Journal of Structural Biology **135**, 18–25 (2001)
doi:10.1006/jsbi.2001.4375, available online at http://www.idealibrary.com on IDEAL®

Composition and Mass of the Bacteriophage ϕ29 Prohead and Virion

Charlene Peterson,* Martha Simon,† James Hodges,‡ Peter Mertens,§ LeeAnn Higgins,¶ Edward Egelman,‖ and Dwight Anderson*,**,[1]

*Department of Oral Science, **Department of Microbiology, ¶Department of Biochemistry, and ‡Division of Biostatistics, University of Minnesota, Minneapolis, Minnesota 55455; †Biology Department, Brookhaven National Laboratory, Upton, New York 11973; §Packard BioScience B.V., Groningen, The Netherlands; and ‖Department of Biochemistry, University of Virginia Health Science Center, Charlottesville, Virginia 22908

Received February 23, 2001, and in revised form May 9, 2001; published online July 25, 2001

The protein composition of the *Bacillus subtilis* bacteriophage ϕ29 prohead and virion was determined by combustion of gel bands of ^{3}H-labeled proteins. Copy numbers of individual proteins were calculated relative to the 12 copies of the head–tail connector protein. The mean numbers of copies of the major capsid protein in the prohead and virion were 241 and 218, respectively, approaching the 235 copies determined previously by cryoelectron microscopy. The mean numbers of copies of the dimeric head fiber on the prohead and virion were 24 and 31, respectively, demonstrating partial occupancy of the 55 fiber binding sites. Measured copies of neck and tail proteins in the virion included 11 of the lower collar, 58 of the appendage, and 9 of the tail; if the true copies of these proteins are 12, 60, and 9, respectively, the entire neck and tail of ϕ29 has quasi-sixfold symmetry. The mass of the fiberless prohead with pRNA was about 14.2 MDa, and the mass of the prohead determined by scanning transmission electron microscopy was consistent with the biochemical data. The mass of the fiberless virion containing the 12.8-MDa DNA genome was about 30.4 MDa. A full complement of dimeric fibers on the prohead or virion would increase the mass of the particle by about 3.2 MDa. The data complement studies relating the structure of ϕ29 components to dynamic functions in morphogenesis and infection. © 2001 Academic Press

Key Words: **bacteriophage ϕ29 structure; ϕ29 composition; ϕ29 mass.**

INTRODUCTION

Knowledge of the stoichiometry of the proteins comprising viruses is integral to structure–function analyses of morphogenesis and infection. Bacteriophage ϕ29 of *Bacillus subtilis* is the simplest of the well-studied double-stranded DNA phages in terms of size and composition, and its morphogenesis is being studied with integrated genetic, biochemical, and biophysical approaches. Recently, the structure and assembly of ϕ29 have been studied in three dimensions by cryoelectron microscopy (cryo-EM) and image reconstruction (Tao *et al.*, 1998). This work demonstrated that both the prolate prohead and the mature head of ϕ29 contain 235 copies of the major shell protein in the $T = 3$, prolate icosahedron. In addition, the structure of the ϕ29 DNA packaging motor, which includes the head–tail connector, prohead RNA (pRNA), and the ATPase gp16, has been determined by X-ray crystallography and cryo-EM (Simpson *et al.*, 2000).

Infection of *B. subtilis* with suppressor-sensitive mutants, where translation is terminated at a nonsense codon, permits the isolation of intermediates on the assembly pathway, including the precursor capsid or prohead (Bjornsti *et al.*, 1981). The prohead is composed of the scaffolding protein (gene product 7, gp7), the major capsid protein (gp8), the head fibers (gp8.5), the head–tail connector (gp10), and the prohead RNA (pRNA) (Guo *et al.*, 1987). The prohead efficiently packages the DNA-gene product 3 complex (DNA-gp3) in a defined *in vitro* system (Guo *et al.*, 1986), and the DNA-filled heads can be assembled to infectious phage in extracts containing the neck lower collar (gp11) and appendages (gp12*, cleavage product of gp12) and the tail (gp9) (Bjornsti *et al.*, 1981; Guo *et al.*, 1987). Alternatively, lower collar, appendage, and tail proteins produced from cloned genes in *Escherichia coli* can be assembled onto the DNA-filled head to yield infectious ϕ29 (Lee and Guo, 1994). The DNA-filled head and intermediates in neck and tail assembly can also be obtained from cells infected with suppressor-sensitive mu-

[1] To whom correspondence should be addressed at the University of Minnesota, 18-242 Moos Tower, 515 Delaware Street S.E., Minneapolis, MN 55455. Fax: (612) 625-1108. E-mail: dlander@tc.umn.edu.

1047-8477/01 $35.00

tants and studied for composition and structure (Hagen *et al.*, 1976). However, the composition and symmetry of the ϕ29 neck and tail have not been known with certainty, such as the number of subunits in each of 12 neck appendages that serve as adsorption organelles of ϕ29 (Tosi *et al.*, 1975).

Here we determine the protein composition of the purified ϕ29 prohead and virion by labeling with [^{3}H]leucine, separation of proteins by SDS–gel electrophoresis, and combustion of excised gel fractions for quantification of radiolabel. The biochemical data on composition of the ϕ29 head are consistent with the $T = 3$ quasi-symmetry determined by cryo-EM (Tao *et al.*, 1998) and with the mass of the prohead determined here by scanning transmission electron microscopy (STEM). The number of copies of ϕ29 structural proteins differs substantially from those estimated previously (Carrascosa *et al.*, 1981); in particular, copies of the lower collar, appendage, and tail are at least twice the earlier estimate. The composition data will complement an ongoing study of the structure of the neck and tail of ϕ29 by cryo-EM and X-ray crystallography.

MATERIALS AND METHODS

Preparation of radiolabeled proheads and phage. To prepare proheads, *B. subtilis* DB431 (RD2) was grown to 4×10^8 cells/ml in 416 medium [2% (w/v) Difco Bacto-tryptone, 1% (w/v) Difco yeast extract, and 0.17 M NaCl], diluted 5-fold into the minimal medium M40 (Polsinelli and Beretta, 1966) supplemented with 50 μg/ml of 20 amino acids (M40s medium) and grown again to 4×10^8 cells/ml. Cells were pelleted and concentrated to 2×10^9 cells/ml in M40s medium with reduced leucine at 9 μg/ml. To produce fibered and fiberless proheads, cells were infected with *sus*16(300)–*sus*14(1241) or *sus*8.5(900)–*sus*16(300)–*sus*14(1241) phage, respectively, at a multiplicity of infection of 15. After 5 min at 37°C, the infected cells were diluted 10-fold into M40s medium with 9 μg/ml leucine and incubated at 37°C. At 10 min [^{3}H]leucine (Amersham Pharmacia Biotech) was added to 5 μCi/ml, and incorporation of radiolabel, measured as trichloroacetic acid-precipitable counts, was linear over the period of infection. The lysate of cells infected with *sus*8.5(900)–*sus*16(300)–*sus*14(1241) phage was clarified by centrifugation, particles were pelleted in the Beckman 50.2 rotor at 35 000 rpm for 3.5 h at 4°C, and the pellet was resuspended in 1/150 volume of TMS buffer (50 mM Tris–HCl, pH 7.5, 10 mM $MgCl_2$, and 100 mM NaCl). Cells infected with *sus*16(300)–*sus*14(1241) phage were pelleted at 65 min after infection and lysed by resuspension in 1/150 volume of 2× Difco antibiotic medium No. 3 (Penassay broth) containing SMM/BSA [0.15 M sucrose, 20 mM $MgCl_2$, 20 mM maleic acid–NaOH (pH 6.5), and 1% bovine serum albumin], 100 mg/ml lysozyme, 10 μg/ml DNase I (Calbiochem), 40 units/ml RNase inhibitor (human placental, Calbiochem), and 2.5% (w/v) chloroform at 37°C. The lysate was vortexed for 1 min and clarified by centrifugation, and the labeled proheads were isolated by centrifugation in a 17–30% (w/v) linear sucrose density gradient containing TMS buffer in the Beckman SW28 rotor at 16 500 rpm for 15 h at 4°C. The particles were pelleted as described above and resuspended in TMS buffer.

Fibered and fiberless phage were produced in *B. subtilis* RD2 cells grown and labeled with [^{3}H]leucine as described above but infected with *sus*14(1241) or *sus*8.5(900)–*sus*14(1241) phage, respectively. Infected cells were lysed after 2.5 h of infection by the addition of lysozyme to 20 μg/ml and DNase I to 5 μg/ml. The lysate was clarified by centrifugation, the phage was pelleted and resuspended in TMS buffer, and the radiolabeled particles were purified by isopycnic centrifugation in CsCl as described (Anderson *et al.*, 1966).

TABLE I
Leucine Content of ϕ29 Proteins

Gene product (gp)	Molecular weight	No. of leucine residues[a]	Leu/Da ($\times 10^{-4}$)
gp7	11 267	13	11.538
gp8	49 847	31	6.219
gp10	35 880	23	6.410
gp8.5	29 491	25	8.477
gp11	33 801	21	6.213
gp12*[b]	73 888	48	6.500
gp9	67 594	41	6.066
gp3	31 051	16	5.150

[a] Vlcek and Paces (1986). For gp12* the number of leucine residues was determined from its mass and N-terminal sequence analysis to show that it is cleaved at the C-terminus.
[b] Cleavage product of the 92.1-kDa gp12; molecular mass determined by MALDI-TOF mass spectrometry (Fig. 2).

Protein separation and quantification. The leucine content of the ϕ29 proteins inferred from the nucleotide sequence is shown in Table I. Prohead and phage proteins labeled with [^{3}H]leucine were separated by SDS–gel electrophoresis (Laemmli, 1970), the gels were stained with Coomassie blue, and excised gel slices were flame combusted in the Packard Model 307 TriCarb oxidizer (Packard Instrument Co.) for scintillation counting. In the oxidizer the ^{3}H-labeled proteins are combusted to water vapor in an oxygen-enriched atmosphere, avoiding problems of insolubility and quenching, with a recovery of more than 99% of the counts. The data were derived from several preparations of each type of particle and multiple gel lanes of each. The number of determinations (band excisions and combustion) of gp7, gp8, gp8.5, and gp10 for proheads was 31, 58, 27, and 58, respectively. For phage, there were 40 determinations each of gp8, gp10, gp11, gp12*, and gp9 and 19 determinations of gp8.5.

Statistical methods. Restricted maximum likelihood (REML) and Bayesian analyses are presented. Each analysis has a single fixed effect, the true number of copies, and distinct components of measurement variation for particles, batches, gels, oxidizer runs, and pure measurement error (i.e., variation between repeat measurements with the same particle, batch, gel, and oxidizer run). The more conventional REML method tends to give zero estimates of variance components for small data sets, as reflected in the analyses for gp9, gp11, and gp12* (Table II). In this respect, the Bayesian analysis is a conservative alternative. REML estimates and standard errors were computed using PROC MIXED in SAS version 6.12 (SAS Institute, Inc.). The Bayesian analysis used a Gibbs sampler (Gilks *et al.*, 1996) custom-written in S+ (version 3.4; Statistical Sciences, Inc.), with the model expressed in the constraint-case formulation (Hodges, 1998). Results are based on the 101st through 5100th draws. The estimated number of copies and the 95% confidence interval are, respectively, the average and the 2.5th and 97.5th percentiles of the Gibbs draws. The prior distributions were as follows: for the number of copies, flat (normal with infinite variance); for the reciprocal of the pure measurement error variance, γ with mean 0.025 and standard deviation 0.05; and for the reciprocals of the other variances, γ

TABLE II

Protein Composition of the ϕ29 Prohead and Virion

Gene product (structure)	Prohead/virion	Molecular weight[a]	Statistical method	Copies[b]	Standard error	95% confidence interval	Mass (MDa)[c]	Mass (MDa) EM/other[d]
gp7 (scaffold)	P	11 267	REML	25.8	3.4	(19.0, 32.6)[e]		
			Bayesian	26.9	8.8	(13.6, 40.9)[f]		
gp8 (major capsid)	P	49 847	REML	242.2	28.6	(185.1, 299.3)	12.07	11.71
			Bayesian	240.8	52.9	(168.0, 318.7)	12.00	
	V	49 847	REML	219.2	13.0	(193.1, 245.3)	10.93	
			Bayesian	217.9	27.8	(169.2, 265.8)	10.86	
gp10 (connector)	P/V	35 880		12				0.43
gp8.5 (head fiber)	P	29 491	REML	46.8	4.9	(37.0, 56.6)		
			Bayesian	47.2	9.3	(30.6, 62.8)		
	V	29 491	REML	62.8	3.1	(56.6, 68.9)		
			Bayesian	62.8	10.2	(49.9, 73.9)		
gp11 (lower collar)	V	33 801	REML	11.2	0.5	(10.3, 12.1)	0.38	
			Bayesian	11.2	3.6	(5.6, 16.6)	0.38	
gp12* (appendage)	V	73 888	REML	58.6	3.4	(51.9, 65.3)	4.33	
			Bayesian	58.2	8.1	(44.7, 71.2)	4.30	
gp9 (tail)	V	67 594	REML	9.1	0.6	(7.9, 10.3)	0.62	
			Bayesian	9.1	5.5	(3.3, 15.2)	0.62	
gp3 (DNA termini)	V	31 051		2				0.06

[a] Protein molecular weights were inferred from the nucleotide sequence, except for the cleavage product of the 92.1-kDa gp12, which was determined by MALDI-TOF mass spectrometry (Fig. 2).

[b] Copies of gp7, gp8, gp8.5, gp11, gp12*, and gp9 are based on the demonstration of 12-fold symmetry of gp10 in the ϕ29 neck (Carrascosa *et al.*, 1982). The number of determinations of gp7, gp8, gp8.5, and gp10 for proheads (P) was 31, 58, 27, and 58, respectively. For the virion (V), 40 determinations were made of gp8, gp10, gp11, gp12*, and gp9, and 19 determinations were made of gp8.5.

[c] Copies × protein molecular weight. The masses of scaffolding protein (gp7) and head fiber (gp8.5) are not given because varying amounts of these proteins were lost in particle isolation and purification (Fig. 3 and text).

[d] Twelve copies of gp10 are from electron microscopy (Carrascosa *et al.*, 1982); 235 copies of gp8 are from electron microscopy (Tao *et al.*, 1998); 2 copies of gp3 are from biochemical analysis (Salas *et al.*, 1978; Ito, 1978; Yehle, 1978).

[e] Copy estimate ± 2 standard errors.

[f] 2.5 and 97.5% quantiles of Gibbs samples.

with mean 0.5 and standard deviation 1. The prior distributions for the variances provide information equivalent to a quarter of a measurement and, apart from pure measurement error, treat the components of variation equally.

Gel densitometry. Stained gels were digitized with an Epson 800 scanner. Densitometry was performed on a Macintosh computer using the public domain NIH Image program (developed at the U.S. National Institutes of Health and available on the Internet at http://rsb.info.nih.gov/nih-image/).

Scanning transmission electron microscopy. Fiberless ϕ29 proheads, with and without scaffolding protein (gp7), were imaged in the STEM. The operation of the STEM and the mass analysis have been described (Wall *et al.*, 1998). A 3-μl aliquot of the prohead solution was injected into a drop of buffer on a thin carbon (2–3 nm) film supported by a holey thick carbon film on a titanium grid to which tobacco mosaic virus had been previously applied as an internal calibration. The grid was washed extensively with 20 mM ammonium acetate, blotted to a thin layer of liquid, and rapidly frozen in liquid nitrogen slush. It was freeze-dried overnight and transferred under vacuum to the microscope.

Mass spectrometry. gp12* was purified from an infected cell lysate prepared for prohead isolation by successive centrifugation in a 10–40% (w/v) linear sucrose density gradient and a 10–60% (v/v) linear Iodixanol density gradient (OptiPrep, Gibco BRL). A 0.75-μl aliquot of gp12* in water was spotted on a MALDI target and mixed with 0.75 μl of matrix. The matrix solution was a saturated solution of sinapinic acid [3,5-dimethoxy-4-hydroxycinnamic acid (530-59-6)] in 30:70 acetonitrile:water, 0.1% TFA. After the sample plus matrix dried on the target, a second layer of sample as well as matrix was pipetted on the same spot, mixed, and allowed to dry.

The instrument used for the collection of matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometric data was a Bruker Biflex III equipped with a N_2 laser (337 nm, 3-ns pulse length) and a microchannel plate detector. The data were collected in the linear mode, positive polarity, with an accelerating potential of 19 kV. Each spectrum is the accumulation of 200–500 laser shots. The data were smoothed with 20 points (Savitzky and Golay, 1964). External calibration was performed using the singly and doubly charged peaks of 15 μM bovine serum albumin.

RESULTS AND DISCUSSION

Protein Composition of the ϕ29 Prohead and Mature Phage

The ϕ29 precursor head or prohead is assembled from the scaffolding protein (gp7), the major capsid protein (gp8), head fibers (gp8.5), and head–tail connector (gp10), and once assembled, the scaffolding

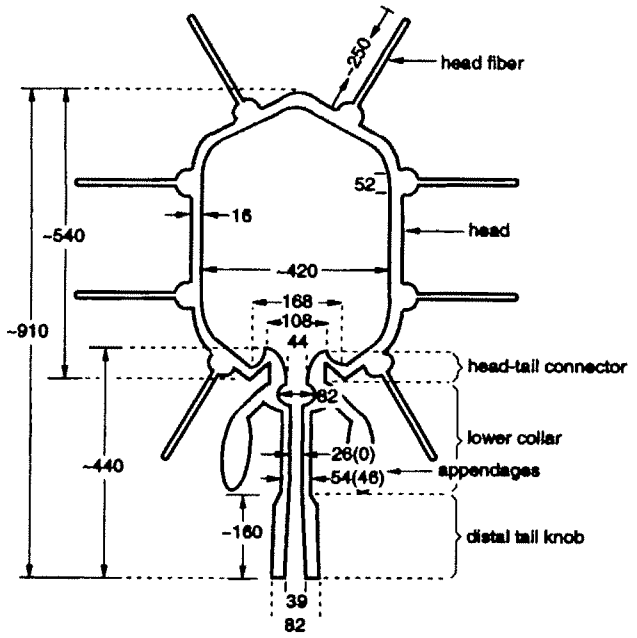

FIG. 1. Schematic diagram of an emptied ϕ29 particle (without DNA). Reprinted from Tao *et al.* (1998) Assembly of a tailed bacterial virus and its genome release studied in three dimensions, *Cell* **95,** 435, with permission from Cell Press. Distances are given in angstroms. Numbers in parentheses correspond to the mature phage. The 420-Å dimension of the prolate head refers to the external diameter.

protein can be removed without loss of biological activity. Fiberless proheads also have full biological activity. Prohead assembly is completed upon binding of multiple copies of 174-base prohead RNA (pRNA) to the connector. Following packaging of the double-stranded genome with covalently bound terminal protein (DNA-gp3), the pRNA and packaging ATPase gp16 are released, and the lower collar (gp11), tail (gp9), and neck appendages (gp12*) are assembled sequentially onto the head to give the mature virion (reviewed by Anderson and Reilly, 1993). A schematic diagram of an emptied ϕ29 particle (without DNA) is shown in Fig. 1.

Twelvefold symmetry of gp10 in the head–tail connector of ϕ29 was demonstrated by computer processing of electron micrographs of isolated necks (Carrascosa *et al.,* 1982). Atomic force microscopy has also demonstrated 12-fold symmetry of connectors (Muller *et al.,* 1997), and 12-fold (Carazo *et al.,* 1986) or 13-fold (Dube *et al.,* 1993; Tsuprun *et al.,* 1994) symmetry has been found by processing micrographs of free connectors. Recently, 12-fold symmetry of connectors has been shown by X-ray crystallography (Guasch *et al.,* 1998; Badasso *et al.,* 2000; Simpson *et al.,* 2000). Since 12 copies of gp10 were found in the neck of the virion, this number was used as an internal standard for quantification of the copies of the other proteins. The copy numbers of individual proteins relative to 12 copies of gp10 were calculated from the leucine content of the proteins inferred from the nucleotide sequence and the [^{3}H]leucine content of the proteins isolated by SDS gel electrophoresis (Table II). Combustion of gel fractions of radiolabeled proteins was required for accurate quantification because of the large range in amounts of individual proteins; for example, the quantity of gp8 was about 27 times that of gp10 (see Materials and Methods; Figs. 3b and 3c, insets).

The mean number of copies of the major capsid protein (gp8) in proheads and mature phage was 241 and 218, respectively, by the Bayesian method (Table II); thus the major capsid protein is conserved during particle maturation. These copy numbers approach the 235 copies of gp8 determined recently by cryo-EM and 3D image reconstruction (Tao *et al.,* 1998). The prolate ϕ29 prohead and virion consist of two icosahedral (half $T = 3$) caps that together contain 20 hexamers of gp8. The caps are separated by an equatorial band of 10 hexamers, giving 30 hexamers of gp8. Eleven pentamers of gp8 form the vertices (one additional pentamer is replaced by the connector at the base of the particles) to contribute another 55 subunits, giving a total of 235 copies of gp8 per prohead or virion. This is in contrast to a previous $T = 1$ model containing 90 dimers or 180 copies of gp8 per particle (Carrascosa *et al.,* 1981).

In situ crosslinking of ϕ29 head fibers (gp8.5) has suggested that the fibers are dimers, and free head fibers purportedly sediment as dimers (Carrascosa *et al.,* 1981). The number of head fiber binding sites on ϕ29 proheads and phage was determined to be 55 by cryo-EM and 3D reconstruction (Tao *et al.,* 1998). The fibers extend from the quasi-threefold axes at the two end caps of the head and are missing in the equatorial region of the capsid where there are no pentamers of gp8. There are 55 rather than 60 fiber binding sites because the replacement of 5 gp8 subunits by the connector precludes the binding of 5 fibers. Each fiber makes contact with a gp8 subunit in a pentamer at each quasi-threefold position, but no contact with hexamer subunits, and therefore the pentamers may differ in structure from hexamers at the quasi-threefold axes. In the present study the measured number of copies of the head fibers in the prohead and mature phage was 47 and 63 (24 and 31 dimers), respectively, by the Bayesian method (Table II). These numbers suggest only about 43 and 57% occupancy on the prohead and virion, respectively, of the fiber binding sites identified by cryo-EM. Variability of head fiber content suggests that fibers are lost in particle purification and particularly from proheads. Head fibers are difficult to observe on the prohead by negative staining and elec-

TABLE III
Estimated Masses of the ϕ29 Fiberless Prohead and Virion

Particle	Proteins and RNA or DNA	Copies[a]	Mass (MDa)[a]
Prohead	gp8	235	11.71
	gp10	12	0.43
	gp7	~147	~1.7
	pRNA	5–6[b]	0.29 or 0.35[b]
	Total mass (MDa) ~14.2		
Virion	gp8	235	11.71
	gp10	12	0.43
	gp11	12[c]	0.41
	gp12*	60[c]	4.43
	gp9	9	0.61
	DNA with gp3		12.76 + 0.06
	Total mass (MDa) 30.4[c]		

[a] See Tables I and II and Fig. 2.
[b] pRNA exhibits fivefold symmetry by cryo-EM (Simpson *et al.*, 2000), while genetic data support hexamers of pRNA (Zhang *et al.*, 1998; Guo *et al.*, 1998).
[c] Copies of gp11 and gp12*, measured at 11.2 and 58.2, respectively (Table II), are assumed to be 12 and 60.

tron microscopy (Nelson *et al.*, 1976) or by cryo-EM (Tao *et al.*, 1998) but are readily observed on the mature head. The difficulty in observing head fibers on proheads may be due to further loss of fibers in preparation for electron microscopy and suggests that fibers are more weakly attached to proheads than to the virion.

Prohead assembly requires the scaffolding protein (gp7), and in its absence the major capsid protein (gp8) forms anomalous particles called "telephones" (Hagen *et al.*, 1976). gp7 is generally lost from proheads to a variable extent during cell lysis or prohead isolation, and its maximum retention requires lysis of cells in highly concentrated form (see Materials and Methods). When unconcentrated infected cells lyse, the entire complement of gp7 can be lost. The estimated number of copies of gp7 reported in Table II was only 27 per particle, with a 95% confidence interval of 14–41. SDS gel electrophoresis of proheads that lack gp7 and exceptional proheads that retained a large amount of gp7 are shown in the insets in Figs. 3b and 3c. In the particles of Fig. 3c, gp7 content was estimated at 147 copies by densitometry and at 107 copies by STEM (see below). The mass of 147 copies of gp7 is about 1.7 MDa, equivalent to about 15% the mass of the major capsid protein (gp8) (Tables II and III). The structure of the scaffolding protein within the capsid is unknown; it may line the inner surface of the capsid or form a fibrous network that occupies the large internal space. A difference in shell structure of phage P22 procapsids with and without scaffolding protein was observed by cryo-EM, providing evidence that scaffolding protein binds specific shell subunits (Thuman-Commike *et al.*, 1999).

The agreement of the biochemical data of 241 and 218 copies of the major capsid protein (gp8) in the prohead and virion, respectively (Table II), with the 235 copies from cryo-EM and 3D reconstruction (Tao *et al.*, 1998) supports the validity of quantification of the lower collar, appendage, and tail proteins of the mature virion (Table II). In addition to the major capsid and connector proteins conserved from the prohead, the mature virion contains about 11 copies of the lower collar protein (gp11) that is attached to the connector, about 58 copies of the appendage protein (gp12*, cleavage product of gp12) that is present on the virion as 12 spindle-shaped structures (Anderson *et al.*, 1966), and about 9 copies of the tail protein (gp9) (Table II). The copy number of gp12* was based on N-terminal amino acid sequencing to demonstrate that gp12 is cleaved at the C-terminus (not shown, but see Table I) and on determination of the mass/charge of purified gp12* by MALDI-TOF mass spectrometry (Fig. 2). A peak at m/z = 73 888 was assigned as the singly charged peak for gp12*, whereas a doubly charged peak was not observed. The broadness of the peak at the descending part of the trace may be due to an unresolved sinapinic acid adduct or molecules of gp12* that are posttranslationally modified and that differ by masses that could not be resolved. The copy numbers for the lower collar, appendage, and tail proteins are at least twofold higher than those estimated previously (Carrascosa *et al.*, 1981).

The connector, lower collar, and tail proteins

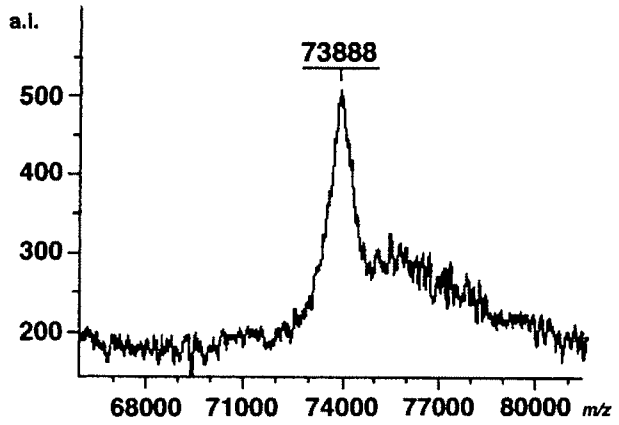

FIG. 2. Linear mode MALDI-TOF mass spectrum of gp12* using sinapinic acid as matrix.

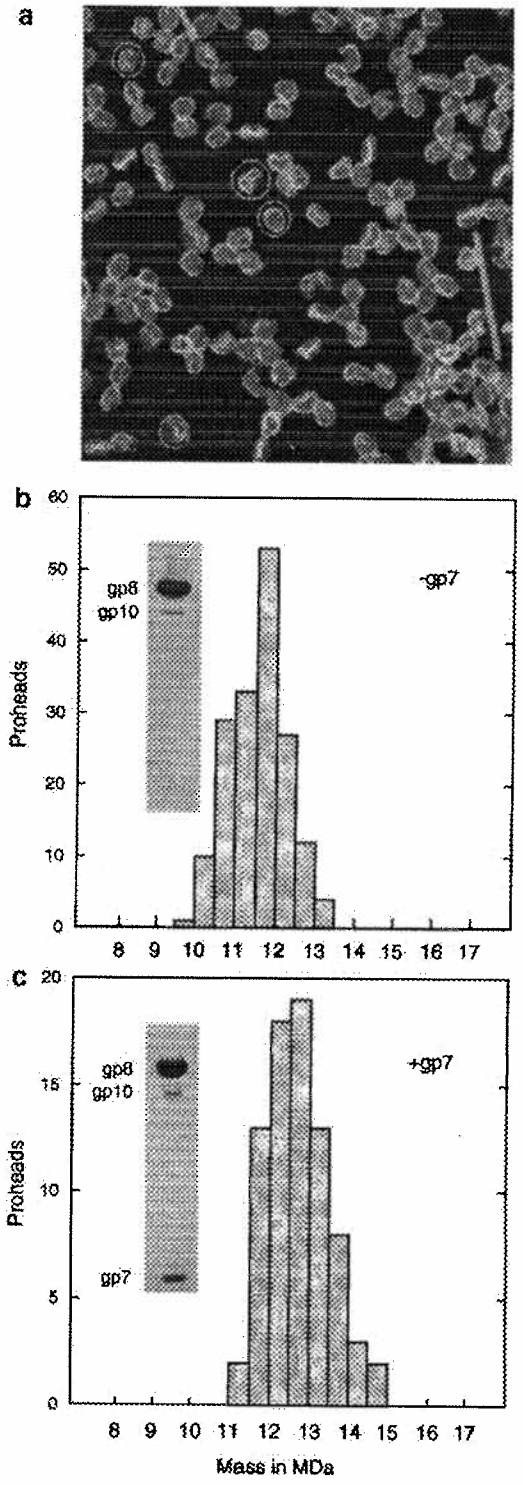

FIG. 3. Mass of the fiberless ϕ29 prohead, without and with scaffolding protein (gp7). (a) A typical dark-field electron micrograph from the STEM, of fiberless proheads without scaffolding protein, is shown. The long thin cylindrical molecule at the right is tobacco mosaic virus. Isolated particles suitable for mass measurements are circled. The full scale of the micrograph is 1.024 μm. (b and c) Histograms of masses of fiberless, pRNA-free proheads without and with scaffolding protein (gp7), respectively. The mean masses and standard deviations in b and c were 11.6 ± 0.9 and 12.8 ± 0.7 MDa, respectively. The difference of 1.2 MDa corresponds to 107 copies of the 11.3-kDa gp7. Insets show SDS–polyacrylamide gel electrophoresis of the particles; gp7 in c was estimated at 147 copies by densitometry.

together form the channel, with an inner diameter ranging from 26 to 44 Å (Tao *et al.*, 1998; Fig. 1), through which DNA–gp3 passes in infection. gp11 forms both a collar and a stem that extends from the connector to the tail (Carrascosa *et al.*, 1983). The neck appendages function as the adsorption organelles of ϕ29 (Tosi *et al.*, 1975). The true number of copies of gp11 and gp12* may be 12 and 60, respectively, and the entire neck/tail of ϕ29 probably has quasi-sixfold symmetry. Then each of the 12 individual appendages is a pentamer of gp12* rather than a dimer as reported previously (Villanueva *et al.*, 1981). The 9 copies of gp9 in the tail are likely arranged with threefold symmetry. The neck and tail structures are under study by cryo-EM, and the estimated volumes of these structures (not shown) agree with the protein content determined here.

Determination of the Mass of the ϕ29 Prohead by STEM

The mass of the ϕ29 prohead was confirmed by STEM. A STEM image of unstained/freeze-dried proheads lacking head fibers (gp8.5), scaffolding protein (gp7), and pRNA is shown in Fig. 3a. Representative mass histograms of these proheads, and counterparts in which scaffolding protein gp7 was retained by lysis of concentrated cells, are shown in Figs. 3b and 3c. The mean masses for these particles and the standard deviations were 11.6 ± 0.9 and 12.8 ± 0.7 MDa, respectively. The difference of 1.2 MDa corresponds to 107 copies of the 11.3-kDa gp7, compared to the 147 copies estimated by densitometry (Figs. 3b and 3c, insets).

The mean masses and standard deviations of four preparations of fiberless proheads without scaffolding protein (gp7) and pRNA determined here were 11.6 ± 0.9, 14.2 ± 2.1, 9.9 ± 0.8, and 13.3 ± 0.6 MDa, respectively, using 169, 244, 163, and 232 particles in the four analyses. The average mass of the four preparations was 12.25 MDa. This compares favorably with the 12.43-MDa mass of fiberless proheads containing gp8 (12.00 MDa) and gp10 (0.43 MDa) calculated from the present biochemical data (Table II). It also compares favorably with the probable true mass of 12.14 MDa of fiberless proheads containing 12 copies of gp10 (0.43 MDa; Carrascosa *et al.*, 1982) and 235 copies of gp8 (11.71 MDa) from cryo-EM and 3D reconstruction (Tao *et al.*, 1998).

Estimates of the Masses of the Intact ϕ29 Prohead and Mature Virion

Four of the five STEM mass values reported above are for proheads without head fibers (gp8.5), scaffolding protein (gp7), and prohead RNA (pRNA). Adding 1.7 MDa for 147 copies of scaffolding protein (gp7) (Fig. 3c) and 0.29 or 0.35 MDa for five or six copies of the 57.6-kDa prohead RNA that are bound to the connector (Zhang *et al.*, 1998; Guo *et al.*, 1998; Simpson *et al.*, 2000) gives a total mass for the fiberless prohead of about 14.2 MDa (Table III). Fifty-five dimeric fibers would increase this mass by 3.2 MDa to about 17.4 MDa.

An estimated mass of the fiberless virion (Table III) is based on determination of the copy numbers of gp8 and gp10 by electron microscopy (Carrascosa *et al.*, 1982; Tao *et al.*, 1998) and assumptions from the present biochemical data that likely provide close approximations of the copy numbers of gp11, gp12*, and gp9 (Table II). Adding the 19 285-bp DNA genome of 12.76 and 0.06 MDa for gp3 at the DNA termini gives a mass of about 30.4 MDa for the fiberless virion. Fifty-five dimeric fibers would increase this mass by 3.2 MDa to about 33.6 MDa.

The number of copies of each protein in the virion, except gp10, is substantially greater than that estimated previously from amino acid analysis of acid-hydrolyzed phage proteins and the particle molecular weight from the sedimentation and diffusion coefficients (Carrascosa *et al.*, 1981). These differences stem from errors in the number of particular amino acid residues in the proteins, now known with certainty from the nucleotide sequence, an underestimation of the mass of the particle, and the use of gp12* as internal standard with the assumption that it is a dimer. The new composition data provide a firm basis for ongoing structure–function analyses of ϕ29 morphogenesis and infection.

We thank Sarah Turnquist and Carmen Varga for technical help. The work was supported by NIH research grants to D.A. and by NIH/NIDCR P30-DE09737 to J.H. The BNL STEM is an NIH Supported Resource Center, NIH P41-RR01777, with additional support provided by DOE, OBER.

REFERENCES

Anderson, D. L., Hickman, D. D., and Reilly, B. E. (1966) Structure of *Bacillus subtilis* bacteriophage ϕ29 and the length of ϕ29 deoxyribonucleic acid. *J. Bacteriol.* **91,** 2081–2089.

Anderson, D., and Reilly, B. E. (1993) Bacteriophage ϕ29 morphogenesis, *in* Hoch, J. A., Losick, R., and Sonenshein, A. L. (Eds.), *Bacillus subtilis* and Other Gram Positive Bacteria: Physiology, Biochemistry, and Molecular Genetics, pp. 859–867, ASM Publications, Washington, DC.

Badasso, M. O., Leiman, P. G., Tao, Y., He, Y., Ohlendorf, D. H., Rossmann, M. G., and Anderson, D. (2000) Purification, crystallization, and initial X-ray analysis of the head–tail connector of bacteriophage ϕ29. *Acta Crystallographica Sect. D* **56,** 1187–1190.

Bjornsti, M. A., Reilly, B. E., and Anderson, D. L. (1981) *In vitro* assembly of the *Bacillus subtilis* bacteriophage ϕ29. *Proc. Natl. Acad. Sci. USA* **78,** 5861–5865.

Carazo, M., Donate, L. E., Herranz, L., Secilla, J. P., and Carrascosa, J. L. (1986) Three-dimensional reconstruction of the connector of bacteriophage ϕ29 at 1.8 nm resolution. *J. Mol. Biol.* **192,** 853–867.

Carrascosa, J. L., Mendez, E., Corral, J., Rubio, V., Ramirez, G., Salas, M., and Vinuela, E. (1981) Structural organization of *Bacillus subtilis* phage ϕ29: A model. *Virology* **111,** 401–413.

Carrascosa, J. L., Vinuela, E., Garcia, N., and Santisteban, A. (1982) Structure of the head–tail connector of bacteriophage ϕ29. *J. Mol. Biol.* **154,** 311–324.

Carrascosa, J. L., Carazo, J. M., and Garcia, N. (1983) Structural localization of the proteins of the head to tail connecting region of bacteriophage ϕ29. *Virology* **124,** 133–143.

Dube, P., Tavares, P., Lurz, R., and van Heel, M. (1993) The portal protein of bacteriophage SPP1: A DNA pump with 13-fold symmetry. *EMBO J.* **12,** 1303–1309.

Gilks, W. R., Richardson, S., and Spiegelhalter, D. J. (1996) Markov Chain Monte Carlo in Practice, Chapman and Hall, London.

Guasch, A., Pous, J., Parraga, A., Valpuesta, J. M., Carrascosa, J. L., and Coll, M. (1998) Crystallographic analysis reveals the 12-fold symmetry of the bacteriophage ϕ29 connector particles. *J. Mol. Biol.* **281,** 219–225, doi:10.1006/jmbi.1998.1928.

Guo, P., Grimes, S., and Anderson, D. (1986) A defined system for *in vitro* packaging of DNA-gp3 of the *Bacillus subtilis* bacteriophage ϕ29. *Proc. Natl. Acad. Sci. USA* **83,** 3505–3509.

Guo, P., Erickson, S., and Anderson, D. (1987) A small viral RNA is required for *in vitro* packaging of bacteriophage ϕ29 DNA. *Science* **236,** 690–694.

Guo, P., Zhang, C., Chen, C., Garver, K., and Trottier, M. (1998) Inter-RNA interaction of phage ϕ29 pRNA to form a hexameric complex for viral DNA transportation. *Mol. Cell* **2,** 149–155.

Hagen, E. W., Reilly, B. E., Tosi, M. E., and Anderson, D. L. (1976) Analysis of gene function of bacteriophage ϕ29 of *Bacillus subtilis:* Identification of cistrons essential for viral assembly. *J. Virol.* **19,** 501–517.

Hodges, J. S. (1998) Some algebra and geometry for hierarchical models, applied to diagnostics (with discussion). *J. R. Stat. Soc. Ser B.* **60,** 497–536.

Ito, J. (1978) Bacteriophage ϕ29 terminal protein: Its association with the 5′ termini of the ϕ29 genome. *J. Virol.* **28,** 895–904.

Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227,** 680–685.

Lee, C.-S., and Guo, P. (1994) A highly sensitive system for the *in vitro* assembly of bacteriophage ϕ29 of *Bacillus subtilis. Virology* **202,** 1039–1042, doi:10.1006/jivo.1994.1434.

Muller, D. J., Engel, A., Carrascosa, J. L., and Velez, M. (1997) The bacteriophage ϕ29 head–tail connector imaged at high resolution with the atomic force microscope in buffer solution. *EMBO J.* **16,** 2547–2553.

Nelson, R. A., Reilly, B. E., and Anderson, D. L. (1976) Morphogenesis of bacteriophage ϕ29 of *Bacillus subtilis:* Preliminary isolation and characterization of intermediate particles of the assembly pathway. *J. Virol.* **19,** 518–532.

Polsinelli, M., and Beretta, M. (1966) Genetic recombination in crosses between *Streptomyces aureofaciens* and *Steptomyces rimosus. J. Bacteriol.* **91,** 63–68.

Salas, M., Mellado, R. P., Vinuela, E., and Sogo, J. M. (1978) Characterization of a protein covalently linked to the 5′ termini of the DNA of *Bacillus subtilis* phage ϕ29. *J. Mol. Biol.* **119,** 269–291.

Savitzky, A., and Golay, M. J. E. (1964) Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* **36,** 1627–1639.

Simpson, A. A., Tao, Y., Leiman, P. G., Badasso, M. O., He, Y., Jardine, P. J., Olson, N. H., Morais, M. C., Grimes, S., Anderson, D. L., Baker, T. S., and Rossmann, M. G. (2000) Structure of the bacteriophage ϕ29 DNA packaging motor. *Nature* **408,** 745–750.

Tao, Y., Olson, N. H., Xu, W., Anderson, D. L., Rossmann, M. G., and Baker, T. S. (1998) Assembly of a tailed bacterial virus and its genome release studied in three dimensions. *Cell* **95,** 431–437.

Thuman-Commike, P. A., Greene, B., Malinski, J. A., Burbea, M., McGough, A., Chiu, W., and Prevelige, P. E., Jr. (1999) Mechanism of scaffolding-directed virus assembly suggested by comparison of scaffolding-containing and scaffolding-lacking P22 procapsids. *Biophys. J.* **76,** 3267–3277.

Tosi, M. E., Reilly, B. E., and Anderson, D. L. (1975) Morphogenesis of bacteriophage ϕ29 of *Bacillus subtilis:* Cleavage and assembly of the neck appendage protein. *J. Virol.* **16,** 1282–1295.

Tsuprun, V., Anderson, D., and Egelman, E. H. (1994) The bacteriophage ϕ29 head–tail connector shows 13-fold symmetry in both hexagonally packed arrays and as single particles. *Biophys. J.* **66,** 2139–2150.

Villanueva, N., Lazaro, J. M., and Salas, M. (1981) Purification, properties and assembly of the neck-appendage protein of the *Bacillus subtilis* phage ϕ29. *Eur. J. Biochem.* **117,** 499–505.

Vlcek, C., and Paces, V. (1986) Nucleotide sequence of the late region of *Bacillus* phage ϕ29 completes the 19285-bp sequence of ϕ29 genome. Comparison with the homologous sequence of phage PZA. *Gene* **46,** 215–225.

Wall, J. S., Hainfeld, J. F., and Simon, M. N. (1998) Scanning transmission electron microscopy of nuclear structures. *Methods Cell Biol.* **53,** 139–164.

Yehle, C. O. (1978) Genome-linked protein associated with the 5′ termini of bacteriophage ϕ29 DNA. *J. Virol.* **27,** 776–783.

Zhang, F., Lemieux, S., Wu, X., St.-Arnaud, D., McMurray, C. T., Major, F., and Anderson, D. (1998) Function of hexameric RNA in packaging of bacteriophage ϕ29 DNA *in vitro. Mol. Cell* **2,** 141–147.