© 2000 Nature America Inc. • http://structbio.nature.com

Novel fold and capsid-binding properties of the λ-phage display platform protein gpD

Fan Yang[1,2], Patrik Forrer[1,3], Zbigniew Dauter[4], James F. Conway[5,6], Naiqian Cheng[5], Mario E. Cerritelli[5], Alasdair C. Steven[5], Andreas Plückthun[3] and Alexander Wlodawer[2]

The crystal structure of gpD, the capsid-stabilizing protein of bacteriophage λ, was solved at 1.1 Å resolution. Data were obtained from twinned crystals in space group $P2_1$ and refined with anisotropic temperature factors to an R-factor of 0.098 (R_{free} = 0.132). GpD (109 residues) has a novel fold with an unusually low content of regular secondary structure. Noncrystallographic trimers with substantial intersubunit interfaces were observed. The C-termini are well ordered and located on one side of the trimer, relatively far from its three-fold axis. The N-termini are disordered up to Ser 15, which is close to the three-fold axis and on the same side as the C-termini. A density map of the icosahedral viral capsid at 15 Å resolution, obtained by cryo-electron microscopy and image reconstruction, reveals gpD trimers, seemingly indistinguishable from the ones seen in the crystals, at all three-fold sites. The map further reveals that the side of the trimer that binds to the capsid is the side on which both termini reside. Despite this orientation of the gpD trimer, fusion proteins connected by linker peptides to either terminus bind to the capsid, allowing protein and peptide display.

Bacteriophage λ (ref. 1) has a linear double-stranded DNA (dsDNA) genome contained within an icosahedral capsid of triangulation number T = 7, and a flexible, noncontractile tail (reviewed in ref. 2). The original *Escherichia coli* strain, K12, from which almost all laboratory strains have been derived, harbored a lysogenic λ-genome[3]. Although this genome is removed in most strains now in use, its finding prompted the use of λ as an experimental system. Phage λ has been extensively studied as a model system for gene regulation — notably for its switch between the lysogenic and lytic states[4] — as well as for virus assembly[5]. It has also become an important cloning vehicle[6] because of its capacity to carry large DNA fragments and its high infectivity. The full sequence of wild type (wt) λ showed that its 48.5 kbp genome codes for at least 46 genes[7]. Recently, λ has been adapted for phage display, providing a valuable alternative to filamentous phages[8–10]. Both the head protein, gpD, and the tail protein, gpV, have been used for this purpose. Because λ assembles in the cytoplasm, it is particularly well suited for the display of cytoplasmic proteins, and its capability for multivalent display allows for selection based on weak interactions.

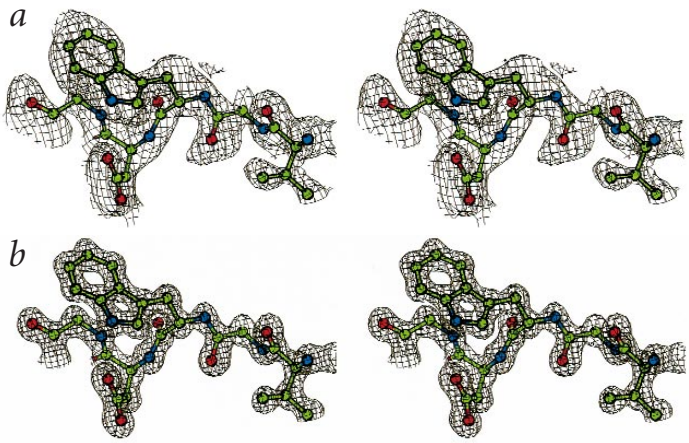

Fig. 1 Electron density maps for gpD. ***a***, Stereo view of the MAD-phased electron density map calculated from Se-Met data collected at the NSLS beamline X9B (Table 1). MADSYS[44] and DM[45] were used for phase calculation and solvent flattening. The electron density contoured at 1σ level and the final refined coordinates of residues 47–51 are shown. ***b***, Stereo view of the same region covered by the final $2F_o - F_c$ map contoured at the 2σ level.

Phage λ assembles its head and tail on separate, convergent pathways. Unlike small icosahedral viruses, whose head proteins co-assemble with their nucleic acid molecules, dsDNA viruses with large genomes (40–200 kbp), such as phage λ, assemble their heads in two steps: prohead assembly and DNA packaging. The latter event is accompanied by a large scale, irreversible, conformational change of the prohead (the expansion transformation) that increases its volume and stability. The virion is then completed by attachment of the tail to the mature head[11–14].

The shell of the λ-prohead is composed mainly of gpE (40 kDa). DNA packaging is accompanied by expansion of the prohead from a rounded to an icosahedral shape and the subsequent binding of gpD[12], a small (11.4 kDa) cysteine-free protein of 109 amino acids (not counting the initial methionine, which is not present in the mature protein). GpD can be efficiently expressed in *E. coli* by itself or with a variety of fusion partners appended to the C-terminus of the molecule[15]. The mature λ-head contains 415 copies of gpE and 405–420 copies of gpD.

[1]These authors contributed equally to this work. [2]Macromolecular Crystallography Laboratory, Program in Structural Biology, National Cancer Institute, Frederick Cancer Research and Development Center, Frederick, Maryland 21702, USA. [3]Biochemisches Institut, Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland. [4]NCI-FCRDC and NSLS, Brookhaven National Laboratory, Bldg. 725A-X9, Upton, New York 11973, USA. [5]Laboratory of Structural Biology, National Institute of Arthritis, and Musculoskeletal and Skin Diseases, Bethesda, Maryland 20892, USA. [6]Current address: Institut de Biologie Structurale, 41 rue Jules Horowitz, 38027 Grenoble Cedex 1, France.

Correspondence should be addressed to A.P. *email: plueckthun@biocfebs.unizh.ch*

© 2000 Nature America Inc. • http://structbio.nature.com

© 2000 Nature America Inc. • http://structbio.nature.com

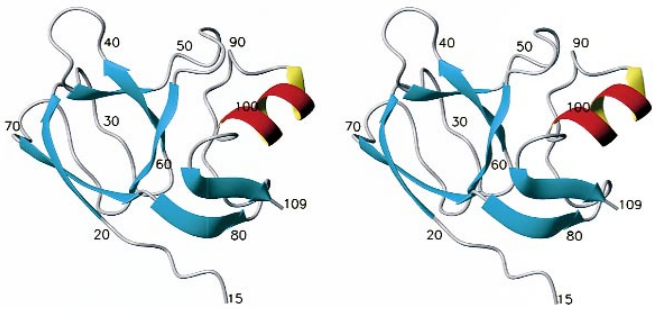

Fig. 2 Stereo view of the structure of the gpD monomer. The starting N-terminal residue is Ser 15, because the first 14 residues are disordered and their positions cannot be determined. The three-fold axis runs from top to bottom, the capsid-distal side is at the top ('top side') and the capsid-proximal face is at the bottom ('bottom side'). Figure prepared with MOLMOL[21].

Visualized by cryo-electron microscopy (EM) at ~34 Å resolution[11], the hexamers and pentamers of gpE were seen to have cartwheel-like structures with skewed arms radiating from central hubs. 'Thimble-shaped' protrusions overlying the trigonal sites, where three arms from neighboring capsomers meet, were interpreted as trimers of gpD, consistent with earlier inferences from negative-staining EM[16].

The binding sites for gpD are created or exposed only after the prohead expands[16,17]. Although gpD is not needed for prohead assembly, it is essential if a full-length genome is to be stably accommodated within the head. In contrast, a gpD$^-$ virus can only retain a DNA molecule of ~80% of wild type length. Furthermore, the mutant phage particles have to be stabilized by the addition of magnesium ions and are highly sensitive to EDTA[18]. These results support the proposed function of gpD in stabilizing the λ-head against the pressure imposed by packaged DNA[17,18].

We have determined the crystal structure of gpD with the aim of achieving a better understanding of how gpD functions as a display platform and to appreciate its further potential in this role, as well to elucidate the mechanism by which it stabilizes the expanded capsid. The same trimeric structure was visualized in two crystal forms. Because prior[17] and current evidence (see below) indicated that in solution, at concentrations up to millimolar, gpD is a monomer and not a trimer, we investigated its organizational state as bound to the capsid surface by cryo-EM. A three-dimensional density map of the mature, empty capsid at 15 Å resolution revealed gpD trimers apparently identical to those seen in the crystals. This map also disclosed the pattern of interactions between gpD and the underlying gpE molecules at the three-fold sites of the capsid surface lattice as well as the locations of the N- and C-termini of gpD used in peptide and protein display.

Structure of the gpD monomer

Crystals of gpD were obtained in two space groups, $P2_1$ and C2 (Table 1); the former were used for the initial structure determination. The crystal structure was determined by the method of multiwavelength anomalous diffraction (MAD) for a selenomethionine (Se-Met) derivative of the protein, and was refined to a final R of 0.098 (R_{free} = 0.132) at 1.1 Å resolution (Table 2). The experimentally phased maps are of high quality (Fig. 1*a*) and could be traced without difficulty. However, of the 109 residues in gpD, the first 14 are disordered and not visible on any of the three molecules in the asymmetric unit. The position of Ser 15 can be observed, but its conformation cannot be determined accurately. The secondary structure is quite irregular but nevertheless well defined, made up of ~65% coils and loops with several short β-strands and only one α-helix (Fig. 2). The strands (residues 21–25, 36–39, 46–48, 57–61, 71–75, 77–80, and 105–108, as assigned by the program RIBBONS[19]), are distributed over the entire primary structure of the protein. None of the strands is longer than five residues, and they make up only ~27% of the total sequence. Strands 1 and 5, 2 and 3, and the C-terminal ends of strands 4 and 5 form short antiparallel β-sheets, whereas strands 6 and 7 form a short parallel β-sheet. All have only a few interstrand hydrogen bonds (Fig. 3). A short and slightly distorted α-helix (eight residues) is located near the C-terminus. The overall packing of the gpD monomer is very tight and without any internal cavities. Its core is hydrophobic, formed by residues Leu 27, Ala 31, Pro 36, Leu 37, Trp 49, Ala 57, Ile 60, Leu 61, Ala 65, Leu 72, Phe 80, Val 85, Trp 87, Ala 100, Phe 101, Ile 106, and Ile 108. The molecule consists of a single domain with two identifiable segments consisting of residues 15–76 and 77–109, respectively (Fig. 3). There is only one main chain interaction between the two segments in addition to a few side chain interactions. Although the 14 N-terminal residues are disordered, the C-terminal residues are fully ordered and the last residue, Val 109, forms short hydrogen bonds with the amide nitrogen of Tyr 82 and the side chain of Arg 98. A data base search using the program DALI[20] did not identify any other pro-

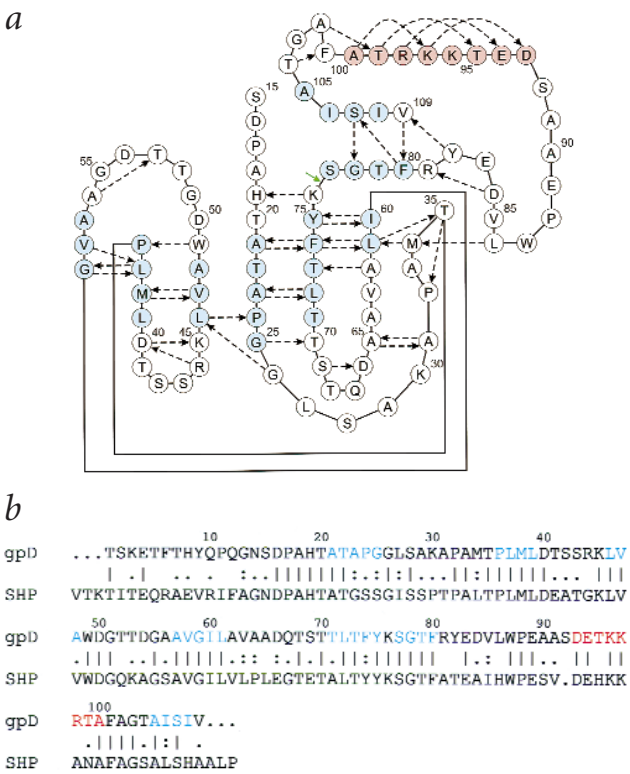

Fig. 3 Primary and secondary structures of gpD. ***a***, Schematic diagram of the main chain interactions of the gpD monomer. The directions of the dashed arrows represent the main chain hydrogen bonds from the amide nitrogen atoms (donors) to the carbonyl oxygen atoms (acceptors). The seven β-strands (blue) include residues 21–25, 36–39, 46–48, 57–61, 71–75, 77–80 and 105–108. The α-helix (red) includes only residues 93–100. The green arrow indicates the segment border. ***b***, Sequence alignment between gpD and the homologous protein SHP from the lambdoid phage 21.

© 2000 Nature America Inc. • http://structbio.nature.com

© 2000 Nature America Inc. • http://structbio.nature.com

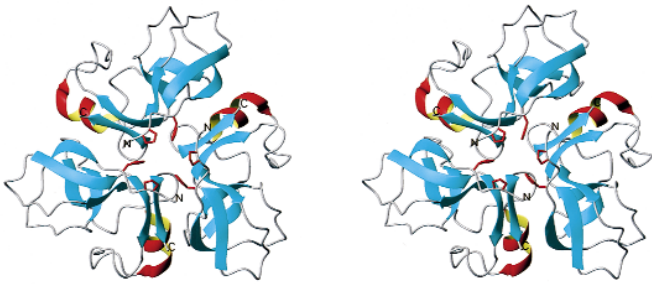

Fig. 4 Stereo view of the gpD trimer viewed from the bottom side. Both termini are on the same side of the trimer. The three C-termini are near the edge of the trimer and have a stable conformation, whereas the N-termini converge to the three-fold center of the trimer and are disordered up to residue 14. Also shown are Pro 17 and His 19, which form a small ring near the three-fold axis. Figure prepared with MOLMOL[21].

teins with similar three-dimensional structures, implying that the gpD structure (Figs 2, 3) describes a new fold.

Quaternary structure of gpD

Noncrystallographic trimers were first observed in the crystals grown in space group $P2_1$. Subsequently, identical trimers were also observed in the C2 crystals. The asymmetric unit of the $P2_1$ crystals contains one trimer: that of the C2 crystals, two trimers related by noncrystallographic symmetry. However, their packing differs markedly between the two crystal forms, giving different intertrimer interactions.

The differences in the atomic positions between the independent molecules in each trimer are small. The root mean square (r.m.s.) deviations of main chain atoms fitted between each pair of monomers in the $P2_1$ trimer are 0.18 Å, 0.16 Å and 0.23 Å, respectively. When the positional deviations and the averaged isotropic temperature factors of the main chain atoms are plotted as a function of residue number, they clearly correlate with the temperature factors, the indices of positional stability (data not shown). Plots of both parameters show high peaks in three segments around residues 28, 41 and 68, besides the disordered N-terminal residues. These three segments correspond to sharp turns in the structure (Fig. 3). Conversely, all seven β-strands and the single α-helix have low positional deviations and low temperature factors.

The gpD trimer has the shape of a hollow triangle with rounded sides (Fig. 4). The thickness at the center is ~28 Å and the length of a triangular edge is ~51 Å. The three monomers are tightly associated because of interactions among surface residues. Except for some hydrogen bonds, the predominant interactions are hydrophobic, involving residues Ala 21, Ala 23, Pro 24, Leu 39, Leu 46, Val 58, Phe 74 and Ala 102. The accessible surface areas of the monomer and trimer are ~5,500 $Å^2$ and ~12,700 $Å^2$, respectively. Thus, ~1,270 $Å^2$ per subunit (~25% of the molecular surface) is buried upon trimerization.

It is noteworthy that all seven β-strands are located near the three-fold axis, although they are distributed over the entire primary structure (Figs 3, 4). The β-strands are close to one side of the trimer, which is slightly convex; we call this the bottom side because it binds to the capsid. The three copies per trimer of the first observed N-terminal residue (Ser 15) converge to the three-fold axis of the trimer, also on the bottom side (Fig. 4). Likewise, the C-termini are on the bottom side, but they are located near the edge of the trimer. On the top side, which is slightly concave, there is a cavity at the center of the trimer that is solvent accessible. Comparing the two surfaces, the bottom side is more neutral and hydrophobic than the top, as calculated with MOLMOL[21] (Fig. 5*a*,*b*). Ten hydrophobic residues (Pro 17, Ala 18, Pro 32, Ala 33, Ala 62, Val 63, Ala 64, Phe 80, Leu 86 and Val 109) are located on the bottom side. Most of them are conserved in SHP, the gpD homolog of phage 21 (Figs 3*b*, 5*c*,*d*)[22]. This conservation is in contrast to the substantially lower similarity between residues found on the top side of the gpD trimer. Another interesting feature on the bottom side is a small ring near the three-fold axis, formed by two ring-containing residues from each monomer, Pro 17 and His 19 (Fig. 4). The members of this ring are evenly separated with nearest-neighbor spacings of 3.5–4.2 Å.

The gpD crystal trimer binds the λ-capsid

Although gpD is clearly trimeric in both crystal forms, it appears to be monomeric in solution (see below). To address the question of its oligomeric status *in situ*, we performed a cryo-EM analysis of the capsid. In a previous study by this technique[11], the limited resolution of the experiments (34 Å) precluded settling the question of whether the three inferred subunits in each 'thimble-shaped' protrusion made contact with each other (as opposed to being close but not touching, and held in place by interactions with three underlying gpE molecules).

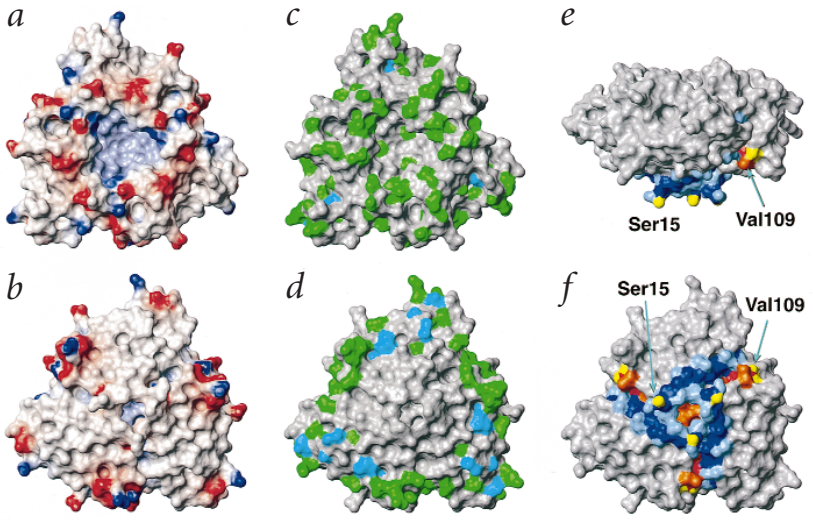

Fig. 5 The molecular surface of the gpD trimer. ***a***,***b***, The surface charge. The surface of the bottom side (*b*) is relatively smooth and uniformly neutral (white), whereas the top side (*a*) has randomly distributed charged residues and a cavity at the three-fold center that may be accessible to solvent. Red indicates negative potential; blue, positive potential. ***c***,***d***, Similarity of gpD and SHP. Residues on the surface of the bottom side (*d*) of the gpD trimer are more conserved than those on the top side (*c*). Identical residues are gray, similar residues (side chains only) are cyan, and different residues (side chains only) are green. ***e***,***f***, Location of the termini in the structure. The N-terminal residues are light blue (side chain atoms) and dark blue (other atoms), whereas the C-terminal residues are colored orange (side chain atoms) and dark red (other atoms). The N-terminal nitrogen atom and the C-terminal oxygen atoms are colored yellow. Side view (*e*), top view (*f*). Figure prepared with MOLMOL[21].

© 2000 Nature America Inc. • http://structbio.nature.com

© 2000 Nature America Inc. • http://structbio.nature.com

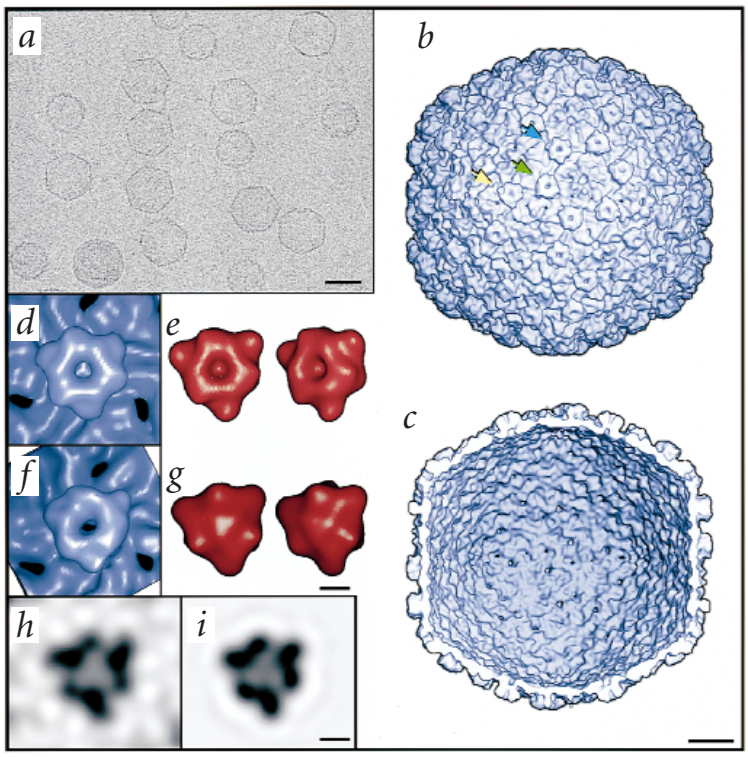

Fig. 6 Electron micrographs of phage λ capsids. ***a***, Field containing both mature capsids (hexagonal shape in cross section) and smaller, rounder, procapsids. The round, denser, particle (bottom left) is a contaminating vesicle. This field represents an area relatively well populated with mature capsids. Bar, 500 Å. ***b***,***c***, Outer and inner surfaces of a reconstruction at 15 Å resolution of the mature capsid, as viewed along an axis of two-fold symmetry. Bar, 100 Å. The 'thimbles' of density at the trigonal sites, identified by Dokland and Murialdo[11] as containing gpD by comparing D^+ and D^- capsids, are resolved into ringlike trimers (arrows). There are three classes of quasi-equivalent trimers, examples of which are indicated by yellow, green and blue arrows, respectively. ***d–g*** Enlargements of the gpD trimer on the three-fold axis (*d*; blue arrow in *b*) and closest to the five-fold axis (*f*; yellow arrow in *b*), are compared with the top (*e*) and bottom (*g*) sides of the crystal structure, as limited to the same (15 Å) resolution. The trimers in (*e*) and (*g*) are viewed along their symmetry axis (at left) and as rotated through a small angle (at right: ~10° about the vertical axis). This comparison makes it clear that the top side of the trimer faces outward and, conversely, the bottom side binds to the capsid. Also shown are 1.8 Å thick sections through the trimer as portrayed in ***h***, the cryo-EM density map and ***i***, the resolution-limited crystal structure. Both sections reveal the division of the trimer into its component subunits. Bars in (*g*, *i*) = 10 Å.

To obtain capsid images unobscured by the coprojection of encapsidated DNA, which is not icosahedrally ordered, we chose to work with the CosQ mutant[23]. Under nonpermissive conditions, about one third of these capsids embark on but are unable to complete DNA packaging. Upon lysis of *E. coli*, the partially packaged capsids lose their DNA and may be recovered as empty capsids, together with the remaining particles, which are still at the prohead stage (Fig. 6*a*). Although our yields of particles were lower than optimal for cryo-EM analysis, by pooling data from 12 focal pairs of sparsely populated micrographs, we were able to calculate a density map from 380 particles to 15 Å resolution (Fig. 6*b*,*c*).

The T = 7 surface lattice of the λ-capsid has three nonequivalent trigonal sites: the global three-fold axis at the center of each icosahedral facet (20 in all), and 60 copies each of two kinds of local three-fold axes (one surrounded by one penton and two hexons, the other by three hexons). Each of these sites (Fig. 6*b*, arrows) is occupied by a triangular complex with rounded sides and a central dimple, whose dimensions and shape are strongly reminiscent of the crystal trimer of gpD. To confirm this relationship and to determine the orientation of the trimer (that is, which side is on the external surface), we band-limited the crystal structure to 15 Å resolution, and calculated surface renderings of it in both orientations (Fig. 6*d*,*e*). The top side has a central dimple that is a near-perfect match for the EM structure (Fig. 6*d*,*e*), whereas the bottom side has a different topography, with a low plateau, even without accounting for the first 14 residues. The other features of the crystal trimer, as viewed from the top side, including its slight but clearly expressed vorticity and the three triangular corner densities set at a lower level than the mass surrounding the central dimple, are faithfully reproduced in the cryo-EM map. Similar comparisons made for the gpD trimers at the other two sites indicated that they have essentially the same structure except for some slight departures from three-fold symmetry, which presumably reflect the adaptation of the trimer to local curvature — that is, meeting the requirements of quasi-equivalence at these sites.

To further confirm the compatibility of the crystal structure and the cryo-EM density map, we modeled the high-resolution structure of the trimer into the density map and obtained an excellent fit (Fig. 7*a*,*b*). Thus, we conclude that (i) the trimer seen in both crystals represents the same form of the protein that is deployed on the capsid surface, and (ii) it binds to the capsid with the dimpled side (which we refer to as the top side) facing outward.

Characterization of gpD in solution

Purified gpD was studied by gel filtration, circular dichroism (CD) and NMR spectroscopy. The short-wavelength part (200–245 nm) of the CD spectrum was weak, showing no evidence for significant α-helical content (data not shown), in agreement with earlier CD measurements of wt-gpD[17]. Furthermore, the spectrum of wt-gpD resembled that of a random coil, consistent with the high percentage of irregular secondary structure elements found in the crystal structure. Preliminary NMR spectroscopy experiments indicated that soluble gpD is a well-structured globular protein. In addition, by measuring its proton relaxation time, a correlation time of ~7 ns was determined, which corresponds to an apparent molecular mass of gpD of 14,000 (ref. 24). Thus, gpD seems to be monomeric, even at the very high concentration of 1.3 mM used in these experiments (H. Iwai, pers. comm.). This result was confirmed by gel filtration and dynamic light scattering (data not shown) and is in agreement with the earlier observation[17] that wt-gpD is monomeric at a concentration of 0.1 mM, as determined by gel filtration and analytical ultracentrifugation. This finding contrasts with the tight trimer of gpD found in the crystals and on the capsid surface.

Properties of other viral accessory proteins

GpD forms the same close knit trimer in both crystal forms and on the capsid. The interaction surfaces between gpD subunits are extensive, and in forming the trimer, some hydrophobic residues (Leu 39, Leu 46, Val 58 and Phe 74) that would be exposed to solvent in the monomer are buried. A distinctive feature of the trimer is the ring of three copies of Pro 17 and His 19 that forms the protruding 'plateau' on the bottom side (Fig. 4). In the monomer, Pro 17 and His 19 would be exposed to solvent, sug-

© 2000 Nature America Inc. • http://structbio.nature.com

© 2000 Nature America Inc. • http://structbio.nature.com

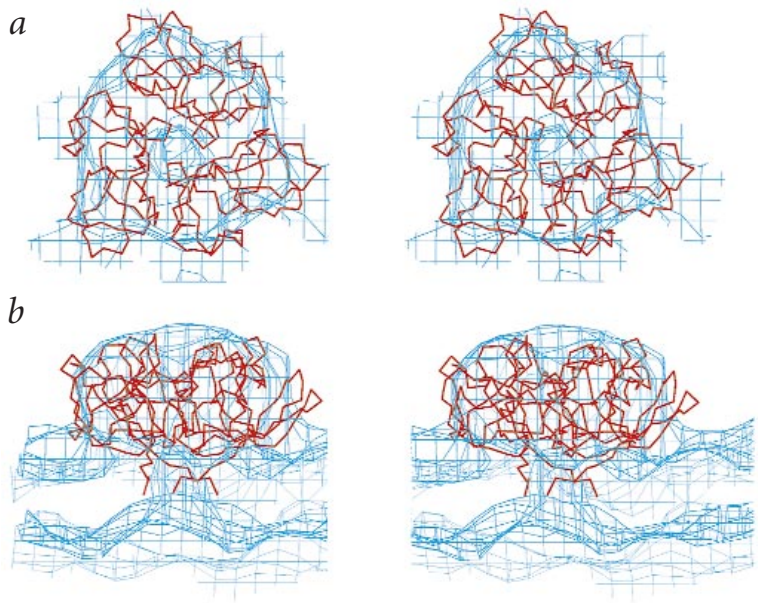

Fig. 7 Modeling of the crystal structure of the gpD trimer (red) into the portion of the cryo-EM density map (blue) that represents the gpD trimer at the three-fold axis of symmetry. Modeling was performed with the program O[46]. The trimer is viewed from ***a***, outside the capsid and ***b***, in a side view.

gesting that formation of this ring contributes significantly to trimerization. The affinity of the subunits must be low in solution as explained above, and to form trimers, they need either to be highly concentrated or to interact with gpE on the virion, which presents a trimerization template. Binding of gpD is found to be cooperative[17], which implies facilitated binding of the second and third gpD of each trimer. In addition, conformational changes may occur in gpD upon binding to gpE that further stabilize the trimeric conformation.

This situation of gpD parallels the cases of two other accessory proteins of dsDNA viruses: gp*soc* of phage T4 (10 kDa) and VP26 (12 kDa) of herpes simplex virus. Like gpD, these proteins are dispensable for assembly[25–27] and bind only to the transformed (expanded) prohead[28,29]. Moreover, gp*soc* also binds around the trigonal sites, whereas six copies of VP26 bind around the outer tip of each HSV hexon[30], where it forms a hexameric ring[31] but displays a monomer/dimer equilibrium in solution[32]. Gp*soc* further resembles gpD in having been developed as a display platform[33,34], and VP26 appears to have potential for such exploitation[32,35].

Interaction of gpD with the gpE capsid shell

Upon comparison of the crystal structure and the EM structure, it is evident that the 'bottom side' of the gpD trimer binds to the capsid (Figs 6, 7). In addition to their respective topographies, the two sides of the trimer differ in polarity and surface charge: the top side is relatively hydrophilic and shows a random distribution of positive and negative charges (Fig. 5*a*,*b*), whereas the bottom side is less charged and more hydrophobic. This property suggests that when the bottom side binds to the capsid, these hydrophobic groups are sequestered from solvent. Moreover, the bottom side is more conserved than the top side when compared with SHP, the gpD homolog from phage 21 (ref. 22), which has 50% amino acid identity to gpD (Figs 3*b*, 5*c*,*d*). On examining the distribution of altered residues, we observed that most substitutions occur around the edges or on the top side of the trimer. Among the most conserved features is the 'trimerization ring — that is, three copies of Pro 17 and His 19, forming the central 'plateau' of the bottom side. Because SHP can complement λ-phage devoid of gpD *in vivo*[36] and *in vitro* (Table 3), it seems reasonable that the more conserved surface is also the capsid-binding surface.

This orientation places both termini of gpD on the side proximal to the virus, with the visible N-terminus (Ser 15) near the center of the gpD/gpE interface. Such placement implies that the N-terminus up to Ser 15 is also involved in interactions with gpE (Figs 5*f*, 6). This proposition is supported by complementation data indicating that a truncated form of gpD, lacking the N-terminal 14 amino acids (gpDΔN1), fails to bind to the capsid (Table 3). This truncated variant is well expressed, highly soluble and monomeric in solution, as is full length gpD. The complementation assay was performed by measuring the concentrations of gpD derivatives at which the phage becomes resistant to exposure to EDTA and able to infect *E. coli* with normal titers[10,18]. A minimum concentration of wt-gpD (0.2 μM) is needed for capsid binding[17], which corresponds well to the concentration dependence for complementation of λD^- (Table 3). This finding shows that complementation is correlated with binding. Thus, it is plausible that gpDΔN1 was unable to complement because it failed to bind to the capsid. The complementation, when observed, was strongly dependent on the concentration of the gpD variant (data not shown). This strong concentration dependence is a sign of cooperative binding, as would be expected for a monomer in solution that forms a trimer on the capsid.

In contrast to gpDΔN1, normal complementation was found for SHP (Table 3). SHP has an N-terminal sequence that is completely unrelated to the corresponding sequence in gpD and is even three amino acids longer (Fig. 3*b*). This feature was unexpected, because the N-terminus has been implicated in capsid binding. Such binding is on the one hand abolished in the N-terminal deletion mutant, but is on the other hand almost identical between SHP and gpD.

Implications for gpD-based phage display

Significantly, all of the gpD fusions that we tested, including those with whole domains (up to 425 amino acids) fused to either terminus, were able to complement gpD^- deficiency (Table 3). Their similar threshold values for EDTA resistance imply that their binding to the capsid is very similar to that of wt-gpD. These experiments indicate that the fusion proteins (all of which have sizable linkers added to the displayed moieties; Table 3) can occupy most, if not all, gpD binding sites. If the observed resistance was conferred by a small amount of normal-sized gpD generated by proteolysis, much higher apparent thresholds should have been observed, because the breakdown product would have to be present at 0.1–0.2 μM, and much higher levels of the fusion protein would be needed to produce this amount. Our data do not allow the conclusion, however, that it will be possible in general to produce fully substituted gpD fusion phages *in vivo*, because a certain degree of proteolysis could possibly occur, generating some wt-gpD. Nevertheless, it appears that at least a highly multivalent gpD display is possible[9,10].

Because both N-terminal[9,10] and C-terminal[8,9] fusions have been successfully displayed, additions at either end of gpD do not prevent the fusion proteins from binding to the capsid. In this context, our finding that the termini of gpD are located on

© 2000 Nature America Inc. • http://structbio.nature.com

© 2000 Nature America Inc. • http://structbio.nature.com

Table 1 Data collection statistics[1]

	Se-Met gpD MAD (NSLS)	Se-Met gpD MAD (APS)	wt-gpD (NSLS)	wt-gpD
Space group	$P2_1$	$P2_1$	$P2_1$	C2
a (Å)	45.46	45.47	45.58	74.39
b (Å)	68.41	68.51	69.07	51.66
c (Å)	45.50	45.50	45.59	155.89
β (°)	104.35	104.29	104.32	95.80
R_{sym} (%)	3.0	3.5	4.6	5.5
Resolution (Å)	1.80	1.75	1.10	2.06
Completeness (%)	95.2	96.0	99.2	88.3
Total reflections	~261,000	~203,000	635,194	368,034
Unique reflections	~48,000 (F+ F- separated)	~52,000 (F+ F- separated)	109,996	32,775

[1]NSLS, Brookhaven National Laboratory; APS, Advanced Photo Source, Argonne National Laboratory.
[2]Pairs of Friedel-related reflections treated as two different reflections.

the capsid-proximal side of the trimer was unexpected. In particular, the visible N-terminus (Ser 15) is located close to the threefold axis (Fig. 4). The requirement that the fused portions be exposed on the outer surface is most readily met for the C-terminus, for which only a short linker is needed to reach the solvent (Fig. 5*e*,*f*). As for display of N-terminal fusions, there are two possibilities. (i) The additional density seen in the cryo-EM density map (Fig. 7) as stalks underlying the crystal trimer and connecting it to the capsid surface could be contributed by the N-termini, with the first residue on the outside of the stalk, suitably positioned to support display. This hypothesis allows the N-termini to become ordered through interaction with gpE. (ii) Alternatively, the N-termini and/or the linkers could somehow be threaded across the gpD/gpE interface from the near-axial location of Ser 15, so as to display the fusions on the outside. Both possibilities may be tested by cryo-EM difference mapping with the D^- capsid and other derivatives of gpD.

Methods

Cloning, expression, and purification of gpD. Bacteriophage λ gene D (ref. 7) was amplified by polymerase chain reaction (PCR) from λcI857Sam7 DNA (Roche Molecular Biochemicals), using the oligonucleotides 5'-ATCAGATCTG GTTCCATGGC GAGCAAAGAA AC and 5'-CATAGCTTTC ATTAAACGAT GCTGATTG as forward and reverse primers, respectively. The forward primer also exchanged the second codon of gene *D* to introduce an *Nco*I site at the expense of a Thr-Ala mutation. After digestion of the PCR fragment with *Nco*I and *Hin*dIII, it was subcloned into the *Nco*I and *Hin*dIII-linearized T7-promoter-based expression plasmid pTFT74 (ref. 37), resulting in the gpD expression plasmid pAT101. Protein D was expressed from this plasmid in BL21(DE3) (ref. 38) by induction with isopropyl-β-D-thiogalactopyranoside (IPTG) for 4 h at 37 °C. For Se-Met labeling of the protein, the construct was expressed in DL41(DE3) (ref. 39), using a synthetic growth medium containing seleno-L-methionine (Se-Met, Sigma)[40] by induction with IPTG for 4 h at 37 °C. Both recombinant proteins were expressed to high levels in a soluble form and could be purified from the cytoplasm by the same procedure, except that for the purification of the Se-Met-labeled protein all buffers were degassed and supplemented with 5 mM dithiothreitol (DTT) and 0.2 mM EDTA to prevent oxidation. Cells were harvested by centrifugation, resuspended in 50 mM Tris pH 9, 500 mM NaCl, and lysed by extensive sonication. After centrifugation, the supernatant was incubated at 50 °C for 15 min and the denatured proteins were removed by centrifugation. Ammonium sulfate was added to the sample to 2.4 M and the precipitate was removed. Additional ammonium sulfate was added to a final concentration of 3.0 M and the precipitate containing gpD was collected and resuspended in 10 mM Tris pH 9, 30 mM NaCl, and dialyzed against the same buffer. Chromatographic purification was achieved by perfusion chromatography on a BioCad system (PerSeptive Biosystems) at 20 °C. The sample was loaded on an anion exchange column (Poros HQ), and the flowthrough was collected, dialyzed against 50 mM sodium acetate pH 4.8, and loaded on a cation exchange column (Poros HS). GpD was eluted from the column with a gradient of 0–150 mM NaCl. Protein D-containing fractions were pooled, concentrated, dialyzed against 20 mM Tris pH 7.5, aliquoted and stored at -80 °C until use. Removal of the N-terminal methionine residue and incorporation of two Se-Met residues per gpD was confirmed by mass spectroscopy. The yield obtained from a 4-liter flask of *E. coli* culture was 100 mg of highly purified gpD.

Production of gpD variants. An expression plasmid for a N-terminal deletion mutant of gpD, gpDΔN1 (containing Ser 15–Val 109 of gpD and an additional N-terminal Gly), was constructed as described for gpD (see above), using corresponding PCR primers. The gene *D* homolog of phage 21, *SHP*, was PCR amplified from λ-21 hybrid 19 (kindly provided by M. Feiss[36]) and subcloned into pTFT74 as described for gene *D*. Purification of these proteins was similar to gpD. Both gpDΔN1 and SHP were found to be very well expressed, highly soluble, monomeric in solution as determined by gel filtration, and not prone to aggregation upon thermal denaturation. The construction and purification of GpHDL-cCrk (a protein D variant with an N-terminal His_6 tag fused to the C-terminal 21.4 kDa domain of cCrk) and gpHDL-JNK2 (a protein D variant with an N-terminal His_6 tag fused to the N-terminal 48 kDa domain of JNK2 were described earlier[15]. The GST-ATF2-gpD fusion protein (this pD variant contains the glutathione *S*-transferase (GST) from *Schistosoma japonicum* at its N-terminus bridged by the first 100 amino acids of ATF2) was produced from a pGEX (Amersham Pharmacia)-derived expression plasmid (P.F., unpublished data) and purified using a GST affinity column (GSH-Sepharose). This fusion protein is a dimer, as is wt-GST.

Crystallization and data collection. The gpD sample was concentrated to about 20 mg ml^{-1} in 20 mM Tris pH 7.5. Crystal Screen kit I (Hampton Research) was used for the initial screening. Needles and a large, bulky crystal were found at conditions 28 and 43. The final crystallization conditions of gpD were refined to 28% PEG 4000, 0.1 M Bis-Tris pH 6.5, 10% glycerol. Two crystal forms, $P2_1$ and C2, were found in the same crystallization droplets. The crystallization conditions for Se-Met gpD are slightly different, with 32% PEG 4000, 0.1 M Bis-Tris pH 6.5, 20 mM DTT and 10% glycerol. Although the shapes of wt-gpD crystals are usually irregular, those of the Se-Met gpD crystals are quite regular, with well-developed faces. The $P2_1$ crystals diffract to higher resolution and have smaller unit cell parameters than the C2 crystals, which makes them preferable for structure determination. Furthermore, the Se-Met derivative produced for MAD experiments could only be crystallized in the $P2_1$ form. Thus, the $P2_1$ crystals were used for the initial structure determination.

Two independent MAD data sets were collected from the Se-Met gpD crystals (Table 1). One data set was obtained using a four-element ADSC CCD detector on synchrotron beamline X9B at the NSLS, Brookhaven National Laboratory, while the other was obtained using a single-element ADCS CCD detector on the synchrotron beam-

Table 2 Statistics of structure solution and refinement

R_{cryst}	0.098
R_{free}	0.132
R.m.s. deviation of bonds (Å)	0.014
R.m.s. deviation of angle distances (Å)	0.031
R.m.s. chiral volumes ($Å^3$)	0.087
Protein atoms	2,088
Solvent atoms	309
Heteroatoms	12

© 2000 Nature America Inc. • http://structbio.nature.com

© 2000 Nature America Inc. • http://structbio.nature.com

line 14-BM-D of the Advanced Photon Source (APS), Argonne National Laboratory (Table 1). Both MAD data sets were collected at four wavelengths each, to the resolution 1.8 Å and 1.75 Å, respectively. A native gpD data set was collected for a $P2_1$ crystal to 1.1 Å resolution at beamline X9B in two passes, by collecting and merging high-resolution and low-resolution data. A data set for a C2 crystal was collected to 2.06 Å resolution on a Mar345 image detector using a Rigaku X-ray generator. Integration and scaling of the data were accomplished with HKL2000 (ref. 41).

Table 3 ***In vitro*** **complementation of λD⁻ by various proteins[1]**

Protein	Molecular mass (kDa)	Threshold (μM)
GpD	11.4	0.1
gpDΔN1[2]	9.8	>1,400
SHP[3]	11.8	0.4
$RGSH_6(GS)_2M$-**gpD**-$(GSGG)_3T$-**cCrk**(120–303)[4]-RSLIS	35	0.3
$RGSH_6(GS)_2M$-**gpD**-$(GSGG)_3T$-**JNK2**(2–410)[5]	62	0.4
GST(1–218)[6]-IEGRGIP-**ATF2**(1–100)[7]-SRGGSGMFM-**gpD**	50	0.4

[1]The proteins were expressed and purified as described in the Methods section. Purified λ-phages devoid of gpD were incubated with the proteins indicated and the minimal concentration for complementation was determined. Complementation refers to rendering the phage resistant against EDTA treatment by binding the protein in question to the λD^- phage. Bold letters indicate protein names and plain letters are single-letter amino acid codes.
[2]GpD (Ser 15–Val 109) with an additional N-terminal Gly.
[3]GpD homolog from phage 21.
[4]cCrk is the adapter molecule CRK, Swiss-Prot Q64010.
[5]JNK2 is the c-Jun N-terminal kinase, Swiss-Prot P45984.
[6]GST is glutathione *S*-transferase, Swiss-Prot P08515.
[7]ATF2 is the cAMP-dependent transcription factor ATF2, Swiss-Prot P15336.

Structure determination and refinement. The Matthews coefficient for the $P2_1$ crystals was calculated as 2.0 $Å^3$ Da^{-1} and the solvent content as 39% under the assumption of three gpD monomers per asymmetric unit, indicating tight packing of the protein. Six selenium atom sites were readily found using either SHELXS[42] or SOLVE[43], in agreement with the presence of two methionine residues per monomer. Refinement of the heavy atom sites using MADSYS[44] resulted in an overall figure of merit (FOM) of 0.45. After solvent flattening with DM[45], the overall FOM improved to 0.85. The initial MAD-phased electron density maps clearly showed the backbone and side chain features of many residues (Fig. 1), even before noncrystallographic symmetry averaging was applied. The initial structural model was easily built with the program O[46]. Further improvement of the electron density maps was achieved by noncrystallographic symmetry averaging, using masks built from the initial model.

The structure was initially refined using CNS[47]. At this stage, the R and R_{free} factors and electron density maps were significantly poorer than expected, necessitating reevaluation of the diffraction data. Subsequent analysis showed that the crystals of both the original protein and the Se-Met derivative were pseudomerohedrally twinned, as a result of the near-identity of the a and c axes of the $P2_1$ crystals. The twinning fraction in both types of crystals was ~0.36, and the twinning matrix was (0 0 1, 0 -1 0, 1 0 0). The presence of such twinning was confirmed by an immediate decrease in both R and R_{free} factors after the twinning was included in SHELXL refinement[48] and by the remarkable improvement in the quality of the electron density maps. The final structural model was refined using SHELXL with twinning and anisotropic temperature factor refinement. Hydrogen atoms were also included in the final refinement. Because of the high resolution of the diffraction data, it was not necessary to utilize noncrystallographic symmetry constraints in the final stages of refinement. The final R and R_{free} were 0.098 and 0.132, respectively, at 1.1 Å resolution (Table 2), and the final map was excellent (Fig. 1*b*).

The structure of gpD in the alternate (C2) crystal form was solved by molecular replacement using the gpD structure previously solved in space group $P2_1$. The program AMoRe[49] gave a clear solution, using the gpD trimer as the structural unit. This structure was not refined further.

Cryo-EM and image reconstruction. Capsids were prepared using the CosQ mutant MF1427 (λ-P1:5R Kn^R *cI*857 *nin*5 Δ*cosQ*), which is simply a *galK*⁻ derivative of the standard *E. coli* strain called C1a[23], kindly provided by M. Feiss. A 1-liter culture of the lysogenic stain was propagated at 30 °C in LB media containing kanamycin (50 mg ml^{-1}). When the cell density reached ~5×10^7 ml^{-1}, the culture was induced by shifting the temperature to 42 °C for 215 min, after which the cells were further incubated at 37 °C until lysis was observed (an additional 90–180 min). The capsids were isolated by precipitation with PEG-8000, followed by density gradient centrifugation on 10–30% (w/v) sucrose gradients prepared in TM buffer (10 mM Tris pH 7.5, 1 mM $MgSO_4$). The capsid-containing fractions were dialyzed against PBS buffer.

After an initial assessment by negative staining, samples were vitrified in thin films of buffer suspended over holey carbon films and observed under 'low-dose' conditions on a Philips CM200-FEG electron microscope equipped with a Gatan 626 cryo-holder, essentially as described[50]. Numerous focal pairs of micrographs were recorded at a magnification of 38,000 and digitized on a SCAI scanner (Zeiss Photogrammetrics) with a pixel size of 7 μm (1.8 Å at the specimen). The semiautomatic particle-picking program X3D[51] was used to extract 393 image pairs from 12 focal pairs of digitized micrographs. The positions of their first contrast transfer function (CTF) zeros varied from 17 Å to 21.5 Å (first image) and from 23 Å to 26 Å (second image). Reconstructions were calculated by Fourier–Bessel methods[52], using the projection-matching PFT algorithm of Baker and Cheng[53] to determine particle orientations, and taking our 25 Å resolution reconstruction of bacteriophage HK97 Head II (ref. 54) as starting model, after rescaling it to match the λ-capsids in size. The program CTFMIX[51] was used to correct for the CTF and to combine focal pairs. The final reconstruction included all particles (n = 380) with correlation coefficients >0.3. This population provided an even distribution of viewing angles over the icosahedral asymmetric unit. The density map had a resolution of 15 Å, as given by the band-limited resolution of the crystal structure at which it most closely resembled the corresponding portion of the EM density map.

In vitro **complementation.** Purified λ-phages devoid of gpD were incubated with crude *E. coli* extracts or with pure gpD, gpD fusion proteins, gpHDΔN1, or SHP (Table 3) as described below. Positive complementation was shown by resistance of the phage against EDTA treatment[10,18]. For each protein, the minimal concentration capable of complementation was determined. λ-Phages devoid of gpD were produced by thermal induction from NS3762 (kindly provided by R. Hoess), a nonsuppressing *E. coli* strain lysogenic for λ*Dam15 b538 cIts857 nin5* (79.5% of the wild type genome size)[10], and purified by PEG precipitation and gel filtration using a Sepharose CL-6B column (Amersham Pharmacia). For complementation, 10^8 phages were incubated for 15 min at 25 °C in SM (50 mM Tris pH 7.5, 10 mM $MgSO_4$, 100 mM NaCl) containing various concentrations of the test protein. The complemented phages were then diluted 10-fold into 10 mM Tris pH 7.5, 10 mM EDTA, incubated for 15 min at 25 °C, diluted 100-fold into SM, and titrated on a suppressing (*supE*44 and *supF*58) *E. coli* strain VCS257 (Stratagene) to count the survivors.

Coordinates. The atomic coordinates for the refined structure have been deposited in the Protein Data Bank (accession code 1c5e).

Acknowledgments
We thank V. Hawkins for assistance with the capsid preparations, H. Iwai and O. Zerbe for performing the NMR experiments and for stimulating discussions, M. Feiss, R. Weisberg, D. Belnap, and R. Hendrix for helpful suggestions, M. Feiss and R. Hoess for supplying material, and A. Arthur for editorial comments.

Received 29 November, 1999; accepted 26 January, 2000.

© 2000 Nature America Inc. • http://structbio.nature.com

© 2000 Nature America Inc. • http://structbio.nature.com

1. Hendrix, R.W., Roberts, J.W., Stahl, F.W. & Weisberg, R.A. *Lambda II.* (Cold Spring Harbor Laboratories, Cold Spring Harbor, New York; 1983).
2. Campbell, A.M. Bacteriophages. In *Escherichia coli and Salmonella: cellular and molecular biology.* (ed. Neidhardt, F.C.) 2325–2338 (ASM Press, Washington, DC; 1996).
3. Lederberg, E.M. Lysogenicity in *E. coli* K12. *Genetics* **36**, 560–560 (1951).
4. Ptashne, M. *A genetic switch: phage lambda and higher organisms.* (Blackwell Science Ltd, Boston; 1992).
5. Hendrix, R.W. & Garcea, R.L. Capsid assembly of dsDNA viruses. *Semin. Virology* **5**, 15–26 (1994).
6. Young, R.A. & Davis, R.W. Efficient isolation of genes by using antibody probes. *Proc. Natl. Acad. Sci. USA* **80**, 1194–1198 (1983).
7. Sanger, F., Coulson, A.R., Hong, G.F., Hill, D.F. & Petersen, G.B. Nucleotide sequence of bacteriophage lambda DNA. *J. Mol. Biol.* **162**, 729–773 (1982).
8. Santini, C. *et al.* Efficient display of an HCV cDNA expression library as C-terminal fusion to the capsid protein D of bacteriophage lambda. *J. Mol. Biol.* **282**, 125–135 (1998).
9. Mikawa, Y.G., Maruyama, I.N. & Brenner, S. Surface display of proteins on bacteriophage lambda heads. *J. Mol. Biol.* **262**, 21–30 (1996).
10. Sternberg, N. & Hoess, R.H. Display of peptides and proteins on the surface of bacteriophage lambda. *Proc. Natl. Acad. Sci. USA* **92**, 1609–1613 (1995).
11. Dokland, T. & Murialdo, H. Structural transitions during maturation of bacteriophage lambda capsids. *J. Mol. Biol.* **233**, 682–694 (1993).
12. Georgopoulos, C., Tilly, K. & Casjens, S.R. Lambdoid phage head assembly. In *Lambda II.* (eds Hendrix, R.W., Roberts, J.W., Stahl, F.W. & Weisberg, R.A.) 279–304 (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York; 1983).
13. Kocsis, E., Greenstone, H.L., Locke, E.G., Kessel, M. & Steven, A.C. Multiple conformational states of the bacteriophage T4 capsid surface lattice induced when expansion occurs without prior cleavage. *J. Struct. Biol.* **118**, 73–82 (1997).
14. Duda, R.L., Martincic, K., Xie, Z. & Hendrix, R.W. Bacteriophage HK97 head assembly. *FEMS Microbiol. Rev.* **17**, 41–46 (1995).
15. Forrer, P. & Jaussi, R. High-level expression of soluble heterologous proteins in the cytoplasm of *Escherichia coli* by fusion to the bacteriophage lambda head protein D. *Gene* **224**, 45–52 (1998).
16. Wurtz, M., Kistler, J. & Hohn, T. Surface structure of in vitro assembled bacteriophage lambda polyheads. *J. Mol. Biol.* **101**, 39–56 (1976).
17. Imber, R., Tsugita, A., Wurtz, M. & Hohn, T. Outer surface protein of bacteriophage lambda. *J. Mol. Biol.* **139**, 277–295 (1980).
18. Sternberg, N. & Weisberg, R. Packaging of coliphage lambda DNA. II. The role of the gene D protein. *J. Mol. Biol.* **117**, 733–759 (1977).
19. Carson, M. RIBBONS 4.0. *J. Appl. Crystallogr.* **24**, 958–961 (1991).
20. Holm, L. & Sander, C. Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**, 123–138 (1993).
21. Koradi, R., Billeter, M. & Wüthrich, K. MOLMOL: A program for display and analysis of macromolecular structures. *J. Mol. Graphics* **14**, 51–55 (1996).
22. Smith, M.P. & Feiss, M. Sequence analysis of the phage 21 genes for prohead assembly and head completion. *Gene* **126**, 1–7 (1993).
23. Cue, D. & Feiss, M. A site required for termination of packaging of the phage lambda chromosome. *Proc. Natl. Acad. Sci. USA* **90**, 9290–9294 (1993).
24. Wider, G. & Wüthrich, K. NMR spectroscopy of large molecules and multimolecular assemblies in solution. *Curr. Opin. Struct. Biol.* **9**, 594–601 (1999).
25. Ishii, T. & Yanagida, M. The two dispensable structural proteins (soc and hoc) of the T4 phage capsid; their purification and properties, isolation and characterization of the defective mutants, and their binding with the defective heads in vitro. *J. Mol. Biol.* **109**, 487–514 (1977).
26. Tatman, J.D., Preston, V.G., Nicholson, P., Elliott, R.M. & Rixon, F.J. Assembly of herpes simplex virus type 1 capsids using a panel of recombinant baculoviruses. *J. Gen. Virol.* **75**, 1101–1113 (1994).
27. Thomsen, D.R., Roof, L.L. & Homa, F.L. Assembly of herpes simplex virus (HSV) intermediate capsids in insect cells infected with recombinant baculoviruses expressing HSV capsid proteins. *J. Virol.* **68**, 2442–2457 (1994).
28. Aebi,U. *et al.* Capsid fine structure of T-even bacteriophages. Binding and localization of two dispensable capsid proteins into the P23* surface lattice. *J. Mol. Biol.* **110**, 687–698 (1977).
29. Newcomb, W.W. *et al.* Isolation of herpes simplex virus procapsids from cells infected with a protease-deficient mutant virus. *J. Virol.* **74**, 1663–1673 (1999).
30. Booy, F.P. *et al.* Finding a needle in a haystack: detection of a small protein (the 12 kDa VP26) in a large complex (the 200 MDa capsid of herpes simplex virus). *Proc. Natl. Acad. Sci. USA.* **91**, 5652–5656 (1994).
31. Zhou, Z.H., He, J., Jakana, J., Tatman, J.D., Rixon, F.J. & Chiu, W. Assembly of VP26 in herpes simplex virus-1 inferred from structures of wild-type and recombinant capsids. *Nature Struct. Biol.* **2**, 1026–1030 (1995).
32. Wingfield, P.T. *et al.* Hexon-only binding of VP26 reflects differences between the hexon and penton conformations of VP5, the major capsid protein of herpes simplex virus. *J. Virol.* **71**, 8955–8961 (1997).
33. Ren, Z.J. *et al.* Phage display of intact domains at high copy number: a system based on SOC, the small outer capsid protein of bacteriophage T4. *Protein Sci.* **5**, 1833–1843 (1996).
34. Jiang, J., Abu-Shilbayeh, L. & Rao, V.B. Display of a PorA peptide from *Neisseria meningitidis* on the bacteriophage T4 capsid surface. *Infect. Immun.* **65**, 4770–4777 (1997).
35. Desai, P. & Person, S. Incorporation of the green fluorescent protein into the herpes simplex virus type 1 capsid. *J. Virol.* **72**, 7563–7568 (1998).
36. Feiss, M., Fisher, R.A., Siegele, D.A., Nichols, B.P. & Donelson, J.E. Packaging of the bacteriophage lambda chromosome: a role for base sequences outside cos. *Virology* **92**, 56–67 (1979).
37. Ge, L., Knappik, A., Pack, P., Freund, C. & Plückthun, A. Expression antibodies in *Escherichia coli*. In *Antibody engineering*.(ed. Borrebaeck, C.A.) 229–266 (Oxford University Press, Oxford, United Kingdom; 1995).
38. Studier, F.W. & Moffatt, B.A. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189**, 113–130 (1986).
39. Qoronfleh, M.W. *et al.* Production of selenomethionine-labeled recombinant human neutrophil collagenase in *Escherichia coli*. *J. Biotechnol.* **39**, 119–128 (1995).
40. Doublie, S. Preparation of selenomethionyl proteins for phase determination. *Methods Enzymol.* **276**, 523–530 (1997).
41. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
42. Sheldrick, G.M. Patterson superposition and *ab initio* phasing. *Methods Enzymol.* **276**, 628–641 (1997).
43. Terwilliger, T.C. & Berendzen, J. Correlated phasing of multiple isomorphous replacement data. *Acta Crystallogr. D* **52**, 749–757 (1996).
44. Hendrickson, W.A. Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. *Science* **254**, 51–58 (1991).
45. Collaborative Computational Project, Number 4. CCP4 Suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
46. Jones, T.A. & Kieldgaard, M. Electron-density map interpretation. *Methods Enzymol.* **277**, 173–208 (1997).
47. Brünger, A.T. *et al.* Crystallography & NMR System: a new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
48. Sheldrick, G.M. & Schneider, T.R. SHELXL: high-resolution refinement. *Methods Enzymol.* **277**, 319–343 (1997).
49. Navaza, J. An automated package for molecular replacement. *Acta Crystallogr. A* **50**, 157–163 (1994).
50. Zlotnick, A. *et al.* Dimorphism of hepatitis B virus capsids is strongly influenced by the C- terminus of the capsid protein. *Biochemistry* **35**, 7412–7421 (1996).
51. Conway, J.F. & Steven, A.C. Methods for reconstructing density maps of "single particles" from cryoelectron micrographs to subnanometer resolution. *J. Struct. Biol.* **128**, 106–118 (1999).
52. Fuller, S.D., Butcher, S.J., Cheng, R.H. & Baker, T.S. Three-dimensional reconstruction of icosahedral particles-the uncommon line. *J. Struct. Biol.* **116**, 48–55 (1996).
53. Baker, T.S. & Cheng, R.H. A model-based approach for determining orientations of biological macromolecules imaged by cryoelectron microscopy. *J. Struct. Biol.* **116**, 120–130 (1996).
54. Conway, J.F., Duda, R.L., Cheng, N., Hendrix, R.W. & Steven, A.C. Proteolytic and conformational control of virus capsid maturation: the bacteriophage HK97 system. *J. Mol. Biol.* **253**, 86–99 (1995).

© 2000 Nature America Inc. • http://structbio.nature.com