposed 52 kDa proteolytic fragment is underlined, the proposed 33 kDa proteolytic fragment doubly underlined. Dashed lines indicate variants or uncertainties.

the resulting protein was 32.7 kDa, as determined by mass spectrometry (Table 1). This suggests that a part of the C-terminus of the protein has also been removed. In view of the measured mass of the resulting fragment, cleavage after residue Arg-395, a trypsin site, appears most likely. The proposed fragment is doubly underlined in Figure 2.

Part of the protein appears to partially unfold at 56 °C, allowing the proteases access to certain sites that were not accessible in the folded protein or in the protein partially unfolded at 37 °C. Whether the protein forms protease-sensitive aggregates after cooling to 37 °C or whether it just does not completely refold is presently not known. The 33 kDa fragment remained trimeric as evidenced by SDS-PAGE without prior boiling of the sample (not shown), behaved well during gel filtration (migrating as a trimer, not shown), could be concentrated by ultrafiltration without problems and subsequently crystallised (see below).

Heating to 75 °C before proteolysis gave the same fragment, but at considerably lower yield. Part of the protein appears to partially unfold like at 56 °C, giving rise to the 33 kDa fragment, while most of the protein appears to completely unfold, subsequently becoming hydrolysed completely.

Electron Microscopic Analysis

The native protein shows eight smaller and bigger globules separated by thinner neck regions and a total length of 330 Å (Figure 3A). From the bottom (see the averaged image in the inset of Figure 3A), the first three small globules appear somewhat grouped together and are separated by a longer neck region from a second group of three small globules. This second group of three small globules is then connected by another longer neck region to the largest of the globules, which is itself closely associated with the second largest globule. The bottom six repeats of 40 Å each comprise a total of around 240 Å. The two larger globular parts at the top together have a length of around 90 Å.

Makhov *et al*. (1993) performed a scanning transmission electron microscopy (STEM) analysis of gp12. The overall shape and dimensions are the same in both analyses although details are different. The picture from STEM appears more regular, while our negative stain analysis shows a structure divided into globules. The STEM analysis showed a long shaft region connected to an arrow-shaped head region, and Makhov *et al*. (1993) concluded the C-terminal end is the head domain, while the N-terminus corresponds to the start of the shaft. It is likely the shaft region observed corresponds to the six small globules we observe and thus the N-terminus would be located at the bottom in the averaged image of Figure 3A. The two larger globules would then correspond to the arrow-shaped head. The top globule being smaller in negative stain electron microscopy is consistent with the arrow-shape observed by the scanning transmission method.

When we compare the native protein with the protein digested at 37 °C (Figure 3B) there are no obvious changes (Figure 3D). The part removed by digestion is

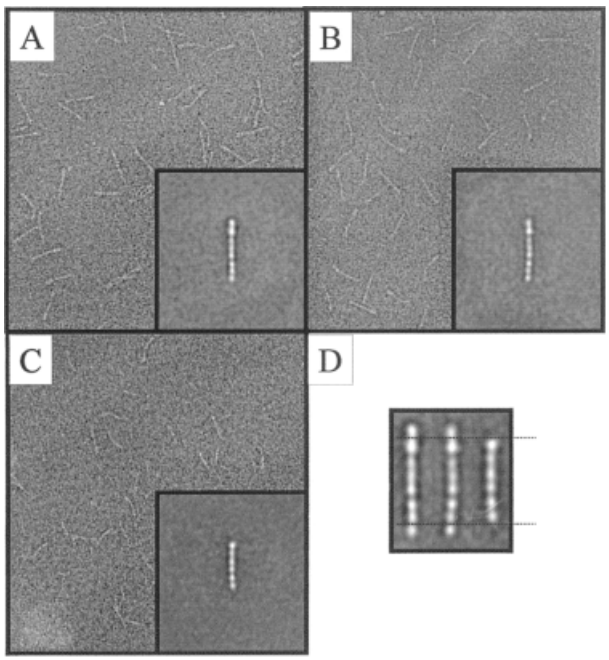

Fig. 3 Electron Micrographs of Full-Length gp12 and Proteolytic Fragments.
Averaged images are shown as insets in parts (A), (B) and (C). (A) Full-length gp12 expressed in *Escherichia coli*. (B) Fragment generated by proteolysis with trypsin at 37 °C (52 kDa fragment). (C) Fragment generated by proteolysis with trypsin after partial unfolding of the protein at 56 °C (33 kDa fragment). (D) Gallery of averaged images with, from left to right: the full-length protein, the 52 kDa fragment and the 33 kDa fragment. All samples were negatively stained with silicotungstate. For full details see Materials and Methods.

too small or too disordered to be seen in electron microscopy. As this region is likely to be involved in attachment to the base plate, it could be that it is disordered in absence of the base plate and not visible in electron microscopic pictures of the full-length protein.

In contrast, the comparison with the protein that was incubated at 56 °C and then digested (Figure 3C) clearly shows that two globular domains have been removed, a small one at the bottom or putative N-terminus and a large one at the top or putative C-terminus (Figure 3D). This leads to a fragment containing five of the small globules and the largest of the globules of the head still remaining, with a length of just under 250 Å (5 × 40 Å plus one just below 50 Å). As the receptor-binding activity is thought to be located in the C-terminal head, our fragment may or may not still contain the receptor binding site.

Large-Scale Preparation and Crystallisation of the 33 kDa Fragment

For large-scale proteolysis, purified and ammonium sulphate precipitated gp12 was redissolved and heated to 56 °C. After cooling, the protein was trypsinised as described in Materials and Methods. The resulting trimeric 33 kDa fragment was subsequently purified by gel filtration.

The 33 kDa fragment was crystallised at pH 5.6 by vapour diffusion from solutions containing tertiary butanol (2-methylpropane-2-ol) and glycerol as described in Materials and Methods. Single crystals belonging to space group P321 appeared overnight and grew to a maximal size around 0.2 × 0.2 × 0.8 mm (Figure 4A) in periods varying from overnight to several weeks. They diffracted X-rays to diffraction spacings of 1.5 Å. Given the size of the unit cell (a = b = 51 Å, c = 249 Å), the space group (P321) and the size of the fragment (around 33 kDa), we expect the asymmetric unit to consist of one gp12 monomer, leaving room for 50% solvent content. The three-fold symmetry axis of the gp12 trimer should then coincide with the crystallographic three-fold symmetry axis.

Electron Microscopy of gp12 Crystals

Negative stain electron microscopy analysis of partially dissolved crystals (Figure 4B) shows that the gp12 fragments are packed co-linearly. The vertical repeat (249 Å, Figure 4C) corresponds well with the length of the fragment, while the horizontal repeat (51 Å, Figure 4C) corre-

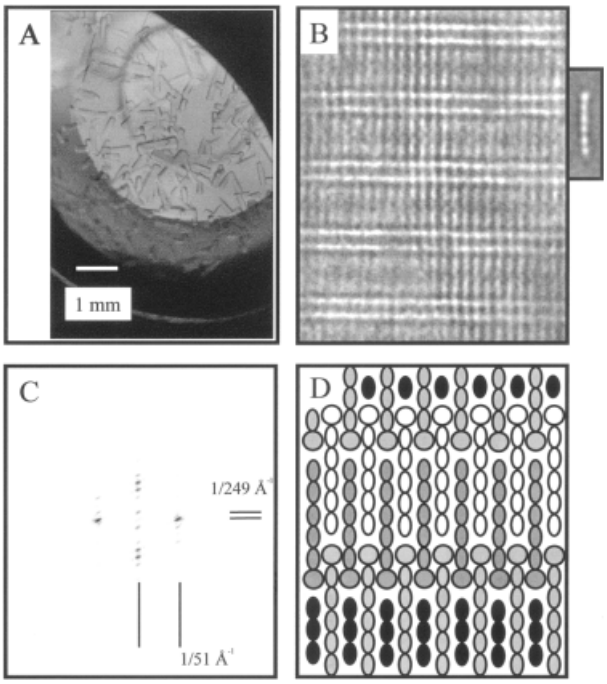

Fig. 4 Crystallisation of the 33 kDa Proteolytic Fragment of gp12.
(A) Photograph of gp12 crystals. They belong to space group p_{321} with cell parameters of a = b = 51 Å, c = 249 Å. The white bar at the bottom left is approximately 1 mm long. (B) Electron micrograph of a small crystal fragment stained with uranyl acetate (negative stain). To the right, the averaged picture of the 33 kDa fragment is shown (identical to Figure 3C). (C) Electron diffraction pattern. Spacings are indicated. (D) Proposed crystal packing. Different layers of the protein are shown from white to black in different shades of grey.

sponds to the largest width of the fragment. The cell parameters are identical to those determined by X-ray crystallography (see below). From the electron microscopic analysis combined with the knowledge of the space group, P321, and cell parameters, we can propose a model for the likely packing of the fragment in the crystals (Figure 4D). This will allow us to determine an approximate mask of the protein in the crystal lattice, which may be useful in future crystallographic phase determination and solvent flattening steps.

Discussion

Electron microscopic or crystallographic structure determination of the bacteriophage T4 particle to atomic resolution is (at least presently) impossible due to its large size and asymmetry. In the past few years, however, progress has been made in structure determination of individual phage structural proteins, such as fibritin (Tao *et al*., 1997), a protein attached to the phage collar thought to be important in long tail fibre assembly and regulation of long tail fibre contraction. The structures of the base plate proteins gp9 (Kostyuchenko *et al*., 1999), the trigger for long tail fibre contraction, and gp11 (Leiman *et al*., 2000), the interface between the base plate and the short tail fibres, have also been determined. All three proteins contain an N-terminal triple α-helical coiled coil attached to C-terminal domains with β-structure. Circular dichroism spectroscopy has shown that the long (Selivanov *et al*., 1987) and short tail fibres (Burda *et al*., 2000) are rich in β-sheets and β-turns, but no structure has yet been determined. We therefore set out to determine the structure of the short tail fibre protein gp12.

Other authors have tried to crystallise full-length gp12 or fragments, but so far no crystals have been reported. For example, Miroshnikov *et al*. (1998) tried to fuse fragments of gp12 and of the long tail fibre protein gp37 to the trimerisation domain of T4 fibritin to obtain crystallisable material. The C-terminal trimerisation domain of fibritin consists of 30 residues from each subunit forming a β-propeller-like structure with a hydrophobic interior (Tao *et al*., 1997).

As we were not successful in obtaining crystals of full-length gp12, we have identified a heat- and protease-stable 33 kDa fragment of the bacteriophage T4 short fibre, starting at Leu-84 and most likely ending at Arg-395, although we cannot be sure about the C-terminus due to uncertainties in the mass spectrometric results. We can also not rule out that the protein is still non-covalently associated with small proteolytic peptides from outside the region spanning amino acids 84 to 395. The difficulties in obtaining accurate mass spectrometry results may be related to the presence of multiple proteolytic fragments close together in size. However, we could crystallise the resulting fragment and the resulting crystals diffracted X-rays very well. It is generally believed that a chemically and conformationally pure sample is necessary to obtain such high-quality crystals. Therefore we consider it more likely that the difficulties result from the intrinsic properties of the protein, leading the unfolded protein to aggregate even in the acidic denaturing conditions in the mass spectrometer. The adenovirus fibre (van Raaij *et al*., 1999), also a very stable trimeric fibrous protein, gives similar difficulties in mass spectroscopy (Jaquinod and van Raaij, unpublished observations). The crystals we obtained should allow the structure of the stable gp12 fragment to be solved.

Thermal unfolding of gp12 shows a small increase in ellipticity at 55 °C (Burda *et al*., 2000). It was suggested that this resulted from partial unfolding at the N-terminus. Our present results show differences in protease susceptibility at 37 °C if the protein is pre-incubated at 37 or 56 °C. The most likely explanation is partial unfolding at 56 °C, taking place at the C-terminus in addition to unfolding at the N-terminus. Jayaraman *et al*. (1997) showed that, in the absence of proteases or SDS, gp12 can refold back into a bacteriophage-reconstitution active conformation when heated to temperatures up to 75 °C for 5 minutes. In our experiments heating to 55 °C was sufficient to irreversibly render part of the N-terminus and C-terminus of the protein sensitive to proteolysis. This may be due to the prolonged incubation time at elevated temperature of 30 min, which would suggest a slow unfolding kinetic. It is also possible that a higher protein concentration as used in our experiments leads to the formation of protease-sensitive aggregates, whereas at lower protein concentrations gp12 can refold into non-aggregated functional trimers.

Proteolytic shortening experiments of the P22 tailspike have also been carried out at elevated temperature using thermostable proteases (Chen and King, 1991; Danner *et al.*, 1993). In these experiments, only N-terminally shortened protein chains were obtained. An N-terminally shortened fragment containing residues 106–666 was subsequently expressed and crystallised (Steinbacher *et al*., 1994). The N-terminal phage-binding domain was later also expressed and crystallised (Steinbacher *et al*., 1997). A similar strategy could be envisioned to obtain structural information regarding the N-terminal and C-terminal portions missing from the fragment we have described here.

Materials and Methods

Expression and Purification of gp12

For expression of the T4 short fibre protein gp12, the plasmid pT4g57g12 (Burda and Miller, 1999) was used, which encodes also the phage chaperone protein gp57 necessary for the correct folding of gp12. For each expression experiment, the plasmid was freshly transformed into *E. coli* strain JM109(DE3) (Promega, Charbonnières, France) and cultures (total volume 3 litre) were grown to an optical density of 0.6–1.0 at 600 nm. Cultures were then cooled to 22 °C and expression was induced with 1 mM isopropyl-β-D-thiogalactopyranoside. Expression was carried out for 12 h at 22 °C.

After harvesting by centrifugation, cells were resuspended in 100 ml TE-buffer [40 mM Tris-hydrochloric acid, 10 mM ethylenediamine tetraacetic acid, pH 8.0; Tris is 2-amino-2-(hydroxymethyl)-1,3-propanediol] containing 0.2 M ammonium sulphate and lysed using a French press. Insoluble material was removed by centrifugation; the soluble fraction was brought to 0.8 M ammonium sulphate and recentrifuged, removing some impurities. The gp12 protein was precipitated by adding ammonium sulphate to a final concentration of 1.4 M. The gp12 precipitate was redissolved in 50 ml TE-buffer containing 0.8 M ammonium sulphate and loaded onto a 75 ml phenyl Sepharose FF high sub column (Pharmacia, Orsay, France) equilibrated with the same buffer. Elution was with a gradient of 0.8 M to 0 M ammonium sulphate; the protein eluted at about 0.2 M. To the protein pool, ammonium sulphate was added (final concentration around 1.8 M) to precipitate gp12. Precipitated gp12 was redissolved in 15 ml TE-buffer containing 0.2 M ammonium sulphate. Up to here, all steps were performed at 4 °C, subsequent steps were performed at room temperature.

The protein was further purified using a Pharmacia Sephacryl S-200 gelfiltration column (16 mm i.d. × 60 cm); three runs were performed with 5 ml sample each. Fractions containing gp12 were pooled, brought to 0.7 M ammonium sulphate and loaded onto a phenyl superose 10/10 column (Pharmacia) equilibrated in TE-buffer containing 0.7 M ammonium sulphate. Elution was with a gradient of 0.7 M to 0.1 M ammonium sulphate; the protein eluted at about 0.5 M. Pure gp12 was precipitated with 1.8 M ammonium sulphate and stored at 4 °C. Protein concentrations were estimated using a specific absorption coefficient of 0.96 $cm^2\ mg^{-1}$ at 280 nm (Burda and Miller, 1999).

Small-Scale Proteolysis Experiments

Ammonium sulphate precipitated gp12 was re-dissolved in HN-buffer [10 mM N-(2-hydroxyethyl)piperazine-N'-(2-ethane)sulphonic acid (HEPES)-sodium hydroxide, 150 mM sodium chloride, pH 7.4] at 0.8 mg/ml. Trypsin and chymotrypsin stocks were at 10 mg/ml and 5 mg/ml, respectively, in 1 mM hydrochloric acid, stored at –20 °C. They were diluted with HN-buffer to 1 mg/ml just before use. Samples of 20 µl of gp12 solution were incubated at varying temperatures (37, 56 and 75 °C) for 20 min, cooled to 37 °C, after which 10 µl of a trypsin or chymotrypsin dilution was added. Incubation was for 20 min at 37 °C, the reaction was stopped by adding 20 µl gel loading buffer containing SDS and protease inhibitors (Complete™, Boehringer Mannheim, Germany).

Mass Spectrometry

Matrix-Assisted Laser Desorption-Ionisation Mass Spectrometry was carried out on a Voyager-Elite XL Biospectrometry workstation (Perspective BioSystems Inc., Framingham, MA, USA). As a matrix 2,5-dihydroxy-benzoic acid was used. Full-length and proteolysed gp12 were dissolved in a 50% (v/v) acetonitrile mixture containing 1% (v/v) trifluoroacetic acid to a final concentration of 5 µM. One µl of the samples was then mixed with 1 µl of saturated matrix solution in 50% (v/v) acetonitrile containing 0.1% trifluoroacetic acid and dried. Spectra were recorded from 256 laser shots (nitrogen laser, 337 nm) with an accelerating voltage of 25 000 volts in the linear mode. The other parameters were: a grid voltage of 94.7%, a guide wire voltage of 0.3% and a pulse source delay time of 300 ns. The instrument was calibrated using bovine myoglobin (Boehringer Mannheim, Germany).

N-Terminal Sequence Analysis

For N-terminal sequence analysis, proteins separated by SDS-PAGE were blotted onto poly-(vinylidene difluoride) membrane using standard procedures. The membrane was stained with Coomassie Brilliant Blue and rapidly destained with methanol. Bands containing the relevant protein were excised from the membrane, completely destained in 100% acetonitrile and immediately washed with water to remove any residual salt. The N-terminus of the protein fragment was sequenced with an Applied Biosystems (Foster City, USA) 494 Procise protein sequencer equipped with a blot cartridge, using the procedures recommended by the manufacturer.

Electron Microscopy

Protein samples at approximately 0.1 mg/ml in HND buffer (10 mM HEPES-sodium hydroxide, 150 mM sodium chloride, 10 mM dithiothreitol, pH 7.4) were applied to the clean side of carbon on mica (carbon/mica interface) and negatively stained with 1% (w/v) sodium silicotungstate (pH 7.0). Micrographs were taken under low-dose conditions with a JEOL 1200 EX II microscope at 100 kV and a nominal magnification of 40 000.

The three best images of each sample were digitised on an Optronics microdensitometer with a pixel size of 25 µm. At the level of the sample, the pixel size corresponds to 6.25 Å. We selected approximately 200 single images of fibre in each sample (native, '37 °C' and '56 °C') according to Crowther *et al*. (1996). The particles were cut out of the field in squares of 128 × 128 pixels and bi-pass filtered between 250 and 15 Å. Using multivariate statistical analysis and the classification method implemented in the program SPIDER (Frank *et al*., 1996), the particles from each condition were classified into ten different subgroups. All the images from each subgroup were added together to give a class average. All the class averages for each of the three samples were similar; the bending of the fibre was different but the length and the number of subunits were identical. One class average of each sample is shown in Figure 3 with a corresponding field of the electron micrographs.

Crystals of the heat- and protease-stable gp12 fragment (see below) were partially dissolved in HND-buffer and applied to carbon/mica grids as above. The best image was scanned, the better part of this image (256 × 256 pixels) was pad in a 512 × 512 image, Fourier transformed and used to calculate the power spectrum. With this power spectrum we were able to determine the cell parameters (249 × 51 Å). The maximum resolution for this image is approximately 25 Å (second order of diffraction along the long axis of the cell and order number ten along the small axis).

Preparation and Crystallisation of the 33 kDa Fragment

For large-scale proteolysis, purified and ammonium sulphate precipitated gp12 was dissolved in HND-buffer at 2–5 mg/ml and heated to 56 °C for 30 min. After cooling to 37 °C, trypsin was added (approximately 1 mg of trypsin per mg of gp12) and incubation was continued at 37 °C. A sample of the reaction mixture was analysed by SDS-PAGE. If the reaction mixture still contained appreciable amounts of the full-length protein (55 kDa) or of the 50 kDa fragment, the heat and trypsin treatment was repeated. If the main component was the 33 kDa fragment, it was purified by gel filtration as described for full-length fibre, precipitated with ammonium sulphate and stored at 4 °C.

The ammonium sulphate precipitated 33 kDa-fragment of gp12 was redissolved in HN-buffer and concentrated using Centricon Plus-20 devices with a nominal molecular weight cut-off of 10 kDa (Millipore, St. Quentin Yvelines, France). Residual am-

monium sulphate was removed by repeated dilution with HN-buffer and re-concentration using the same devices. Crystals of the gp12 fragment were obtained by the sitting drop vapour diffusion method at 20 °C in MVD/24 Crystal Growth Chambers (Charles Supper Co., Natick, USA) with reservoir solutions containing 100 mM sodium citrate at pH 5.6, 10% (v/v) glycerol and 15–25% (v/v) tertiary butanol (2-methylpropane-2-ol). Two µl of protein solution (10–40 mg/ml) were mixed with 2 µl of reservoir solution. Large (up to $0.2 \times 0.2 \times 0.5$ mm) single crystals were obtained overnight or after a few days to a few weeks of incubation, depending on protein and precipitant concentrations. For electron miscroscopy, crystals were partially dissolved in HND-buffer.

Acknowledgements

We are grateful to Stephen Cusack and Rob Ruigrok for their support of this work. M.v.R. and G.S. were supported by EMBL postdoctoral fellowships and M.v.R. also by the Dutch Royal Academy of Science.

References

Bishop, R.J., Conley, M.P., and Wood, W.B. (1974). Assembly and attachment of bacteriophage T4 fibers. J. Supramol. Struct. *2*, 196–201.

Burda, M.R., and Miller, S. (1999). Folding of coliphage T4 short tail fiber *in vitro*. Analysing the role of a bacteriophage-encoded chaperone. Eur. J. Biochem. *265*, 771–778.

Burda, M.R., Hindennach, I., and Miller, S. (2000). Stability of bacteriophage T4 short tail fiber. Biol. Chem. *381*, 255–258.

Cerritelli, M.E., Wall, J.S., Simon, M.N., Conway, J.F., and Steven A.C. (1996). Stoichiometry and domainal organization of the long tail-fiber of bacteriophage T4: a hinged viral adhesin. J. Mol. Biol. *260*, 767–780.

Chen, B., and King, J. (1991). Thermal unfolding pathway for the thermostable P22 tailspike endorhamnosidase. Biochemistry *30*, 6260–6269.

Crowther, R.A., Henderson R., and Smith J.M. (1996). MRC image processing programs. J. Struct. Biol. *116*, 9–16.

Danner, M., Fuchs, A., Miller, S., and Seckler, R. (1993). Folding and assembly of phage P22 tailspike endorhamnosidase lacking the N-terminal, head-binding domain. Eur. J. Biochem. *215*, 653–661.

Frank J., Radermacher M., Penczek P., Zhu J., Li Y., Ladjadj M., and Leith. A. (1996). SPIDER and WEB: processing and visualisation of images in 3D electron microscopy and related fields. J. Struct. Biol. *116*, 190–199.

Green, N.M., Wrigley, N.G., Russel, W.C., Martin, S.R., and McLachlan, A.D. (1983). Evidence for repeating cross-β sheet structure in the adenovirus fibre. EMBO J. *2*, 1357–1365.

Hashemolhosseini, S., Stierhof, Y.D., Hindennach, I., and Henning, U. (1996). Characterization of the helper proteins for the assembly of tail fibers of coliphages T4 and lambda. J. Bacteriol. *178*, 6258–6265.

Jayaraman, A., De Bernardez Clark, E., and Goldberg, E. (1997). Thermal unfolding of bacteriophage T4 short tail fibers. Biotechnol. Prog. *13*, 837–843.

Kells, S.S., and Haselkorn, R. (1974). Bacteriophage T4 short tail fibers are the product of gene 12. J. Mol. Biol. *83*, 473–485.

King, J., and Laemmli, U.K. (1971). Polypeptides of the tail fibres of bacteriophage T4. J. Mol. Biol. *62*, 465–477.

Kostyuchenko, V.A., Navruzbekov, G.A., Kurochkina, L.P., Strelkov, S.V., Mesyanzhinov, V.V., and Rossmann, M.G. (1999). The structure of bacteriophage T4 gene product 9: the trigger for tail contraction. Structure *7*, 1213–1222.

Leiman, P.G., Kostyuchenko, V.A., Shneider, M.M., Kurochkina, L.P., Mesyanzhinov, V.V., and Rossmann, M. G. (2000). Structure of bacteriophage T4 gene product 11, the interface between the baseplate and short tail fibers. J. Mol. Biol. *301*, 975–985.

Makhov, A.M., Trus, B.L., Conway, J.F., Simon, M.N., Zurabishvili, T.G., Mesyanzhinov, V.V., and Steven, A.C. (1993). The short tail-fiber of bacteriophage T4: molecular structure and a mechanism for its conformational transition. Virology *194*, 117–127.

Mason, W.S., and Haselkorn, R. (1972). Product of T4 gene 12. J. Mol. Biol. *66*, 445–469.

Miroshnikov, K.A., Marusich, E.I., Cerritelli, M.E., Cheng, N., Hyde, C.C., Steven, A.C., and Mesyanzhinov, V.V. (1998). Engineering trimeric fibrous proteins based on bacteriophage T4 adhesins. Protein Eng. *11*, 329–332

Mitraki, A., Barge, A., Chroboczek, J., Andrieu, J.P., Gagnon, J., and Ruigrok, R.W.H. (1999). Unfolding studies of human adenovirus type 2 fibre trimers. Evidence for a stable domain. Eur. J. Biochem. *264*, 599–606.

Riede, I. (1987). Receptor specificity of the short tail fibres (gp12) of T-even type *Escherichia coli* phages. Mol. Gen. Genet. *206*, 110–115.

Selivanov, N.A., Venyaminov S.Y., Golitsina, N.L., Nikolaeva L.I., and Mesyanzhinov, V.V. (1987). Structure and regulation of the assembly of bacteriophage T4 fibrillar proteins. I. Isolation and spectral properties of long fibers. Mol. Biol. (Moscow) *21*, 1258–1267.

Steinbacher, S., Seckler, R., Miller, S., Steipe, B., Huber, R., and Reinemer, P. (1994). Crystal structure of P22 tailspike protein: interdigitated subunits in a thermostable trimer. Science *265*, 383–386.

Steinbacher, S., Miller, S., Baxa, U., Budisa, N., Weintraub, A., Seckler, R., and Huber, R. (1997). Phage P22 tailspike protein: crystal structure of the head-binding domain at 2.3 Å, fully refined structure of the endorhamnosidase at 1.56 Å resolution, and the molecular basis of O-antigen recognition and cleavage. J. Mol. Biol. *267*, 865–880.

Stouten, P.F., Sander, C., Ruigrok, R.W.H., and Cusack, S. (1992). New triple-helical model for the shaft of the adenovirus fibre. J. Mol. Biol. *226*, 1073–1084.

Tao, Y., Strelkov, S.V., Mesyanzhinov, V.V., and Rossmann, M.G. (1997). Structure of bacteriophage T4 fibritin: a segmented coiled coil and the role of the C-terminal domain. Structure *5*, 789–798.

van Raaij, M.J., Mitraki, A., Lavigne, G., and Cusack, S. (1999). A triple beta-spiral in the adenovirus fibre shaft reveals a new structural motif for a fibrous protein. Nature *401*, 935–938.

Ward, S., and Dickson, R.C. (1971). Assembly of bacteriophage T4 tail fibers. 3. genetic control of the major tail fiber polypeptides. J. Mol. Biol. *62*, 479–492.

Received January 19, 2001; accepted April 2, 2001