Structure of bacteriophage $\phi 29$ head fibers has a supercoiled triple repeating helix-turn-helix motif

Ye Xiang and Michael G. Rossmann¹

Department of Biological Sciences, Purdue University, 240 S. Martin Jischke Drive, West Lafayette, IN 47907-2032

Edited by Wayne A. Hendrickson, Columbia University, New York, NY, and approved February 8, 2011 (received for review December 6, 2010)

The tailed bacteriophage $\phi 29$ capsid is decorated with 55 fibers attached to quasi-3-fold symmetry positions. Each fiber is a homotrimer of gene product 8.5 (gp8.5) and consists of two major structural parts, a pseudohexagonal base and a protruding fibrous portion that is about 110 Å in length. The crystal structure of the C-terminal fibrous portion (residues 112–280) has been determined to a resolution of 1.6 Å. The structure is about 150 Å long and shows three distinct structural domains designated as head, neck, and stem. The stem region is a unique three-stranded helix-turn-helix supercoil that has not previously been described. When fitted into a cryoelectron microscope reconstruction of the virus, the head structure corresponded to a disconnected density at the distal end of the fiber and the neck structure was located in weak density connecting it to the fiber. Thin section studies of *Bacillus subtilis* cells infected with fibered or fiberless $\phi 29$ suggest that the fibers might enhance the attachment of the virions onto the host cell wall.

infection | $\phi 29$ | supersecondary structure | X-ray crystallography

Bacteriophages are viruses that infect bacteria and are probably the most abundant organisms on Earth (1). More than 96% of bacteriophages are tailed bacteriophages that consist of either an isometric or a prolate head with an attached tail (2). The head capsids of bacteriophages are tough shells with complex elastic properties (3). The major capsid protein is the primary unit from which the head capsid is assembled and in most tail phages adopts the “HK97” fold, named for the major capsid protein of phage Hong Kong 97 (4–6). Other than the major capsid protein, phage capsids are often decorated with different types of minor proteins. One type of minor protein that is frequently found in many phages is an immunoglobulin-like protein (7). Although many minor capsid proteins are not essential for infection in laboratory conditions (8–11), they may have provided favorable properties for the survival of phage in hostile environments. For instance, the minor capsid protein Soc of phage T4 provides enhanced capsid stability (9).

Bacteriophage $\phi 29$ infects the Gram-positive bacteria *Bacillus subtilis* using a short, 380 Å-long, noncontractile tail (12, 13). The *B. subtilis* cell wall teichoic acid is the primary receptor recognized by the $\phi 29$ tail appendages for initiating infection (14). After recognition of the host, the tail is brought closer to the peptidoglycan by digesting the cell wall teichoic acids. With the aid of gene product 13 (gp13) (15, 16), the phage tail can penetrate the cell wall peptidoglycan layer and cell membrane to release the genomic DNA that had been packaged in the prolate capsid. The mature $\phi 29$ capsid consists of two proteins: the major capsid protein gp8 and the head fiber protein gene product 8.5 (gp8.5), assembling into a $T = 3$, $Q = 5$ icosahedral shell elongated along a 5-fold axis of symmetry (Fig. 1) (5, 17). The major capsid protein gp8 has two domains that resemble the HK97 capsid protein and the group 2 bacterial immunoglobulin-like (BIG2) protein. The HK97-like domains of $\phi 29$ gp8 pack tightly against one another in the capsid assembly. The BIG2 domains protrude from the surface of the capsid and provide the attachment sites for the head fibers at quasi-3-fold symmetry positions. Cryoelectron microscopy (cryoEM) maps indicate that

Fig. 1. The fibered structure of phage $\phi 29$. (Left) A 25 Å resolution cryoelectron microscope reconstruction of $\phi 29$. The fibers are colored in gold. Note the lower density connecting the distal ends of the fibers. (Right) Diagrammatic representation of the icosahedral and quasi-symmetry elements ($T = 3$, $Q = 5$) of the phage head. The fibers attach only to quasi-3-fold axes (green dots). The icosahedral 5-fold and 3-fold axes are represented by black and red dots, respectively. The 6-fold symmetry axes in the cylindrical part of the prolate head are represented by orange dots.

the head fibers possess two distinct structural domains: a pseudohexagonal base that attaches to the capsid and a protruding stem that is about 110 Å long. Additional unconnected weak densities at the distal ends of the fibers may be part of the fibers or noise (5, 17). The fibers are not essential for phage infection in laboratory conditions (10, 11).

Here we report the crystal structure of the gp8.5 C-terminal fibrous portion, which has a unique type of supersecondary structure, helix-turn-helix supercoil. The fibers might enhance the attachment of phage $\phi 29$ to the host cells during infection.

Results and Discussion

The gp8.5 Structure. A construct containing the full length gp8.5 and a C-terminal His₆ tag was expressed in *Escherichia coli* cells. The expressed gp8.5 was cleaved in vitro with chymotrypsin. A stable fragment was obtained and purified. Western blotting showed that the fragment has the His₆ tag and, therefore, was identified as being a C-terminal fragment.

Author contributions: Y.X. and M.G.R. designed research; Y.X. performed research; Y.X. and M.G.R. analyzed data; and Y.X. and M.G.R. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The $\phi 29$ gp8.5 coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 3QC7).

¹To whom correspondence should be addressed. E-mail: mr@purdue.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1018097108/-DCSupplemental.

compared to the rest of the structure (for these the equivalent value is 18 Å²), indicating that it is flexible. Fitting of the crystal structure into the cryoEM map of the virus showed that the stem domain can be fitted into the extended density of the head fiber (Fig. S2), placing the head into the disconnected densities at the distal end and the neck into the weak connecting densities. Thus, the disconnected densities in the cryoEM reconstruction were identified as being a part of the gp8.5 structure. The cryoEM densities corresponding to the head and neck parts of gp8.5 are weak, probably as a result of averaging the flexible head and neck structures. Dali searches (http://ekhidna.biocenter.helsinki.fi/dali_server/) did not find any significant similar structure to gp8.5.

A Unique Super Helix-Turn-Helix Coiled Coil. The stem domain of the structure contains short α -helices, each with two to three turns. In each gp8.5 monomer the helices are connected with “ β -turns.” Although, as in all other defined β -turn structures, the β -turns in the stem domain change the direction of the main chain by approximately 180° , they do not correspond to any previously defined β -turn groups (18) (Table S1). Each β -turn raises the fol-

lowing α -helix by about 10 Å. Of the nine helices in each monomer, three helices at the proximal end of the fiber are almost antiparallel to each other. The remaining six helices make a right-handed super helix. The basic repeat unit in the super helix is a helix-turn-helix (Fig. 2B and Table 1). Each repeat unit corresponds to a 45° turn and a translation of 19.5 Å along the super helical axis. There would be eight helix-turn-helix repeats per turn with a pitch of 156 Å and an external radius of 24 Å if the helix were extended.

The three super helices in the trimer are coiled around the central vertical 3-fold axis to form a super helical coiled coil (Fig. 2A). The core of the coiled coil is made of hydrophobic side chains from horizontal helices that belong to different molecules (Fig. 2C). The vertical interactions in the coiled coil are mainly hydrogen bonds between side chain atoms of the antiparallel helices within the same monomer. Sequence analysis shows that the super coiled coil contains the repeat motif HPPXPHXXH-HPXPXPPPPXHXHPPHH, where H represents a hydrophobic residue, P represents a polar residue and X represents any residue (Fig. 2D). A search for the presence of this motif in sequence databases showed that similar sequences exist in many other proteins. However, none of these proteins contains sequential repeats of this motif except for the phage ϕ 29 fiber-like proteins.

The α -helix is a basic secondary structural element of proteins. Helix-helix interactions are ubiquitous in the structure of proteins and protein assemblies. The most common supersecondary structure built on helix-helix interactions is the coiled coil. Coiled coils have a hydrophobic core and are characterized with a heptad sequence repeat. The helix-turn-helix motif is another type of helical supersecondary structure that has been frequently found in DNA-binding proteins (19). The gp8.5 super helix coiled coil is a combination of the coiled coil and the helix-turn-helix supersecondary structures (Fig. 3). Although the exact function of this supersecondary structure is not clear, it could provide information for protein structure prediction and protein design as well as the folding mechanism of fibrous proteins.

Structural Comparisons with Other Fiber Structures in Viruses. Many viruses encode proteins that often form homotrimeric fibrous structures of various folds. A few of these proteins, such as the T4 fibrin (20) and the P22 tail needle (21), fold into α -helical fibrous structures that each has an α -helix coiled coil fold (Fig. 4A). However, most of these viral fiber molecules consist entirely, or nearly entirely, of β -structures and usually function as host-cell recognition molecules (22). Among the viral fibrous β -structures, phage tail spike proteins contain a “ β -helix” fold and are probably the most ubiquitous viral fibers (Fig. 4A) (14). The bacteriophage T4 tail needle fiber (23), the T4 long tail fiber (24), and the adenovirus fiber (25) structures each represent a different β -structured fold (Fig. 4A). Each monomer of the adenovirus triple β -spiral structure contains eight short antiparallel β -strands of which two adjacent β -strands are a repeating unit that makes up a left-handed spiral. Although the secondary

Fig. 3. Model of a trimeric super helix-turn-helix coiled coil. (Top) The monomers are colored red, green, and blue with black dots on the red monomer showing the positions of the helix-turn-helix repeating motifs. (Middle) One repeat motif. (Bottom) Diagrammatic representation of the coiled coil structure showing individual helices.

structural components of gp8.5 are quite different from that of the adenovirus fiber, the overall organization of the gp8.5 α -helices is surprisingly similar to that of the adenovirus fiber antiparallel β -strands (Fig. 4B). Gp8.5 uses a pair of α -helices whereas the adenovirus fiber uses a pair of β -strands as a repeating unit to make a super helix. However, these super helices have opposite hands and slightly different helical parameters. The adenovirus triple β -spiral structure repeats by translating a β -hairpin motif along the 3-fold axis by about 13 Å and rotating it clockwise by about 50°. In contrast, the α -helical gp8.5 fibrous structure has a displacement of about 20 Å between two neighboring α -helical repeats and an anticlockwise rotation of about 45° (Fig. 2).

Possible Role of the Fibers During Phage Infection. The minor surface decorating protein Soc of bacteriophage T4 enhances the stability of the capsid. Thus, possibly, gp8.5 might have a similar function for bacteriophage ϕ 29. This possibility was investigated by treating fibered and fiberless ϕ 29 particles each with buffers at pH 3, pH 10.0, or in 3 M guanidinium chloride. The treated particles were then inspected using cryoEM. The results showed that almost all the capsids were still intact. However, most of them had lost their packaged DNA (Fig. S3). Furthermore, the treated particles produced fewer plaques compared to untreated particles in a standardized plaque assay (Fig. S3). Both types of particles were affected about equally by the above conditions. Thus, gp8.5 does not appear to be required for stabilizing a ϕ 29 capsid.

Crystallization of the fiber proteins showed that phosphate was essential for obtaining well-diffracting crystals, suggesting that the protein may have weak interactions with the phosphate buffer. The major component of the *Bacillus subtilis* cell wall teichoic acid is the glycosylated polyglycerol phosphate. Thus the head fibers may also have weak interactions with the cell teichoic acids and function as an enhancer for the attachment of the phage to the cell wall. To further investigate the role of the fibers during infection, fibered or fiberless ϕ 29 infected cells were frozen under high pressure 30 min after infection. The frozen samples were then slowly substituted with resin. The fixed cells were sliced and examined with an electron microscope. The results showed that most of the fibered particles were well aligned on the cell wall, whereas the fiberless particles attached to the cell wall in a roughly random fashion (Fig. 5 and Fig. S4). Although the infectivity of the fiberless particles had apparently not been affected, more of the particles with fibers recognize the cell successfully or can reach the next stage of infection more efficiently. The difference between the fibered and fiberless particles may have a great impact on the evolution of the virus when the concentration of the virus is low.

Materials and Methods

Protein Expression and Purification. Gene 8.5 was amplified by PCR from the genome DNA of bacteriophage ϕ 29 using the following primers, 5'-GGTGGTTCATATGATGGTTTCATTACT-3' and 5'-GGTGGTCTCGAGTGATGATTCCTCGCTT-3'. Preparation of the genome DNA has been reported previously (26). Purified PCR products were cloned into pET30b (Novagen) using the NcoI-XhoI sites that introduce a C-terminal His₆ tag into the recombinant

Table 1. Super secondary structure helical parameters

	U2*	U3
U1	trans: 19.6 Å	trans: 38.8 Å
α 4- α 5	rot: 48°	rot: 90°
Residues 149–174	α rmsd: 0.28 Å	α rmsd: 0.63 Å
	sequence identity: 40%	sequence identity: 38%
U2		trans: 19.4 Å
α 6- α 7		rot: 43°
Residues 175–200		α rmsd: 0.60 Å
		sequence identity: 28%
U3		
α 8- α 9		
Residues 201–226		

*See Fig. 2B for the definitions of U1, U2, and U3.

Fig. 4. Trimeric virus fiber structures. (A) A comparison of different virus fiber structure folds. (B) Stereographic detail structure comparison of a fibrous β -structure (Left) and a fibrous α -structure (Right).

protein. The recombinant 8.5 gene was expressed in *E. coli* BL21(DE3) Codon-Plus-RIL cells at 20 °C. The produced recombinant protein was affinity-purified using cobalt charged BD TALON™ resins and was eluted from the cobalt beads by using an elution buffer containing 50 mM sodium phosphate at pH 8.0, 300 mM sodium chloride and 200 mM imidazole. The eluted protein was concentrated to approximately 3 mg/mL and then digested overnight at 37 °C with chymotrypsin (final concentration 0.1 mg/mL). The digested gp8.5 was further purified with a Superdex 200 column (GE) and was eluted with a buffer containing 50 mM Tris at pH 8.0 and 100 mM sodium chloride. Size exclusion chromatography purification of the gp8.5 indicated that it exists mainly as a multimer, most probably a trimer, in solution.

Phage Production and Thin Section Sample Preparation. Fibered and fiberless ϕ 29 particles were produced in *Bacillus subtilis* su44⁺ or SpoOA12(sup⁻¹) cells

infected with the mutants *sus16(300)–sus14(1241)* or *sus14(1241)–sus8.5(900)*, respectively, and purified by centrifugation in an isopycnic 60% (w/v) CsCl gradient. The effect of pH or guanidinium chloride on ϕ 29 infection was tested by using standard plaque assays. Prior to phage infection, the phage particles were variously incubated for approximately 20 min in buffers at pH 3.0, pH 8.0, pH 10.0, or in 3 M guanidinium chloride solution. The treated phage particles were also examined by using a CM300 transmission electron microscope.

To prepare thin sectioned samples of infected cells, *Bacillus subtilis* su44⁺ cells were grown in 5 mL LB media to an OD₆₀₀ value of approximately 1.2. The cells were centrifuged at 200 $\times$ g for 3 min and then were resuspended in 1 mL buffer containing 50 mM Tris HCl pH 8.0, 100 mM sodium chloride, and 10 mM magnesium chloride. The resuspended cells were split into two equal portions infected with fibered or fiberless ϕ 29 particles,

Fig. 5. Transmission electron micrographs of thin sections showing a *B. subtilis* bacterial cell surrounded by ϕ 29 phages. (Left) Fibered phages that have arranged themselves in ordered arrays on the bacterial surface. (Right) Fiberless phage that associate with the bacterial surface in a random manner.

respectively, at a multiplicity of infection of 100. About 30 min after the infection, the infected cells were flash-frozen at a high pressure and were fixed at -90°C in acetone containing 2% (w/w) O_3Q_4 and 1% (w/w) uranyl acetate. The fixed samples were slowly substituted over 80 h with resins and then sliced into approximately 100 nm thin sections.

Crystallization. All crystals were obtained by hanging-drop vapor diffusion at 20°C , using 3 μL protein (15 mg/mL) mixed with an equal volume of well solution. Crystals of the gp8.5 C-terminal fragment were grown in 100 mM sodium phosphate at pH approximately 6.8 and 8% w/v PEG 3000. It took about one week for crystals to appear. Crystals were soaked for 30 s in the well solution containing a final concentration of 15% w/v

PEG 3000 and 15% v/v MPD (2-methyl-2,4-pentanediol) to flash-freeze in a N_2 stream at 100 K. The iodine derivative of the gp8.5 C-terminal fragment was prepared by soaking the crystals in the cryowell solution containing 1 M sodium iodine for 1 min.

X-ray Data Collection, Processing, Structure Determination, Refinement, and Analysis. X-ray diffraction data were collected using either synchrotron radiation or home source $\text{Cu K}\alpha$ radiation (Table S2). Crystals of the gp8.5 C-terminal fragment diffracted to $1.6\text{ \AA}$ and belong to space group $R32$ with one molecule in the asymmetric unit and cell parameters of $a = b = 47\text{ \AA}$ and $c = 425\text{ \AA}$. All data were integrated and scaled with the HKL2000 suite (Table S2) (27).

The structure of the gp8.5 C-terminal fragment was determined by SIRAS using the anomalous signals of iodine atoms measured at the $\text{Cu K}\alpha$ wavelength. Heavy atom sites were located by using the program SHELX (28). The heavy atom parameters were refined and initial phases were calculated using reflections in the resolution region of $50\text{ \AA}$ to $2.2\text{ \AA}$ with the program SHARP (29, 30). The calculated phases were gradually improved and extended to higher resolution by using solvent density flattening with the program DM (31). The resultant electron density maps are of good quality in which most of the residues can be clearly recognized. See [SI Text](#) for further model building procedures.

ACKNOWLEDGMENTS. We thank Siyang Sun, Petr Leiman, and Daisuke Kihara for stimulating discussions. We thank Dwight Anderson, Shelley Grimes, and Paul Jardine for their early support on this project. We thank Sheryl Kelly for help in preparing the manuscript. We thank the staff at Argonne National Laboratory's Advanced Photon Source (APS) (GM/CA-CAT) for their help. The APS beamline facility is supported by the US Department of Energy and the National Institutes of Health. This work was supported by a National Science Foundation Grant MCB-1014547 (M.G.R.).

- Wommack KE, Colwell RR (2000) Virioplankton: Viruses in aquatic ecosystems. *Microbiol Mol Biol R* 64:69–114.
- Ackermann HW (2003) Bacteriophage observations and evolution. *Res Microbiol* 154:245–251.
- Ivanovska IL, et al. (2004) Bacteriophage capsids: Tough nanoshells with complex elastic properties. *Proc Natl Acad Sci USA* 101:7600–7605.
- Wikoff WR, et al. (2000) Topologically linked protein rings in the bacteriophage HK97 capsid. *Science* 289:2129–2133.
- Morais MC, et al. (2005) Conservation of the capsid structure in tailed dsDNA bacteriophages: The pseudoatomic structure of ϕ 29. *Mol Cell* 18:149–159.
- Fokine A, et al. (2005) Structural and functional similarities between the capsid proteins of bacteriophages T4 and HK97 point to a common ancestry. *Proc Natl Acad Sci USA* 102:7163–7168.
- Fraser JS, Maxwell KL, Davidson AR (2007) Immunoglobulin-like domains on bacteriophage: Weapons of modest damage? *Curr Opin Microbiol* 10:382–387.
- Choi KH, et al. (2008) Insight into DNA and protein transport in double-stranded DNA viruses: the structure of bacteriophage N4. *J Mol Biol* 378:726–736.
- Qin L, Fokine A, O'Donnell E, Rao VB, Rossmann MG (2010) Structure of the small outer capsid protein, Soc: A clamp for stabilizing capsids of T4-like phages. *J Mol Biol* 395:728–741.
- Reilly BE, Nelson RA, Anderson DL (1977) Morphogenesis of bacteriophage ϕ 29 of *Bacillus subtilis*: mapping and functional analysis of the head fiber gene. *J Virol* 24:363–377.
- Salas M, Vasquez C, Mendez E, Vinuela E (1972) Head fibers of bacteriophage ϕ 29. *Virology* 50:180–188.
- Anderson D, Reilly B (1993) Morphogenesis of bacteriophage ϕ 29. *Bacillus subtilis and Other Gram-Positive Bacteria: Biochemistry, Physiology, and Molecular Genetics*, eds AL Sonenshein, JA Hoch, and R Losick (American Society for Microbiology, Washington, DC), pp 859–867.
- Xiang Y, et al. (2006) Structural changes of bacteriophage ϕ 29 upon DNA packaging and release. *EMBO J* 25:5229–5239.
- Xiang Y, et al. (2009) Crystallographic insights into the autocatalytic assembly mechanism of a bacteriophage tail spike. *Mol Cell* 34:375–386.
- Xiang Y, et al. (2008) Crystal and cryoEM structural studies of a cell wall degrading enzyme in the bacteriophage ϕ 29 tail. *Proc Natl Acad Sci USA* 105:9552–9557.
- Cohen DN, Erickson SE, Xiang Y, Rossmann MG, Anderson DL (2008) Multifunctional roles of a bacteriophage ϕ 29 morphogenetic factor in assembly and infection. *J Mol Biol* 378:804–817.
- Tao Y, et al. (1998) Assembly of a tailed bacterial virus and its genome release studied in three dimensions. *Cell* 95:431–437.
- Hutchinson EG, Thornton JM (1994) A revised set of potentials for β -turn formation in proteins. *Protein Sci* 3:2207–2216.
- Brennan RG, Matthews BW (1989) The helix-turn-helix DNA binding motif. *J Biol Chem* 264:1903–1906.
- Tao Y, Strelkov SV, Mesyanzhinov VV, Rossmann MG (1997) Structure of bacteriophage T4 fibrin: A segmented coiled coil and the role of the C-terminal domain. *Structure* 5:789–798.
- Olia AS, Casjens S, Cingolani G (2007) Structure of phage P22 cell envelope-penetrating needle. *Nat Struct Mol Biol* 14:1221–1226.
- Mitraki A, Papanikolopoulou K, van Raaij MJ (2006) Natural triple β -stranded fibrous folds. *Adv Protein Chem* 73:97–124.
- Kanamaru S, et al. (2002) Structure of the cell-puncturing device of bacteriophage T4. *Nature* 415:553–557.
- Bartual SG, et al. (2010) Structure of the bacteriophage T4 long tail fiber receptor-binding tip. *Proc Natl Acad Sci USA* 107:20287–20292.
- van Raaij MJ, Mitraki A, Lavigne G, Cusack S (1999) A triple β -spiral in the adenovirus fibre shaft reveals a new structural motif for a fibrous protein. *Nature* 401:935–938.
- Anderson DL, Hickman DD, Reilly BE (1966) Structure of *Bacillus subtilis* bacteriophage ϕ 29 and the length of ϕ 29 deoxyribonucleic acid. *J Bacteriol* 91:2081–2089.
- Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Method Enzymol* 276:307–326.
- Sheldrick GM (2008) A short history of SHELX. *Acta Crystallogr A* 64:112–122.
- Bricogne G, Vonrhein C, Flensburg C, Schiltz M, Paciorek W (2003) Generation, representation and flow of phase information in structure determination: Recent developments in and around SHARP 2.0. *Acta Crystallogr D* 59:2023–2030.
- de La Fortelle E, Bricogne G (1997) Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Method Enzymol* 276:472–494.
- Cowtan KD, Zhang KYJ, Main P (2000) DM/DMMULTI: Software for phase improvement by density modification. *International Tables for Crystallography*, eds MG Rossmann and E Arnold (Kluwer, Dordrecht), Vol F, pp 507–510.