Structural and functional similarities between the capsid proteins of bacteriophages T4 and HK97 point to a common ancestry

Andrei Fokine[†‡], Petr G. Leiman[†‡], Mikhail M. Shneider[§], Bijan Ahvazi[¶], Karen M. Boeshans[¶], Alasdair C. Steven[‖], Lindsay W. Black[††], Vadim V. Mesyanzhinov[§], and Michael G. Rossmann[†‡‡]

[†]Department of Biological Sciences, Purdue University, 915 West State Street, West Lafayette, IN 47907; [§]Laboratory of Molecular Bioengineering, Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, 16/10 Miklukho-Maklaya Street, 117997 Moscow, Russia; [¶]X-Ray Crystallography Facility/Office of Science and Technology, National Institute of Arthritis and Musculoskeletal and Skin Diseases, Building 50, Room 1345, 50 South Drive, MSC 8023, Bethesda, MD 20892; [‖]Laboratory of Structural Biology, National Institute of Arthritis and Musculoskeletal and Skin Diseases, Building 50, Room 1517, 50 South Drive, MSC 8025, Bethesda, MD 20892; and [††]Department of Biochemistry and Molecular Biology, University of Maryland School of Medicine, Baltimore, MD 21201

Contributed by Michael G. Rossmann, March 17, 2005

Gene product (gp) 24 of bacteriophage T4 forms the pentameric vertices of the capsid. Using x-ray crystallography, we found the principal domain of gp24 to have a polypeptide fold similar to that of the HK97 phage capsid protein plus an additional insertion domain. Fitting gp24 monomers into a cryo-EM density map of the mature T4 capsid suggests that the insertion domain interacts with a neighboring subunit, effecting a stabilization analogous to the covalent crosslinking in the HK97 capsid. Sequence alignment and genetic data show that the folds of gp24 and the hexamer-forming capsid protein, gp23*, are similar. Accordingly, models of gp24* pentamers, gp23* hexamers, and the whole capsid were built, based on a cryo-EM image reconstruction of the capsid. Mutations in gene 23 that affect capsid shape map to the capsomer's periphery, whereas mutations that allow gp23 to substitute for gp24 at the vertices modify the interactions between monomers within capsomers. Structural data show that capsid proteins of most tailed phages, and some eukaryotic viruses, may have evolved from a common ancestor.

evolution | gene product 24 | major capsid protein

The protein shells of viral capsids are remarkably stable, yet dynamic structures. They have to protect the genome during its transfer between hosts, withstand the high pressure of the condensed nucleic acid, and be able to release the genome once a susceptible host has been recognized. To reconcile both stability and dynamic requirements, assembled procapsids of many viruses undergo large conformational changes during genome packaging and maturation (1).

The capsid of the dsDNA tailed bacteriophage T4 is a prolate icosahedron (Fig. 1) whose capsomers form a $T_{end} = 13$ laevo hexagonal lattice in the end caps and a $T_{mid} = 20$ lattice in the cylindrical midsection (2). The protein shell consists of the major capsid protein gene product (gp) 23*, the pentameric vertex protein gp24*, the portal protein or "connector" gp20, and the two accessory proteins, gp hoc (highly antigenic outer capsid protein) and gp soc (small outer capsid protein) (3), that decorate the outside of the shell. The dodecameric connector replaces a pentamer of gp24* at one of the 12 vertices and serves as a special portal for DNA packaging, tail attachment, and DNA exit (4, 5).

During procapsid assembly, gp23, gp24, and gp20 form a shell around the core structure composed primarily of the scaffolding protein gp22 and assembly protease gp21 (3). The protease activates once procapsid assembly has been completed and cleaves the proteins of the core into small peptides, most of which leave the maturing procapsid, freeing space for the genome. The gp21 protease also cleaves a 65-residue-long, amino-terminal fragment from the 56-kDa gp23, generating the 48.7-kDa gp23* (3). In addition, the 10-residue, amino-terminal region of the 48.7-kDa gp24 is cleaved, giving rise to the 47.6-kDa gp24* (3). These cleavages trigger a large conformational rearrangement in the procapsid, resulting in expansion and causing the capsid's facets to flatten and the capsid's wall to become thinner.

A number of mutations in different genes coding capsid proteins lead to various abnormalities in assembly ranging from formation of isometric and/or giant capsids to substitution of all gp24 pentamers with gp23 pentamers at the vertices (3, 6). The latter observation suggests that the structures of gp23 and gp24, as well as the character of interactions between them, are probably similar.

In the present study, we have crystallized recombinant, assembly-naïve gp24, which, unlike gp23 (7), does not have the propensity to assemble into large polymers. These crystals were used to solve the structure of gp24 to 2.9-Å resolution from which it was possible to build a homology model of gp23*. The gp24 structure was found to be similar to the mature capsid protein of phage HK97 (8), apart from having a 60-residue insertion domain. The structures of gp24 and gp23 were combined with the cryo-EM map of the T4 capsid (2) to build a pseudoatomic-resolution model, allowing identification of the interactions within and between gp24* pentamers and gp23* hexamers. The model also provided a framework for rationalizing the morphogenetic mutants that have been characterized for the T4 capsid proteins (3). The association of the insertion domain with symmetry-related subunits in the gp24 pentamer might play a role similar to the chain-mail interactions between the HK97 capsid subunits (9). Furthermore, the presence of hexameric and pentameric building blocks of HK97-like molecules in several other tailed phages establishes that capsid proteins of tailed phages probably have evolved from a common ancestor.

Methods

Protein Expression, Purification, Crystallization, and Data Collection. DNA encoding gp24 was cloned into the pET-23 expression vector (Novagen) without using the C-terminal His tag. The Se-methionine (SeMet) protein was expressed in the B834(DE3) strain of *Escherichia coli*, which is a methionine-requiring auxotroph (10). The modified M9 medium used for expression of the SeMet protein contained each amino acid at a concentration of 40 $\mu g \cdot ml^{-1}$, the EAGLE medium vitamin complex (ICN) at 1 $\mu g \cdot ml^{-1}$, 0.4% (wt/vol) glucose, 0.1 mM $CaCl_2$, and ampicillin at 200 $\mu g \cdot ml^{-1}$. MALDI-MS showed that there were five SeMet residues in the protein, of six in the sequence, thus confirming the cleavage of the amino-terminal methionine. The protein was purified by ammo-

Abbreviations: gp, gene product; soc, small outer capsid protein; E loop, elongated loop; P domain, peripheral domain; A domain, axial domain.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 1YUE and 1Z1U).

[‡]A.F. and P.G.L. contributed equally to this work.

[‡‡]To whom correspondence should be addressed. E-mail: mr@purdue.edu.

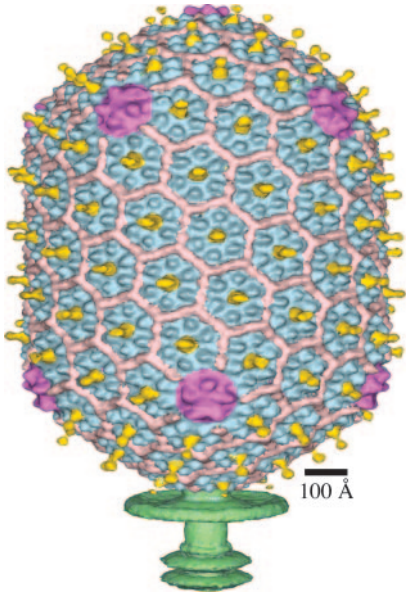

Fig. 1. A cryo-EM reconstruction at 22.5-Å resolution of the T4 head capsid is shown. The gene products 23, 24, hoc, and soc are colored blue, magenta, yellow, and pink, respectively. The reconstruction used 5-fold averaging about the long axis of the head. The features of the tail (green) appear blurred because the tail has six-fold symmetry. [Adapted from Fokine *et al.* (2).]

nium sulfate precipitation and ion exchange chromatography by using a DEAE-cellulose column.

The initial crystallization screening was performed by using Hampton Crystal Screens 1 and 2 (11) and the sitting-drop vapor-diffusion method at 20°C. Drops containing 1 μl of gp24 in 10 mM Tris·HCl (pH 7.5) were mixed with 1 μl of reservoir solution and equilibrated against 100 μl of reservoir solution. The protein crystallized in four different conditions. The best crystals were obtained in 2–3 days from 0.1 M Hepes-Na (pH 7.5), 10% (vol/vol) isopropanol, and 0.2 M sodium citrate. The crystals were washed in a cryoprotectant solution [20% glycerol, 0.1 M Hepes-Na (pH 7.5), 30% (vol/vol) isopropanol, and 0.2 M sodium citrate] and flash-frozen in liquid nitrogen. The multiwavelength anomalous dispersion diffraction data (Table 1) were collected at BioCARS beamlines 14C and 14D of the Advanced Photon Source (Argonne National Laboratory, Argonne, IL). The data were processed with the program HKL2000 (12). The crystals belonged to space group $P6_5$, with cell dimensions $a = 107.614$ Å and $c = 82.636$ Å.

The diffraction data for an Au derivative of another crystal form were collected at beamline X9B at the Brookhaven National Laboratory Synchrotron Light Source (Upton, NY) (Table 1) and processed with the program HKL2000. The crystals belonged to space group 6_5, but, with $a = 107.8$ and $c = 111.8$ Å, had a much longer c axis. Although these crystals also contained only one monomer per asymmetric unit, the longer c axis implied a much higher solvent content.

Structure Solution, Model Building, and Refinement. The determination of selenium positions and the initial phasing were performed with the program SOLVE (13), by using four-wavelength data collected at beamline 14D (Table 1). The phases were then improved by solvent flattening (14), using the CNS program (15). The resulting electron density map was interpreted in terms of an atomic model, using the program XFIT (16). Crystallographic refinement of the atomic model was accomplished with the program CNS (15) by using the data collected to 2.9-Å resolution at beamline 14C (Table 1). The resulting $R_{working}$ and R_{free} values were 0.27 and 0.30, respectively. The rms deviations of bond lengths and bond angles from idealized values were 0.01 Å and 1.66°, respectively.

The structure phases for the Au-containing crystals were determined by using three-wavelength anomalous dispersion followed by solvent flattening. There appears to be only little difference in the map quality produced by the two types of crystals, although the map based on the initial Se phasing was perhaps a little easier to interpret.

Results and Discussion

Crystal Structure of gp24. The T4 gene 24, coding for amino acid residues 1–427, was cloned into a high-expression *E. coli* vector pET-23 and overexpressed in BL21(DE3) cells, as well as in the methionine-auxotrophic strain B834(DE3). The identity of the purified recombinant gp24 was confirmed by N-terminal sequencing and MS. The native and SeMet-substituted proteins were purified to homogeneity and crystallized. The structure was solved by the multiwavelength anomalous dispersion technique applied to two crystal forms. One of these was a crystal of the SeMet derivative; the other was a crystal of the WT protein soaked in 10 mM $KAu(CN)_2$.

The gp24 polypeptide amino acid sequence, from residues 2–422, could be built into the electron density map of the SeMet crystals, although residues 63–81 and 143–154 were found to be disordered (Fig. 2). The electron density derived from the Au-containing crystals was entirely consistent with this interpretation. The structure of gp24, except for residues 63–154, could be superimposed onto the structure of the HK97 capsid protein (8, 17) with an rms deviation of 2.7 Å for 149 equivalenced C_α atoms (Figs. 2 and 3). However, the amino acid identity of the structurally aligned sequences was only 17%. The HK97 capsid protein structure consists of the amino-terminal arm (N-arm), the elongated loop (E loop), the peripheral domain (P domain), and the axial domain (A domain) (8) (Fig. 2). The P domains are on the periphery of a capsomer, surrounding a central ring of wedge-shaped A domains. The superimposable parts between T4 gp24 and the capsid protein of HK97 consist of the P and A domains.

In the mature HK97 capsid protein, the N-arm assumes an

Table 1. Crystallographic data for the SeMet solution and Au derivative

	SeMet solution					Au derivative		
Radiation source		APS, 14BMD, BioCARS			APS, 14BMC, BioCARS		NSLS, X9B	
Wavelength, Å	0.9414 (high-energy remote)	0.9791 (peak)	0.9794 (inflection point)	1.0188 (low-energy remote)	0.9000	1.0345 (high-energy remote)	1.0385 (peak)	1.0419 (inflection point)
Detector		ADSC Quantum-4			ADSC Q315		ADSC Quantum-4	
Resolution, Å		3.40			2.90	3.35	3.35	3.30
No. of unique reflections	14,579	14,580	14,676	14,721	11,621	20,931	22,990	21,492
Redundancy, ratio	6.24	6.29	6.32	6.24	14.12	2.8	2.6	3.1
Completeness, %	97.6	97.6	98.2	98.4	97.7	98.9	98.7	98.3
R_{merge}, %	5.4	6.7	5.6	5.5	7.3	7.9	9.5	9.2
$\langle I/\sigma(I)\rangle$	28.1	23.0	25.2	24.8	24.6	12.78	14.75	10.07

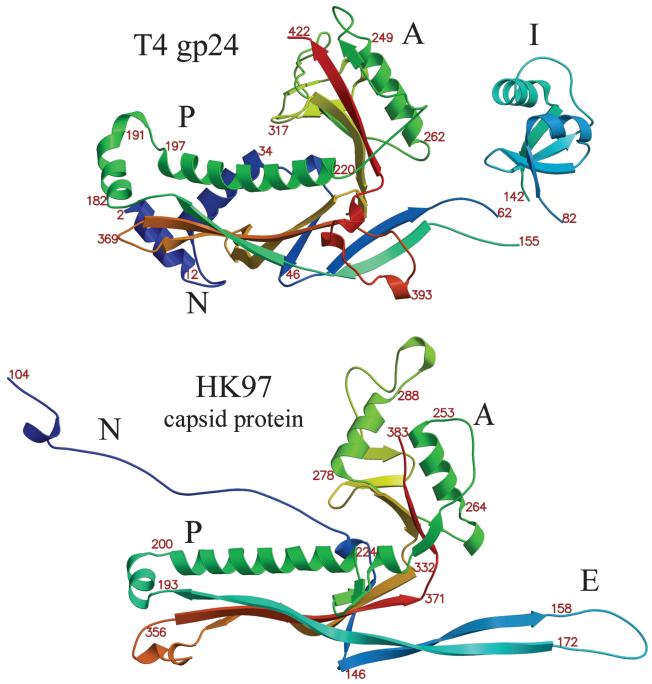

Fig. 2. Ribbon diagrams compare the structure of T4 gp24 (*Upper*) with the structure of the HK97 capsid protein (*Lower*). Colors range from blue at the amino end to red at the carboxyl end of the polypeptides. Amino acid sequence numbers are marked in strategic positions. The P, A, and insertion (I) domains, as well as the E loop (E) and the N-terminal domain (N), are indicated. The entrance and exit linker polypeptides to the insertion domain of gp24 are disordered.

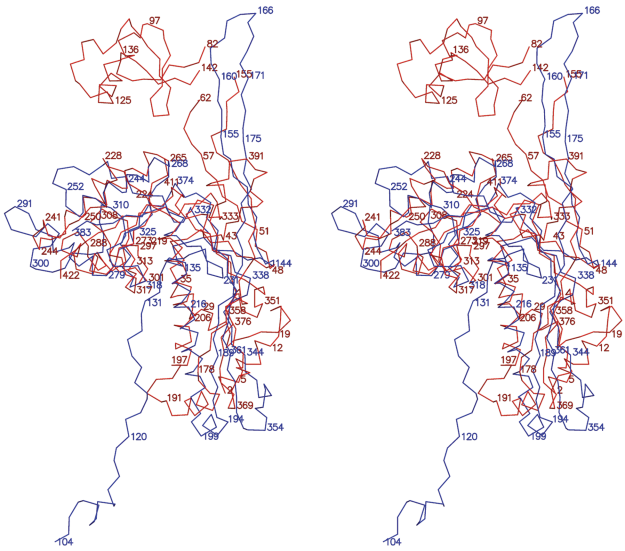

Fig. 3. Stereo diagram shows the superposition of the Cα backbones of T4 gp24 (red) and the HK97 capsid protein (blue).

extended ordered conformation (Fig. 2) (8) and invades two neighboring subunits, one in the same capsomer and the other in a neighboring capsomer. However, in the assembly-naïve state, the conformation of the N-arm is unknown, and it is preceded by the 102-residue-long delta domain, which is cleaved during capsid maturation. The N-terminal region of assembly-naïve gp24 (residues 2–40) is compact and contains two α-helices that pack against the β-sheet and the long α-helix of the P domain (Fig. 2) within the same subunit. The consensus cleavage site in gp24, inferred from experimentally determined sites in other T4 capsid proteins, is after Glu-10 (3). This position corresponds to an only partially exposed site close to the end of the first helix, consistent with only some of the gp24 molecules in the mature T4 capsid having been cleaved (6, 18).

A striking feature of the HK97 capsid protein is that during maturation the E loop becomes covalently crosslinked to a hairpin loop in a neighboring subunit. This crosslinking generates hexameric and pentameric rings of subunits catenated like the Olympic rings, producing a "chain-mail" capsid structure (8, 19). The E loop in gp24 does not crosslink with another subunit, but it has an additional domain (residues 82–142) inserted between two linker regions that are disordered in both crystal forms. A DALI search (20) showed that the 60 aa of the gp24 insertion domain has a topological similarity to the chitin binding domain of chitinase (21, 22), although the residues implicated in chitin binding are not the same as those in gp24. On superposition, it was found that there are rms deviations of 1.9 and 1.8 Å for 41 and 43 equivalenced C_α atoms for chitinase A1 and B, respectively. The corresponding levels of amino acid identity are 22% and 12%.

The insertion domain interacts with the neighboring, 6_5 symmetry-related subunit in both the Se- and Au-containing crystals via matching hydrophobic patches, flanked by polar and charged residues (Fig. 4). The area of contact between the insertion domain and the HK97-like fold is 788 Å^2, which constitutes ≈19% of the total surface area of the insertion domain. The distributions of electric potential (23) calculated for the two contacting surfaces were found to be complementary. The hydrophobic nature and the surface area of this interface are typical for oligomeric interfaces, but not for protein crystal lattice contacts (24) and, therefore, are likely to be of biological significance, suggesting that interactions of the insertion domain with the HK97-like domain in the crystal and in the capsid are similar.

The connecting peptides leading to and from the insertion domain are disordered and are sufficiently long to create some uncertainty as to which symmetry-related insertion domain is connected to a given HK97-like domain. This ambiguity can be resolved by comparing the gp24 crystal packing with the crystal structure of the HK97 mature head, where in each hexamer and pentamer the E loop of a given monomer extends to and over the neighboring counterclockwise symmetry-related monomer when viewed from outside the virus. Analogously, the insertion domain in gp24 extends to and over a molecule related by a crystallographic 6_5 screw axis. Thus, the insertion domain interacts with the counterclockwise, symmetry-related molecule, creating an arrangement in the gp24 crystal similar to the intersubunit interactions as in an HK97 capsomer.

Comparison of the gp24 Crystal Structure with the Cryo-EM Reconstruction of the T4 Capsid. Based on the similarities in structure and function between gp24 of T4 and the capsid protein of HK97, pentamers of gp24 were modeled by superimposing the HK97-like domain of the gp24 structure onto the HK97 pentamers found in Prohead II and mature Head II (1, 8). The position of the insertion domain was determined by assuming the same interaction between it and the HK97-like domain of the neighboring subunit, as found in the crystal structure. This modeling procedure substantially altered the spatial relationship between the HK97-like domain and insertion domain belonging to the same polypeptide chain in the gp24* pentamer model derived from Head II, but not in the model derived from the Prohead II pentamers. These structural alterations are justified by the flexibility of the linker peptides that connect the two domains, as observed in the crystal structure. These considerations also suggest that the spatial relationship between the HK97-like domain and insertion domain in the gp24 crystal structure corresponds to their arrangement in the immature T4 capsid.

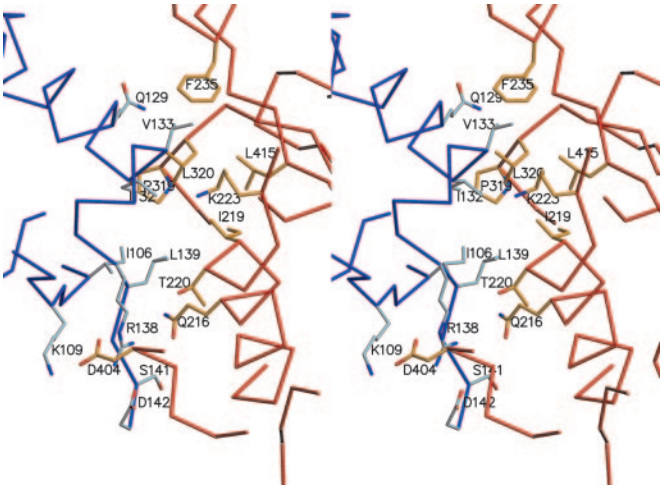

Fig. 4. Stereo diagram shows the interactions of the insertion domain with the HK97-like domain of the crystallographic, 6_5 symmetry-related molecule. The insertion domain of the blue molecule interacts with the A and P domains of the red molecule.

As the mature T4 virion contains the posttranslationally cleaved gp24* structure rather than the complete gp24 molecule, the first 10 aa (corresponding to the first helix of the N-terminal region) were removed from the HK97 Prohead II-based and Head II-based pentameric models. These gp24* pentamers were fitted into the 22-Å resolution cryo-EM map of the WT T4 head (2) (Fig. 5 *A* and *B*). The fitting procedure starting with the structure derived from the HK97 Prohead II pentamers did not result in a satisfactory fit. However, the fitting procedure starting with the structure derived from the HK97 Head II pentamers was consistent with visual and computational interpretation of the map.

The protein shells of many large dsDNA viruses with isometric or prolate capsids are composed of a major hexameric or pseudohexameric capsid protein with a different pentameric protein at the five-fold vertices. In the case of T4, gp23 (the major capsid protein) and gp24 have similar molecular weights and have 21% overall sequence identity, with 25%, 18%, and 7% sequence identity corresponding to the P, A, and insertion domains, respectively (Fig. 6). Therefore, the structures of these proteins should be similar with both proteins containing a HK97-like domain and an insertion domain. This conclusion is supported by the observation that point mutations in gene 23 allow gp23 to substitute for gp24 (6). A precedent for the major capsid protein and the vertex protein having a similar fold occurs in the lipid-containing icosahedral phage PRD1, in which both molecules are composed of two consecutive jellyrolls (25). Accordingly, a model of the gp23* hexamer was built based on sequence alignment with gp24*, the HK97 mature virus structure, and the gp24 crystal structure. This hexamer (Fig. 5*C*) was fitted into the cryo-EM reconstruction giving a pseudoatomic model of the capsid shell minus gp hoc, gp soc, and the portal protein.

In the pseudoatomic model of the T4 capsid, the center-to-center

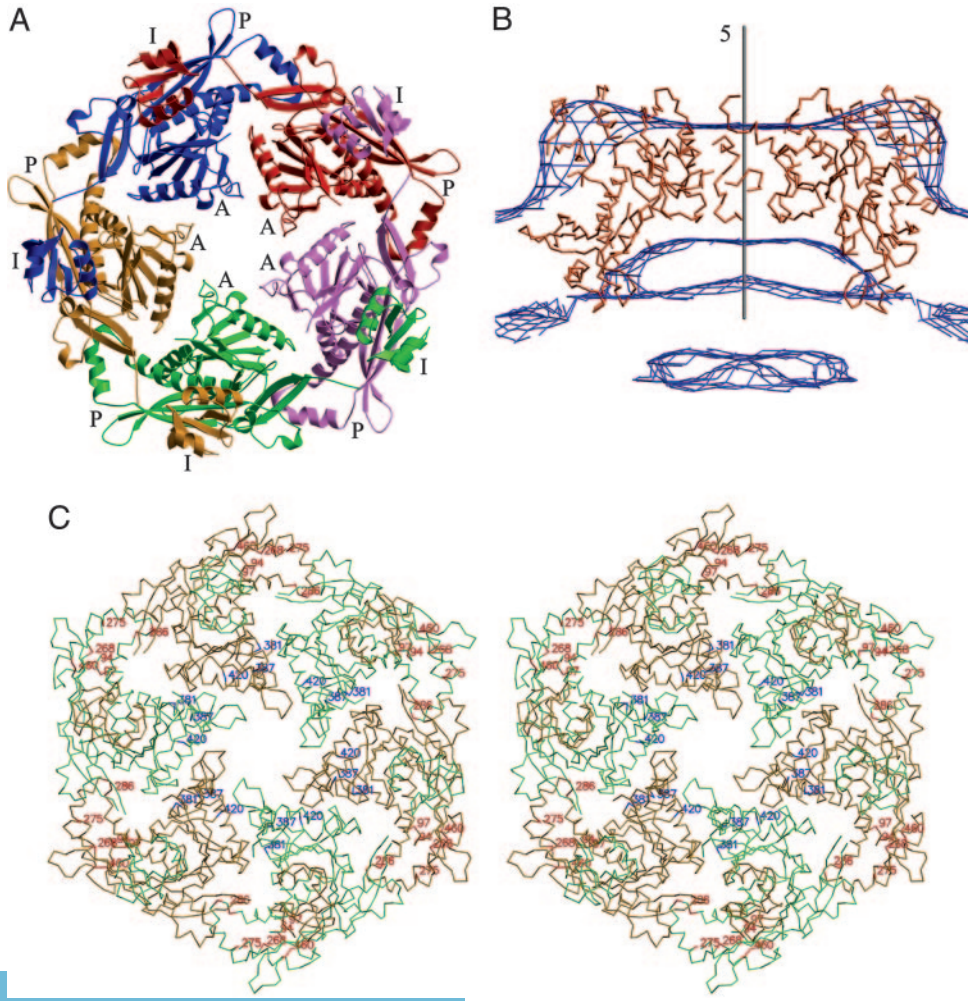

Fig. 5. The pentamer and hexamer structures that form the building blocks of the T4 capsid are shown. (*A*) Ribbon drawing of a gp24* pentamer determined by superimposing the gp24* HK97-like domain structure onto an HK97 pentamer. Each of the five molecules is shown in a different color. The P, A, and insertion (I) domains are shown. (*B*) Fit of the gp24* pentamer into the cryo-EM vertex density of the T4 capsid. (*C*) Stereo diagram of the C_α backbone showing one hexamer. Residues affecting the head morphology are shown in red, and residues that obviate the need for the gp24 vertex protein are shown in blue.

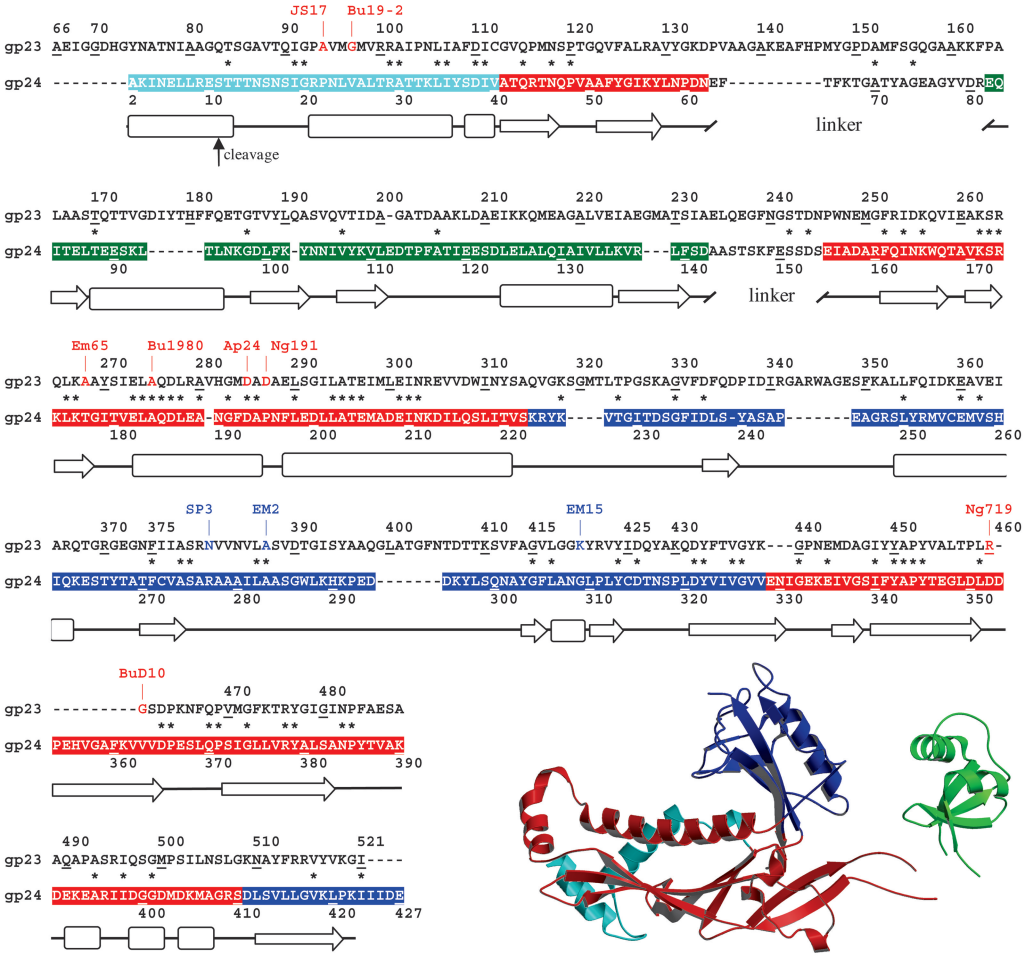

Fig. 6. Sequence alignment of gp24 and gp23 is shown. Residues belonging to the P, A, and insertion (I) domains are colored in red, blue, and green, respectively. Residues belonging to the N-terminal region are colored in cyan. Mutations in gp23 that affect head shape are marked in red within the sequence and mutations that obviate the need for gp24 are marked in blue within the sequence, using the nomenclature of Black *et al.* (3). Secondary structural elements are also indicated by arrows (β-strands) and open boxes (α-helices). The maturation cleavage site is indicated by an arrow. (*Inset*) A ribbon diagram in which the different domains are colored with the same code is shown.

distance between a pentamer and a neighboring hexamer is 135 Å, whereas the average distance between adjacent hexamers is 140 Å. The distance between the closest C_α atoms of a hexamer and a pentamer is ≈8 Å, whereas the distance between the closest C_α atoms between hexamers is ≈16 Å. The gp soc molecules were found to bind to the interfaces between hexamers, but not between hexamers and pentamers (2, 18). Nevertheless, the cryo-EM structures of the mature capsid and the gp soc$^-$ capsid are essentially the same except for the absence of the gp soc molecules. Thus, this additional space between the hexameric capsomers could be caused by differences in the structure of loops around the periphery of the gp24* and gp23* capsomers, and possibly error in the calibration of the electron microscope magnification. However, gp soc does stabilize the capsid against the effects of high pH or elevated temperatures (26, 27), presumably by serving as a molecular clamp to strengthen the contacts between gp23* capsomers.

Spatial Distribution of Mutations That Affect Capsid Morphology. Point mutations in gene 23 alter the shape of the head, producing petite and giant heads (28–30). These are mostly confined to the better conserved P domain (Figs. 5*C* and 6). As this domain determines the interactions between capsomers, it is consistent with the observation that mutations in this domain alter the shape of the capsid. Other mutations in gene 23 allow gp23 to substitute for gp24 at the vertices, in the absence of gp24 (6, 29). These mutations are positioned mostly in the A domain where they are likely to affect the association of monomers into pentamers or hexamers or both. As a special vertex protein probably has evolved more recently (6), and as these changes are now demonstrated to be mostly in the A domain, the A domain must have accepted more changes in the course of evolution to allow the specialization of gp24.

Evolution of Tailed Bacteriophages. The major capsid protein of the P22 (31), ϕ29 (32), and T4 (this study) tailed phages have been shown to have HK97-like folds. However, whereas HK97 uses crosslinking and the chain-mail mechanism for capsid stabilization, other phages have an additional protein domain that is probably used to achieve the same result. Furthermore, the tailed phage λ has many characteristics in common with HK97, and these characteristics are likely to extend to the fold of the major capsid protein (33–35). T4 not only has an insertion domain, but also uses the capsid protein gp soc to bind adjacent capsomers together, analogous to the glue proteins in large icosahedral dsDNA viruses whose capsomer structures are based on a pseudohexameric arrangement of jellyroll domains (25).

Apart from the similarity of their capsid protein structures, there are other similarities that point to a common ancestor for the assembly and structure of dsDNA phage heads. In particular, these viruses have a portal protein at one vertex that has been shown to be structurally similar for ϕ29 (36) and SPP1 (37), and probably also for human herpes simplex virus (38). Furthermore, they possess similarly structured and functioning scaffolding proteins. Although absent from the mature virus, these proteins participate in determining the head shape during assembly. Similar structures of scaffolding proteins have been discovered for the T4 (39), P22 (40), and ϕ29 (41) tailed phages.

It has long been recognized that 3D structure is conserved over much longer timespans than nucleotide or amino acid sequences

(42). Thus, the large number of structural and functional similarities of tailed phage capsids makes it probable that these and the capsids of other related viruses have diverged from a common ancestral source. Considering the high abundance of phages on Earth (43), the HK97 capsid fold must be one of the most frequently occurring structures in the Earth's biomass.

The HK97-like motif creates perfectly symmetric hexameric capsomers separated by ≈140 Å (2, 8). In contrast, numerous other types of icosahedral viruses use two repeated copies of the jellyroll motif within a single polypeptide to create trimers with pseudohexameric symmetry that are separated by ≈73 Å (44–47). The walls of capsids built with the HK97 motif are invariably quite thin, with a width of ≈24 Å, whereas capsids using the jellyroll motif are much thicker with a width of ≈40 Å.

The concept of quasiequivalence developed by Caspar and Klug (48) was based on having exact hexagonal lattices folding into icosahedra with infinitely thin walls, thus approximating the case of the HK97-based capsids. The jellyroll-based capsomers, as a result of their large thickness, will require greater flexibility to accommodate themselves to quasiidentical environments made possible by the departure from true hexagonal symmetry. However, these thick-walled capsids should produce larger binding surfaces between subunits and, hence, greater stability. Thus, the thin-walled HK97-based capsomers need strengthening by such devices as crosslinking, insertion domains, and glue proteins, such as gp soc. Both classes of viruses, however, require further stabilization for very large capsids, encapsulating large genomes, by evolving vertex proteins from their major capsid protein to accommodate the accumulation of strain analogous to the buttresses in Gothic churches.

We dedicate this article to the memory of Eduard Kellenberger, a pioneer of bacteriophage T4 morphogenesis studies and the mentor and friend of three of us (A.C.S., L.W.B., and V.V.M.). We thank Paul Chipman and Anthony Battisti for providing the cryo-EM data from which the original 3D T4 head was constructed (2) and used here for fitting the crystal structure of gp24 and the homology-built structure of gp23; Keith Brister and his staff at the Advanced Photon Source BioCARS beamline for their help with the SeMet-related crystals; Zbigniew Dauter, Miroslawa Dauter, and Peter Zwart for their help with the Au derivative at the Brookhaven National Laboratory; Marc Morais for helpful discussions; Lyuben Marekov, Shyh-Ing Jang, and William Idler for help with protein analysis, vector construction, and protein purification; and Sheryl Kelly, Cheryl Towell, and Sharon Wilder for help in the preparation of the manuscript. The work was supported by grants from the National Science Foundation (to M.G.R.), the Human Frontier Science Program (to M.G.R., V.V.M., and Fumio Arisaka), the National Institutes of Health (to L.W.B.), and the Howard Hughes Medical Institute (to V.V.M.).

1. Conway, J. F., Wikoff, W. R., Cheng, N., Duda, R. L., Hendrix, R. W., Johnson, J. E. & Steven, A. C. (2001) *Science* **292**, 744–748.
2. Fokine, A., Chipman, P. R., Leiman, P. G., Mesyanzhinov, V. V., Rao, V. B. & Rossmann, M. G. (2004) *Proc. Natl. Acad. Sci. USA* **101**, 6003–6008.
3. Black, L. W., Showe, M. K. & Steven, A. C. (1994) in *Molecular Biology of Bacteriophage T4*, ed. Karam, J. D. (Am. Soc. Microbiol., Washington, DC), pp. 218–258.
4. Driedonks, R. A., Engel, A., tenHeggeler, B. & van Driel, R. (1981) *J. Mol. Biol.* **152**, 641–662.
5. Leiman, P. G., Kanamaru, S., Mesyanzhinov, V. V., Arisaka, F. & Rossmann, M. G. (2003) *Cell. Mol. Life Sci.* **60**, 2356–2370.
6. McNicol, L. A., Simon, L. E. & Black, L. W. (1977) *J. Mol. Biol.* **116**, 261–283.
7. van Driel, R. (1977) *J. Mol. Biol.* **114**, 61–72.
8. Wikoff, W. R., Liljas, L., Duda, R. L., Tsuruta, H., Hendrix, R. W. & Johnson, J. E. (2000) *Science* **289**, 2129–2133.
9. Duda, R. L. (1998) *Cell* **94**, 55–60.
10. Ramakrishnan, V., Finch, J. T., Graziano, V., Lee, P. L. & Sweet, R. M. (1993) *Nature* **362**, 219–223.
11. Jancarik, J. & Kim, S.-H. (1991) *J. Appl. Crystallogr.* **24**, 409–411.
12. Otwinowski, Z. & Minor, W. (1997) *Methods Enzymol.* **276**, 307–326.
13. Terwilliger, T. C. & Berendzen, J. (1999) *Acta Crystallogr. D* **55**, 849–861.
14. Wang, B. C. (1985) *Methods Enzymol.* **115**, 90–112.
15. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., *et al*. (1998) *Acta Crystallogr. D* **54**, 905–921.
16. McRee, D. E. (1999) *J. Struct. Biol.* **125**, 156–165.
17. Helgstrand, C., Wikoff, W. R., Duda, R. L., Hendrix, R. W., Johnson, J. E. & Liljas, L. (2003) *J. Mol. Biol.* **334**, 885–899.
18. Iwasaki, K., Trus, B. L., Wingfield, P. T., Cheng, N., Campusano, G., Rao, V. B. & Steven, A. C. (2000) *Virology* **271**, 321–333.
19. Lata, R., Conway, J. F., Cheng, N., Duda, R. L., Hendrix, R. W., Wikoff, W. R., Johnson, J. E., Tsuruta, H. & Steven, A. C. (2000) *Cell* **100**, 253–263.
20. Holm, L. & Sander, C. (1993) *J. Mol. Biol.* **233**, 123–138.
21. Ikegami, T., Okada, T., Hashimoto, M., Seino, S., Watanabe, T. & Shirakawa, M. (2000) *J. Biol. Chem.* **275**, 13654–13661.
22. van Aalten, D. M. F., Synstad, B., Brurberg, M. B., Hough, E., Riise, B. W., Eijsink, V. G. H. & Wierenga, R. K. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 5842–5847.
23. Nicholls, A., Sharp, K. A. & Honig, B. (1991) *Proteins* **11**, 281–296.
24. Dasgupta, S., Iyer, G. H., Bryant, S. H., Lawrence, C. E. & Bell, J. A. (1997) *Proteins* **28**, 494–514.
25. Abrescia, N. G. A., Cockburn, J. J. B., Grimes, J. M., Sutton, G. C., Diprose, J. M., Butcher, S. J., Fuller, S. D., San Martin, C., Burnett, R. M., Stuart, D. I., *et al*. (2004) *Nature* **432**, 68–74.
26. Ishii, T., Yamaguchi, Y. & Yanagida, M. (1978) *J. Mol. Biol.* **120,** 533–544.
27. Steven, A. C., Greenstone, H. L., Booy, F. P., Black, L. W. & Ross, P. D. (1992) *J. Mol. Biol.* **228**, 870–884.
28. Lane, T. & Eiserling, F. A. (1990) *J. Struct. Biol.* **104**, 9–23.
29. Doermann, A. H., Pao, A. & Jackson, P. (1987) *J. Virol.* **61,** 2823–2827.
30. Mooney, D. T., Stockard, J., Parker, M. L. & Doermann, A. H. (1987) *J. Virol.* **61**, 2828–2834.
31. Jiang, W., Li, Z., Zhang, Z., Baker, M. L., Prevelige, P. E., Jr., & Chiu, W. (2003) *Nat. Struct. Biol.* **10**, 131–135.
32. Morais, M. C., Fisher, M., Kanamaru, S., Przybyla, L., Burgner, J., Fane, B. A. & Rossmann, M. G. (2004) *Mol. Cell* **15**, 991–997.
33. Campbell, A. M. (1994) *Annu. Rev. Microbiol.* **48**, 193–222.
34. Hendrix, R. W., Roberts, J. W., Stahl, F. W. & Weisberg, R. A. (1983) *Lambda II* (Cold Spring Harbor Lab. Press, Plainview, NY).
35. Yang, F., Forrer, P., Dauter, Z., Conway, J. F., Cheng, N., Cerritelli, M. E., Steven, A. C., Plückthun, A. & Wlodawer, A. (2000) *Nat. Struct. Biol.* **7,** 230–237.
36. 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., *et al*. (2000) *Nature* **408,** 745–750.
37. Orlova, E. V., Gowen, B., Droge, A., Stiege, A., Weise, F., Lurz, R., van Heel, M. & Tavares, P. (2003) *EMBO J.* **22**, 1255–1262.
38. Trus, B. L., Cheng, N., Newcomb, W. W., Homa, F. L., Brown, J. C. & Steven, A. C. (2004) *J. Virol.* **78**, 12668–12671.
39. Mesyanzhinov, V. V., Sobolev, B. N., Marusich, E. I., Prilipov, A. G. & Efimov, V. P. (1990) *J. Struct. Biol.* **104**, 24–31.
40. Tuma, R., Parker, M. H., Weigele, P., Sampson, L., Sun, Y., Krishna, N. R., Casjens, S., Thomas, G. J., Jr., & Prevelige, P. E., Jr. (1998) *J. Mol. Biol.* **281,** 81–94.
41. Morais, M. C., Kanamaru, S., Badasso, M. O., Koti, J. S., Owen, B. A. L., McMurray, C. T., Anderson, D. L. & Rossmann, M. G. (2003) *Nat. Struct. Biol.* **10**, 572–576.
42. Rossmann, M. G., Moras, D. & Olsen, K. W. (1974) *Nature* **250**, 194–199.
43. Hendrix, R. W. (2002) *Theor. Popul. Biol.* **61**, 471–480.
44. Benson, S. D., Bamford, J. K. H., Bamford, D. H. & Burnett, R. M. (2004) *Mol. Cell* **16**, 673–685.
45. Nandhagopal, N., Simpson, A. A., Gurnon, J. R., Yan, X., Baker, T. S., Graves, M. V., Van Etten, J. L. & Rossmann, M. G. (2002) *Proc. Natl. Acad. Sci. USA* **99**, 14758–14763.
46. Simpson, A. A., Leiman, P. G., Tao, Y., He, Y., Badasso, M. O., Jardine, P. J., Anderson, D. L. & Rossmann, M. G. (2001) *Acta Crystallogr. D* **57**, 1260–1269.
47. Yan, X., Olson, N. H., Van Etten, J. L., Bergoin, M., Rossmann, M. G. & Baker, T. S. (2000) *Nat. Struct. Biol.* **7**, 101–103.
48. Caspar, D. L. D. & Klug, A. (1962) *Cold Spring Harbor Symp. Quant. Biol.* **27,** 1–24.