DNA Organization and Thermodynamics during Viral Packing

C. Rebecca Locker,* Stephen D. Fuller,† and Stephen C. Harvey*
*School of Biology, Georgia Institute of Technology, Atlanta, Georgia 30332-0230; and †Division of Structural Biology, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, United Kingdom

ABSTRACT An elastic DNA molecular mechanics model is used to compare DNA structures and packing thermodynamics in two bacteriophage systems, *T7* and ϕ29. A discrete packing protocol allows for multiple molecular dynamics simulations of the entire packing event. In *T7*, the DNA is coaxially spooled around the cylindrical core protein, whereas the ϕ29 system, which lacks a core protein, organizes the DNA concentrically, but not coaxially. Two-dimensional projections of the packed structures from *T7* simulations are consistent with cryo-electron micrographs of *T7* phage DNA. The functional form of the force required to package the ϕ29 DNA is similar to forces determined experimentally, although the total free energy change is only 40% of the experimental value. Since electrostatics are not included in the simulations, this suggests that electrostatic repulsions are responsible for ~60% of the free energy required for packaging. The entropic penalty from DNA confinement has not been computed in previous studies, but it is often assumed to make a negligible contribution to the total work done in packing the DNA. Conformational entropy can be measured in our approach, and it accounts for 70–80% of the total work done in packing the elastic model DNA in both phages. For ϕ29, this corresponds to an entropic penalty of ~35% of the total work observed experimentally.

INTRODUCTION

Energy stored along double-stranded (ds) DNA chains in bacteriophage capsids must facilitate injection of the genome into host cells. As a consequence, the structure and energetics of the packaged DNA in these systems are necessarily coupled. Although both have been the focus of research for decades (1,2), they are not fully characterized for any virus system.

The inherent symmetry in icosahedral bacteriophages has facilitated detailed three-dimensional images of phage capsids (3), but nucleic acid structural detail is limited to a handful of viruses. Historically, it has been assumed that the DNA structure is similar in all bacteriophages regardless of interior protein content (1), since the packing mechanism (ATP-driven motor) and protein structure at the portal vertex (a 20–30 Å hole in a protein dodecamer) are similar among the phages (4).

Cryo-electron micrograph (cryoEM) maps of tailless bacteriophage *T7* mutants show concentric rings of DNA (5) that are consistent with lower-resolution images of a negatively stained *T7*-*T3* hybrid genome (6), and with phage models where the DNA is coaxially spooled around the core protein (5,7,8). Recent cryoEM reconstructions of epsilon15 and P22 show concentric rings organized like those in *T7* (9,10). Concentric, but not coaxial, rings were observed in minimum energy conformations of simple elastic models of DNA in phage capsids (11). These could give projections that are consistent with the EM images. Regardless of the specific DNA conformation, the genome is highly ordered within the capsid for almost all of the fully packaged cryoEM images, suggesting that the degree of ordering is reproducible in different *T7* phages, i.e., the DNA structures within multiple independently packaged *T7* phages are similar. This study examines the universality of DNA structure in phages with ATP-driven packing motors by comparing the structure of DNA packaged in two different systems. One capsid (*T7*) has a hollow cylindrical core extending from the connector protein on the capsid wall almost to the center (5), whereas the other capsid (ϕ29) has a connector protein that doesn't extend significantly into the interior of the capsid (12). Our goal is to identify the key physical properties that determine the organization and reproducibility of the DNA structure.

The structural organization of the genome during packaging is not well understood. CryoEM images of a deletion mutant *T7* phage show that concentric rings are already well-formed when 86% of the genome is packed, although packing density is low (5). Negatively stained *T7*-*T3* hybrid genomes with an added temperature-sensitive gene show disordered DNA when 40% of the genome is packaged (6). The point during packaging when the genome becomes visibly ordered has not been experimentally determined, but it can be probed via simulations here.

The dominant factors involved in genome packaging probably include DNA electrostatic repulsions, bending strain, stretching strain, torsional strain, and entropy loss due to confinement within the capsid (13). To our knowledge, no previous study has estimated the entropic penalty of packaging, and this is one of the principal purposes of this study. Since there is a high packing density of DNA in fully packed phage capsids, the largest contribution to the phage system energy is generally believed to be the DNA-DNA electrostatic repulsions, and this study also provides an estimate of that important quantity.

Submitted August 9, 2006, and accepted for publication May 17, 2007.

Address reprint requests to Stephen C. Harvey, School of Biology, Georgia Institute of Technology, Atlanta, GA 30332-0230. E-mail: steve.harvey@biology.gatech.edu.

Editor: Ivet Bahar.

0006-3495/07/10/2861/09 $2.00

doi: 10.1529/biophysj.106.094771

One approach to predicting the energetics associated with packaging DNA into capsids is to assume a model for the final structure, and to treat the DNA with a continuum elastic approximation (14–16). These models assume that phage DNA is hexagonally close-packed and use phenomenological DNA-DNA interaction potentials derived from experimental osmotic pressure data of highly condensed dsDNA (17,18). This approach gives packaging force curves that resemble the experimental data (19) fairly closely, and the adjustment of two ''fitting parameters'' allows one to reproduce the force curves quite well (15,16). Thermodynamic information from these studies must be interpreted with caution, however, for two reasons: First, DNA may not be hexagonally close packed in early packing stages because of strong electrostatic repulsions. Second, all three of these studies assumed that the entropic penalty for confining the DNA within the capsid is negligible; here we reexamine this assumption.

An alternative approach is to computationally model the DNA by a set of discrete beads connected by springs that are parameterized to match the experimental elastic moduli for stretching and bending. DNA-capsid interactions and/or long-range DNA-DNA interactions may be modeled by other terms in the energy function. The DNA is then driven into a model capsid using a suitable molecular mechanics algorithm for equilibrating and refining the structure at each stage of injection. No final structure is presumed, so this method allows an unbiased prediction of the final structure. Under some conditions, one can also estimate the forces and energetic costs of packaging. Several laboratories have published studies of this kind (11,20–23). Here we briefly review the results of those earlier studies.

Brownian dynamics simulations of minimalist λ-phage DNA with a potential that mimics a solvent with multivalent cations yield organized, concentrically spooled DNA (20). However, the DNA is disorganized without the attractive force of polyvalent cations. These results imply that DNA would be disorganized in capsids lacking polyvalent cations, although some viruses have been observed to package their genomes without polyvalent cations (24). These same authors were the first to directly measure the force resisting packaging in a molecular mechanics simulation, finding values of $\sim$10 pN for a model small capsid. This compared favorably with the values they obtained from a simple continuum model (20). An extension of this model incorporated torsional stiffness (21). The entering end of the DNA was anchored to the capsid at a point directly opposite the point of injection, to permit the calculation of twist and writhe. These simulations were carried out with a reduced capsid size, to reduce computational burden, but they produced coaxial spooling without an attractive DNA-DNA term when the DNA injection was coupled to rotation of the DNA relative to the capsid. The injection force, calculated by taking the derivative of the total energy with respect to the length of the injected DNA, was qualitatively similar to the experimental force for ϕ29. It is unclear what effects, if any, the small capsid size has in these simulations, and what effect might arise from anchoring the DNA in the Spakowitz model (21).

The recent Langevin dynamics simulations of Forrey and Muthukumar (23) introduced an icosahedral capsid, rather than the spherical model used in previous studies, and they investigated the effect of a cylindrical core protein on the structure, dynamics, and energetics. They found that the core organizes the DNA in a coaxial spool. In the absence of the core, they observe that the DNA ''has no discernible symmetry'' unless an attractive DNA-DNA interaction is included, which agrees with the earlier prediction of Kindt et al. (20). These authors estimated the packaging force by taking the derivative of the energy with respect to distance, finding a peak force of $\sim$20 pN. They also measured the force directly at various points along the trajectory, by determining the force required to hold the last DNA bead fixed at the entrance to the capsid, finding a peak force of $\sim$40 pN. They discussed the possible source of the discrepancy between these two values, to which we will return in the Concluding Remarks.

Several of the foregoing studies have assumed the packaging entropy to be negligible compared to the other contributions to the free energy (14–16,23). Although energetic analyses of hexagonally packed wormlike polyelectrolytes do show that mixing entropy is negligible (25,26), the conformational entropic penalty due to chain confinement within the capsid has not been quantified. Since the radius of gyration of viral DNA free in solution is more than 10 times greater than the radius of gyration for the packed DNA, DNA packing within a viral capsid represents a volume compression of more than 1000-fold (22). It would seem likely that the entropic penalty of confinement might be significant and bears investigation.

We have also carried out several molecular mechanics studies on DNA packaging (11,22,27). We will discuss our methods and previous results as appropriate in the remainder of this article.

In this study, we use a simplified model of DNA to simulate packing into phage capsids via molecular dynamics (MD). Forces required to pack the DNA are computed directly, compared to experimentally measured forces, and decomposed to determine how much entropy and the various components of the internal energy contribute to the free energy cost of packing. Representative structures along the packing pathway of both systems provide detail unavailable in even the highest resolution EM images. These structures are compared to ascertain what structural characteristics of the packaged genomes are independent of viral class. Projections of the packed *T*7 structures are compared to cryoEM images to test the legitimacy of the model. We emphasize that this model is based entirely on the physical properties of DNA. No assumptions are made about the final structures, and the model contains no free parameters that have to be optimized by comparison to experimental observables measured in phage systems.

MODEL AND METHODS

DNA model

The dsDNA is represented as a discretized approximation to a continuum elastic model, as detailed elsewhere (11,22,27). It is modeled as a homopolymer where each monomer (pseudo-atom) represents six consecutive basepairs (bp). The ϕ29 19,300 bp genome is modeled as a 3,217-mer, and the *T*7 39,937 bp genome is modeled as a 6,656-mer. Bending and stretching energies are included explicitly, with parameters derived from Young's modulus and persistence length of free DNA. There is also a ''soft sphere'' repulsion term to prevent strand crossing. Torsional strain from DNA twisting is assumed to be negligible and it is not included. This assumption receives support from simulations in which torsional stiffness was included, where it was found to make almost no contribution to the deformation energy (21). Long-range electrostatics are not included in the current model, but short-range electrostatic interactions and hydration forces are included implicitly in the bending stiffness, and in the soft-sphere excluded volume energy term.

The potential energy for a given conformation of the DNA is (11,27)

$$E(\{r_i,\theta_i\}) = \sum_i (\mathcal{E}_{\vec{r}_i} + \mathcal{E}_{\theta_i}) + \sum_{i,j'} \mathcal{E}_d(\vec{r}_i,\vec{r}_j). \quad (1)$$

The bond vector, $\vec{r}_i$ is the bond between monomers i and $i+1$, θ_i is the acute angle formed between $\vec{r}_i$ and $\vec{r}_{i+1}$, and the prime on the second summation excludes i, j pairs that are less than three monomers away along the chain. The local energetic terms, $\mathcal{E}_{\vec{r}_i}$ and $\mathcal{E}_{\theta_i}$, are modeled as harmonic potentials ($\mathcal{E}_\alpha = k_\alpha/2(\alpha_i - \alpha_0)^2$) representing the stretching ($\alpha = r$) and bending ($\alpha = \theta$) energies of DNA in physiological conditions, respectively. The nonlocal term in Eq. 1 represents the excluded volume, or effective diameter, of the DNA chain under physiological conditions, and is modeled with a semiharmonic repulsive potential

$$\mathcal{E}_d(\vec{r}_i,\vec{r}_j) = \begin{cases} \frac{k_d}{2}(d_{ij} - d_0)^2 & \text{if } d_{ij} < d_0 \\ 0 & \text{otherwise} \end{cases}, \quad (2)$$

where k_d is the force constant for the overlap of monomers i and j, d_{ij} is the distance between monomers i and j, and d_0 is the effective diameter of the DNA chain.

Although the form of the energy function defined in Eq. 1 is identical to that reported earlier (11), the parameterization procedure and simulation temperature differ. Thus, the force constants differ somewhat. Parameters for Eq. 1 are first derived from experimental observables of B-DNA (28) for a 1 bp per monomer model following the procedure described in our previous work (22). A Gaussian distribution in the average rise between bp is assumed (28), and it is equated to the statistical distribution of bond lengths, $P(r) \propto r^2 e^{-(k_r/2RT)(r-r_0)^2}$, to parameterize the bond stretching force constant. This force constant represents the elastic modulus for stretching. The bending force constant is derived from the DNA persistence length, $P_\infty = \langle r_0 \rangle / (1 - \langle \cos\theta \rangle)$, where $\langle \cos\theta \rangle = \int_{\theta=0}^{\pi} \cos\theta \sin\theta e^{-(k_\theta/2RT)\theta^2}$ (29), and P_∞ is 510 Å (30). By design, this model reproduces the DNA elasticity observed experimentally. The effective chain diameter is 25.0 Å, a value that is consistent with the interaction distance of fully packed viruses (5) and with light scattering measurements on dsDNA under condensing conditions (17). The force constant k_d is set to k_r. This choice is somewhat arbitrary and was made to guarantee that the relaxation times associated with interchain repulsions are much smaller than the relaxation times associated with the other degrees of freedom. We have varied k_d and find, as expected, that the structures and packaging energies are insensitive to this parameter. These nonbonded repulsions are, of course, very short-range and ignore the long-range hydration and electrostatic effects. Those will be the subject of a later study. We have shown (22) that macroscopic properties of the model, such as the mean square end-to-end distance, scale similarly to other DNA models, like the Random Φ (29) and Wormlike Chain (31) models.

Next, parameters for the 6 bp per monomer model are determined by the following procedure (22): i), A chain containing 1.7×10^7 monomers is generated following statistical distributions of bond lengths and bending angles for the 1 bp per monomer model. ii), A new chain is developed where each consecutive monomer is every sixth monomer in the 1 bp per monomer model (i.e., six monomers' degrees of freedom are averaged out). iii), New values of the force constants and ideal values for bond lengths and angles are computed from least-squares fits to probability distribution functions of r and θ. Specifically, the values are $k_r = k_d = 3.49$ kcal/mol·Å^2, $r_0 = 19.9$ Å, $d_0 = 25.0$ Å, $k_\theta = 22.4$ kcal/mol·rad^2, and $\theta_0 = 0$ rad. The parameters in Eq. 1 are derived from experimental observables, so the energy of the model chain includes not only intrastrand DNA forces, but also hydration forces and local electrostatic interactions.

Although DNA twisting contributes to the overall potential energy of the chain and may play a role in DNA structure, the change in energy associated with torsional degrees of freedom upon packaging the DNA into a viral capsid is small (21). The DNA in our model is packed into the capsid slowly enough to allow for relaxation of all energetic terms (see below). Thus, twist is assumed to make a minimal contribution to the overall chain energy, and it is consequently neglected in this model.

Interstrand electrostatic repulsions are likely strong, and they will increase as the DNA is packed into the capsid. Phenomenological electrostatic potentials have been developed (14–16,20) that assume a close-packed DNA geometry. This study does not require any assumptions about the geometry, so these potentials would not be suitable here. Two previous molecular mechanics simulations have investigated a Debye-Huckel treatment of the electrostatics (21,23), but they were restricted to relatively small model viruses, because of the added computational burden. We want to examine models for larger systems, so long-range DNA-DNA electrostatic interactions are not explicitly included in our model. They will be the subject of a forthcoming study.

Capsid models

The ϕ29 and *T*7 capsids are approximated by spheres, inside of which the DNA is restrained by a semiharmonic capsid restraint term (11),

$$\mathcal{E}_c(\vec{r}_i) = \begin{cases} \sum_{i=1}^{N} \frac{k_c}{2}(r_i - r_c)^2 & \text{if } r_i > r_c \\ 0 & \text{otherwise} \end{cases}, \quad (3)$$

with a force constant of $k_c = 8.75$ kcal/mol·Å^2. The assumption that DNA-capsid interactions are purely repulsive is consistent with experimental observations of phage HK97 (32), and it is probably appropriate for all phages, since the DNA must be bound to the capsid loosely enough to inject into the host cell upon infection. The capsid radii for ϕ29 ($r_c = 225$ Å) and *T*7 ($r_c = 267$ Å) are based on the dimensions of the mature capsids determined from cryoEM images (6,12,33). The mature *T*7 icosahedral capsid can be approximated as a sphere with little error (1), whereas the mature ϕ29 capsid is somewhat elongated (12) and the spherical approximation might yield different structures and packing energetics than those reported here. Spherical capsids used a first-order approximation in both theoretical models (15, 16) and computational models (11,21–22). Realistic capsid shapes have recently been introduced by Forrey and Muthukumar (23), and examination of the effects of capsid shape will be the subject of forthcoming work.

To model the 190 × 260 Å cylindrical core protein with a hollow channel diameter of 35 Å in the *T*7 system (33), 24 nonchain pseudo-atoms are introduced that exclude the DNA chain from that region through a DNA-core protein potential given in Eq. 2, where $k_{i,j} = 3.49$ kcal/mol·Å^2, $d_0 = 51.25$ Å for DNA-core interactions, and $d_0 = 40.0$ Å for core-core interactions. The core protein pseudo-atoms are arranged as three consecutive rings, with each ring composed of eight pseudo-atoms evenly spaced around the axis of entry to mimic the actual cylindrical core protein.

Discretized packing

To model the ATP-driven packaging process, energy terms are added to the Hamiltonian to drive the chain into the capsid. The in silico packaging process is discretized into packing steps by harmonically coupling four monomers (''studs'') with the following potential:

$$\mathcal{E}_s = \sum_{j=1,4} \frac{k_s}{2}(s_j - s_{0,j})^2, \quad (4)$$

where $k_s = 0.6$ kcal/mol·Å^2, and the studded monomers have instantaneous vector positions, s_j, and ideal vector positions, $s_{0,j}$, located at the entrance to the capsid core near the ATP-driven molecular motor site. The ideal positions move incrementally forward through the core, providing a driving force for packing the DNA chain along one axis of entry. The chain is advanced in half steps at intervals of 12,000 ps. The system is allowed to relax to thermal and energetic equilibrium, and then the stud positions are moved $r_0/2$ farther into the capsid. This process is repeated until all monomers are packaged. YAMMP (34) is used for all MD simulations (time step=1.0 ps), and the packing protocol is identical to that reported previously (22). Temperature is maintained at 300 K through coupling to a Berendsen temperature bath (35) with a coupling time constant of 250 ps. Through extensive simulations, we have determined that this is approximately equal to the natural relaxation time of the slowest modes in the system (22).

Like other studies using molecular mechanics approaches, we face two major questions about equilibration: First, is the structure at equilibrium after each packaging step? In the absence of explicit solvent, equilibration is facilitated by the absence of viscous effects, and molecular dynamics offers an advantage over Monte Carlo and Langevin dynamics, because the intertial component of the motion in MD facilitates structural reorganization (11,27). We have determined the relaxation times of many structural and energetic parameters (22) and have used those to design quite conservative protocols. Several molecular mechanics modeling studies from our laboratory and others have led to similar structural and energetic conclusions, suggesting that we and others are probably dealing with equilibrium structures.

A second question is whether or not the packaging force we are measuring is an equilibrium force. Like others who have measured the force (20,23), we do so at various fixed lengths along the trajectory, by averaging the fluctuating force necessary to hold the DNA in the capsid. There is no net DNA motion over the very long timescale of the measurement (5 μs), so this is an equilibrium measurement.

Although the discrete packing protocol precludes direct comparison of time-dependent processes to experiment, it offers the following advantages: i), It reduces CPU time to enable simulations of the entire packing process for multiple packing simulations. ii), Equilibration at each packing step facilitates ergodicity. This is particularly important at later stages in packing where the dynamics are significantly slowed due to crowding within the capsid. iii), It mimics DNA packaging via portal proteins (1) by restricting the packing to one axis of entry.

Packing thermodynamics

Since the packaging forces that we measure are equilibrium values, integration of the force versus distance curve gives the work done during packaging. This is equal to the change in Helmholtz free energy, ΔA, since there is no $P\Delta V$ work. The internal energy change ΔE is calculated by summing all the terms of the molecular mechanics force field, Eqs. 1–4, and subtracting $\langle E_0 \rangle$ for the free chain ($\langle E_0 \rangle = (3N - 6/2)RT$, where N is the number of pseudo-atoms defining the DNA chain). We obtain the entropic penalty by subtraction, $-T\Delta S = \Delta A - \Delta E$.

To calculate thermodynamic properties, we average values from 10,000 uncorrelated conformations (collected at 500 ps intervals) for each point along the packing simulation. Then, the property is averaged over all runs, $\langle G_\chi \rangle_N = (1/N)\sum_{g=1}^{N} G_\chi(g)$, where N is the total number of runs, and $G_\chi(g)$ is the average value of the property from run g at χ% packaged. One ϕ29 trajectory was equilibrated 10 times longer than the others at every packing step. The thermodynamic properties of this trajectory are within the error range of the other 10 trajectories (22), thereby supporting the assumption of equilibration at each packing step. Packing energies, forces, and free energies are computed from ϕ29 ($N = 11$) and $T7$ ($N = 10$) packing simulations.

Simulated cryoEM images

A pseudo-density for each fully packed $T7$ DNA chain was created by generating points at random locations within each cylindrical bond connecting successive DNA pseudo-atoms, at a density of 2000 points per bond. This was converted into a 200 × 200 × 200 density map (3 Å voxels) by counting the number of points within each voxel. The maps for the chains of 10 models were then added together to create an average density map. Two-dimensional projections were generated with SPIDER (36), Fourier transformed (FT), and the effect of the contrast transfer function was included by multiplying the FT by the complex transfer function for the experimental conditions reported by Cerritelli et al. (5) (Cs = 2 mm; df = −25,000 m source size 0.0005–1). The model images in Fig. 4 were then obtained by reverse FT.

RESULTS AND DISCUSSION

Packing transitions

To look for differences in the range of accessible structures at different points along the packaging trajectory, we have determined the probability distribution functions (PDFs) of the potential energy per monomer ($P(E/m)$). These are computed at 10% packaging increments in the ϕ29 and $T7$ model systems. Each distribution histogram is fit to a normalized Gaussian function, and the results are presented in Fig. 1. The PDFs narrow upon packing in both systems, since the chains are confined to spherical volumes less than the average radius of gyration for the corresponding free DNA. This crowding leads to fewer available conformational states upon packing.

The PDFs shown for ϕ29 and $T7$ in Fig. 1 highlight three main packing regimes. Until ~30% of the chain is packed, $T7$ PDFs shift to lower average energies, whereas ϕ29 PDFs shift to higher average energies. Increases in the ϕ29 bending, capsid restraint and stud energies give an overall increase in $P(E/m)$ upon packing 10–30% of the ϕ29 genome. The ϕ29 system has a higher $P(E/m)$ than the $T7$ system while the first 70% of the chain is packed. Since the diameter of the ϕ29 capsid (475 Å) is smaller than the DNA persistence length (510 Å), the overall bending strain in the system increases with chain confinement, even at early packing stages. The bending strain, and thus the $P(E/m)$, are lower in the initial packing phase of $T7$ since the diameter of the $T7$ capsid is larger (558 Å).

When 40–60% of the genome is packaged, there is a roughly linear increase in $P(E/m)$ in both systems. This linear regime is due to increases in both the capsid restraint and stud energies, whereas all of the other energy terms,

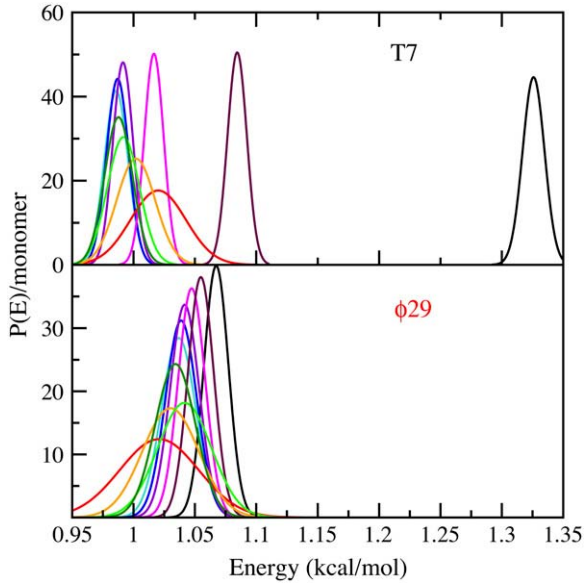

FIGURE 1 Probability distributions of energy per monomer for the $\phi 29$ and the $T7$ phage systems. Individual distribution functions represent 10% (*red*), 20% (*orange*), 30% (*light green*), 40% (*dark green*), 50% (*turquoise*), 60% (*blue*), 70% (*violet*), 80% (*magenta*), 90% (*maroon*), and 100% (*black*) packed. All PDFs are presented as Gaussian fits to averages of all packing trajectories.

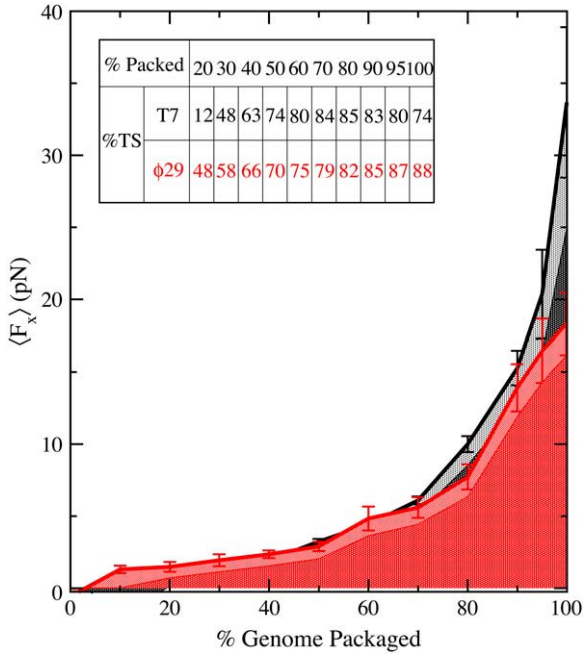

FIGURE 2 Force required to pack DNA as a function of percent genome packed for $T7$ (*black*) and $\phi 29$ (*red*) phage systems. Error bars are shown for each computed force. Data points are numerically integrated using the trapezoid rule to estimate the reversible work at every packing stage, through $W = \int_{\chi=0}^{\chi'} F_x d\chi$. The contribution of the change in internal energy, ΔE, is represented by the light gray and light red areas for $T7$ and $\phi 29$, respectively. The entropic contribution to the work is computed by subtraction and represented by the dark gray and red areas. The table shows the percent contribution of the entropic term ($-\Delta TS$) to the total work done in packaging the genome.

including the excluded volume term, remain relatively constant. This indicates that the capsids are not significantly overcrowded at this stage. The linear trend continues until all of the $\phi 29$ DNA is packaged, but a nonlinear high density regime develops during the last 30% of $T7$ DNA packing. In this phase, the rates of increase in the excluded volume, stud, and capsid restraint energies per monomer rise rapidly, surpassing the average energy per monomer of the fully packaged $\phi 29$ system by the time the $T7$ DNA is between 80% and 90% packed. This indicates that $T7$ DNA is significantly crowded in the last packing stage.

DNA packaging is essentially trajectory-independent for the first 70% of genome packaged for both systems (i.e., the PDFs are similar for all runs). Thus, all conformations are explored in one packing run. By ~80% packed, the PDFs are different for different runs. This is indicative of a glassy regime where it would take longer than our simulation time to explore all available conformations, which is likely a result of high packing densities. The statistical error in thermodynamic properties introduced by this phenomenon is reduced by averaging over 10 independent packing trajectories.

Forces during packaging

Fig. 2 shows the average force required to package the $T7$ and $\phi 29$ model genomes as a function of percent of genome packaged. As intuitively expected from crowding in the capsid, the force increases rapidly in both systems as more DNA is packaged. The force in the $\phi 29$ model is ~40% of that determined experimentally (19).

The roughly exponential dependence of the force in Fig. 2 on the percent of DNA packaged is similar to experiment (19), and to an analytical expression for an idealized theoretical model (15). It also resembles those found in previous simulations (20, 23). Like the earlier simulations, the force is smaller than the experimental value. One interpretation of these results is that much of the experimental force is due to electrostatic repulsion, whereas the exponential dependence of the force is largely due to confinement of the DNA into a small volume.

Free energy of packing

The total free energy required to pack the DNA is given in Table 1, along with a breakdown of the free energy into its various components. More free energy is required to completely pack the $T7$ genome than the $\phi 29$ genome, because the $T7$ genome is twice as long as the $\phi 29$ genome. However, the free energy of packing per monomer is greater for $T7$ than for $\phi 29$ (1.67 and 1.46 kcal/mol·monomer, respectively). This is surprising since the actual void volumes within the capsids are similar as are the changes in entropy per monomer. Most of the difference between the free energies is due to

TABLE 1 Change in Helmholtz free energy upon packing, ΔA, averaged over all packing trajectories for ϕ29 and *T*7 phages

Energy, (kcal/mol)	ϕ29	*T*7
ΔA	4,700	11,100
ΔE	550 (12%)	2,850 (26%)
ΔE_θ	360 (8%)	960 (9%)
ΔE_r	110 (2%)	650 (6%)
ΔE_d	70 (1%)	1,030 (9%)
ΔE_c	10 (0%)	90 (1%)
ΔE_s	4 (0%)	120 (1%)
$T\Delta S$	−4,140 (88%)	−8,250 (74%)
ΔG_{expt}	12,000	—

The contribution from each term is listed with the percent contribution in parentheses, where $\Delta E_\nu = E_{\nu,\ 100\%} - E_{\nu,\ 0}$. $E_{\nu,\ 0}$ is the average ν component of the internal energy of the unconfined chain, where ν represents the bending ($\nu = \theta$), stretching ($\nu = r$), excluded volume ($\nu = d$), capsid restraint ($\nu = c$), and studded monomer ($\nu = s$) energies as defined in the Model and Methods section. ($E_{b,0} = (1/2)N_bRT$ and $E_{\theta,\ 0} = N_\theta RT$, where R is the universal gas constant, T is the absolute temperature, and N_b and N_θ are the numbers of bonds and angles, respectively. $E_{d,\ 0} = E_{c,\ 0} = E_{s,\ 0} = 0$.) The total change in potential energy is computed as $\Delta E = \sum_\nu \Delta E_\nu$. The entropic penalty is obtained by subtraction ($-T\Delta S = \Delta A - \Delta E$). The free energy of packing ϕ29, ΔG_{expt}, computed from experimentally measured forces, is shown for comparison (19).

overcrowding in the *T*7 system, as shown by the large change in excluded volume energy, ΔE_d, and this is attributed to the presence of the internal core protein pseudo-atoms in the *T*7 system. The large protein core in *T*7 DNA gives the capsid interior a highly irregular shape, which must prevent the DNA from packing as efficiently as the DNA in the ϕ29 system.

The entropic penalty is very large and is the dominant source of resistance in packaging these elastic models. For ϕ29, it increases steadily to a maximum of 4,140 kcal/mol (~88% of the total work) when the capsid is fully packed (Fig. 2, *inset*). Interestingly, the entropic portion of the work required to pack *T*7 DNA reaches a maximum of ~85% of the total work when the capsid is between 70% and 80% packed. This implies that there is an entropic bottleneck in packing the *T*7 DNA that is absent in the ϕ29 system, and is consistent with the transition to a nonlinear packing regime seen in the PDFs in Fig. 1. This bottleneck may be caused by the geometric strain induced by the interior core protein in the *T*7 capsid.

The computed free energy required to pack ϕ29 is ~40% of that found in pulling experiments (19), where a value of ~12,000 kcal/mol was reported. Our model does not include electrostatics, so an estimate of the long-range electrostatic contribution can be obtained by subtraction as 12,000−4,700 = 7,300 kcal/mol. If this is correct, the electrostatic interactions account for ~60% of the total experimental free energy of packing, and the entropic penalty accounts for ~35%. We recognize, of course, that the relative contributions of the electrostatic, entropic, and elastic components may be different in a model that includes explicit electrostatic interactions, but the current model predicts that electrostatics should account for ~½−⅔ of the total free energy cost, with the entropic penalty accounting for most of the remainder.

Representative conformations

Conformations from one representative packing trajectory of ϕ29 (*top four panels*) and *T*7 (*bottom four panels*) are shown in Fig. 3. The chains exhibit concentric, but not coaxial, inside-outside spooling when 50% of both genomes are packaged (*top panels* of each system). This is evident in all 11 ϕ29 packing trajectories, and seven out of 10 *T*7 trajectories. (The other three exhibited coaxial spooling.) Concentric spooling resolves to full coaxial spooling in six out of the seven *T*7 cases by 100% packed, as is the case for the trajectory shown in Fig. 3. In all 11 ϕ29 trajectories, fully packaged DNA is concentrically, but not coaxially, spooled. These results support the conclusion that a cylindrical core is required to produce coaxial spooling (23). They contradict

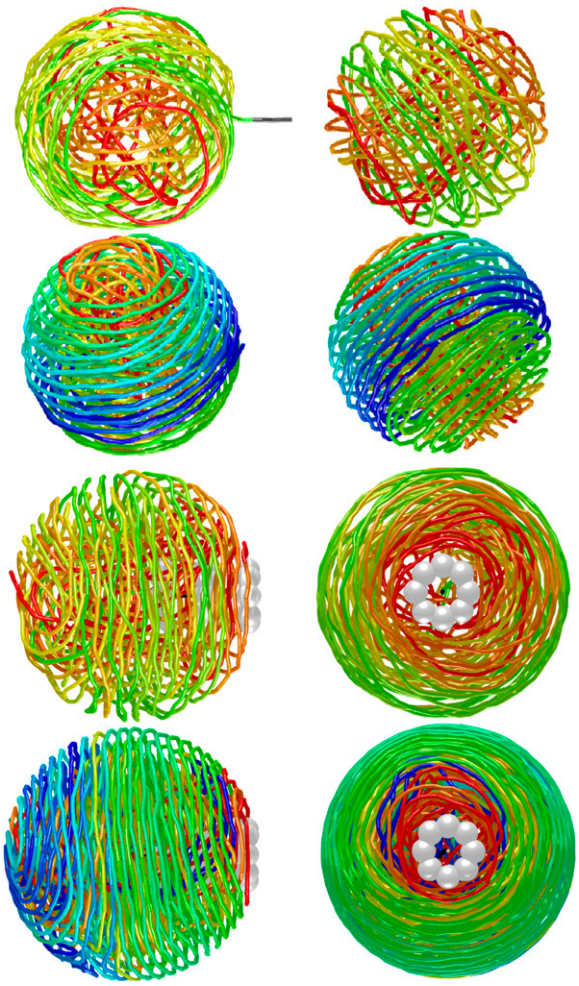

FIGURE 3 DNA conformations from one packing trajectory each, for ϕ29 (*four top panels*) and for *T*7 (*four bottom panels*). Top images show 50% of the DNA packaged for each system, and the bottom images show 100% of the genome packaged, with the last monomer (colored *white*) held just outside the capsid. Left images are side views, and right images are views down the axis of entry. The diameter of the DNA is shown at 35% of the actual value to allow visualization of the interior structure. The first monomer to enter the capsid is colored red, and subsequent monomers are colored according to the color spectrum.

director field models that predict inverse spooling in phage capsids based on energy minimizations of coarse-grained DNA potentials confined to a lattice (37,38), and a simple model where the leading end of the DNA chain was fixed directly opposite to the entry portal (21). The degree of organization is consistent with Brownian dynamics simulations of λ-phage DNA with an attractive-repulsive electrostatic potential, but not with simulations with a purely repulsive term, which give rise to disordered structures (14). Although fully packaged *T*7 DNA preferentially spools about the interior core protein in all cases, the axis of spooling in ϕ29 varies from trajectory to trajectory.

Both systems display long-range order in the fully packed DNA chains. This implies that packing dsDNA into bacteriophages is a deterministic process. The ϕ29 phage system is deterministic because the genome is small enough to limit the conformational search space. The *T*7 system is deterministic because the internal core protein geometrically limits the conformational search space. In all simulations, packing proceeds with smooth global bending of the model DNA. This is consistent with analytic studies of phage systems wherein an estimate of the energetic penalty of random kinks in the DNA is shown to be much greater than a continuous bending energy (13).

Reconstructed two-dimensional electron microscopy images

Projected images for *T*7 are shown for views perpendicular to the entry axis (*top left panel*) and down the axis (top right panel) in Fig. 4. The punctate pattern in the perpendicular view is caused by increased density in and out of the plane as the DNA wraps around the core protein. The core protein location is identified by the low density region on the right side of the projection. The eight concentric rings in the view down the entry axis are spaced ~25 Å apart, equal to the hard-core repulsive interstrand distance. Both views are consistent with cryoEM images of DNA in tailless *T*7 mutants (5) shown in the two bottom panels in Fig. 4. This similarity further supports the coaxial spool model for *T*7 DNA structure. These images are invariant to small changes in projection angle, suggesting that the structure is well defined in all 10 independent trajectories. Although it may be surprising that such a simple model can predict viral DNA structure, these results strongly suggest that DNA structure is predominantly determined by its elastic properties in the confinement of a small capsid.

CONCLUDING REMARKS

Structural and thermodynamic experimental properties are qualitatively reproduced from a simple elastic model of phage DNA with no explicit electrostatic interactions, and no a priori knowledge of the packed structure. Simulations of multiple independent packing events show that the high degree of order in the DNA structures in ϕ29 and *T*7 model

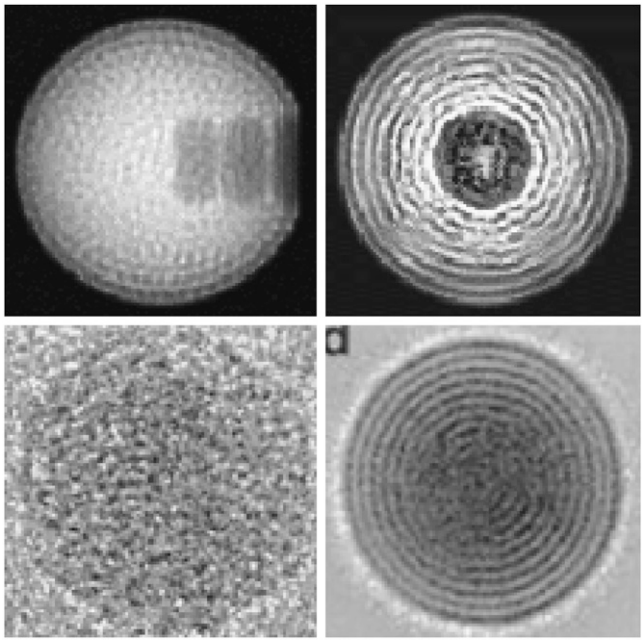

FIGURE 4 Comparison of model (*top*) and experimental (*bottom*) two-dimensional cryoEM images of *T*7 DNA. Model images were obtained by averaging densities of fully packed structures from 10 independent packing trajectories. A punctate image is produced from projections orthogonal to the axis of entry (*top left panel*), and concentric rings are evident from projections down the axis of the cylindrical core protein (*top right panel*). The experimental punctate pattern in the side view of one *T*7 tailless mutant DNA (*bottom left panel*), and the average over 77 views down the entry axis (*bottom right panel*) (5) are shown for comparison. Experimental images reproduced with permission from Alasdair Steven.

systems are reproducible. The ϕ29 genome packs concentrically, but not coaxially. The *T*7 DNA packs coaxially around the interior core protein, in agreement with cryoEM images of *T*7. The force dependence on the amount of genome packed is comparable to that observed for ϕ29, but the force required to pack the model DNA is only ~40% that of the real system. This suggests that more than half of the experimentally observed force arises from electrostatic effects. The conformational entropic penalty accounts for more than half of the free energy cost of confining this elastic model DNA to the capsid, so it should not be neglected in free energy, force, and other thermodynamic calculations. Although more complex interactions like electrostatic repulsions certainly play a role in packing thermodynamics and kinetics, the results from this study imply that the three-dimensional structure is primarily determined by elastic effects of confinement to the capsid. This may be a general result for all viruses in which the capsid radius is comparable to or smaller than the DNA persistence length.

The structures observed here are similar to those produced by recent Langevin dynamics simulations of a more coarse-grained model of phage *T*7 (23). That study introduced an icosahedral capsid, and two-dimensional projections of chain density are similar to our two-dimensional cryoEM

reconstructions. Although the potential and dynamics protocols differ, the interior core protein within the *T*7 capsid is found to order packing in both cases. This further validates the use of simple models to study viral packing, and suggests that structure is largely determined by the presence or absence of a core, and by bending strain.

To our knowledge, we are the first to suggest that the entropic penalty of packaging is substantial. Several previous authors have assumed that it is negligible (14–16,23). Interestingly, one of these studies provides direct evidence that supports our contention. Forrey and Muthukumar (23) calculated the force by two different methods: first, by taking the derivative of the potential energy; and second, by a direct measurement of the restraining force. The latter is identical to our procedure. They found that the first method gives a force that is about half that obtained from the direct method. The molecular mechanics energy corresponds to the internal energy, which has no contribution from conformational entropy, whereas the total force is derived from the free energy, which does. We believe their result supports our conclusion that a substantial fraction of the total free energy change is due to conformational entropy.

The use of phage systems to address features common to viruses in general is becoming both common and productive (39–41). Here we take advantage of the regularity of genome packaging in several phages to draw a correspondence between the predictions of our model of DNA packaging and observations by cryo-electron microscopy and single molecule experiments. Although the mechanism of DNA packing is likely different from that in an animal virus, the energetic challenges faced in forming the final packaged core are similar. Indeed, recent work on Herpes virus reveals that such familiar concepts from phages as the involvement of portal vertices in packaging are also represented in the animal virus system (41). Recently, the first atomic model of a membrane virus was obtained by taking advantage of the convenience afforded by a phage system for crystallization of the entire complex (40). Our work is another example of how a phage system can be used to conveniently address questions that would be more challenging in a less ordered assembly such as an animal virus.

We thank Anton Petrov for stimulating discussions, and Robert Tan for discussions and help with YAMMP.

This work has been supported by grants from the National Institutes of Health to S.C.H. (GM70785) and from the Wellcome Trust Program to S.D.F. (H5RCZRO). S.C.H. is a Georgia Research Alliance Eminent Scholar, and S.D.F. is a Wellcome Trust Principal Research Fellow.

REFERENCES

1. Black, L. W. 1989. DNA packaging in dsDNA bacteriophages. *Annu. Rev. Microbiol.* 43:267–292.
2. Prasad, B. V., and P. E. Prevelige. 2003. Viral genome organization. *Adv. Prot. Chem.* 64:219–258.
3. Baker, T. S., N. H. Olson, and S. D. Fuller. 1999. Adding the third dimension to virus life cycles: Three-dimensional reconstruction of icosahedral viruses from cryo-electron micrographs. *Microbiol. Mol. Biol. Rev.* 63:862–922.
4. Bazinet, C., and J. King. 1985. The DNA translocating vertex of dsDNA bacteriophage. *Annu. Rev. Microbiol.* 39:109–129.
5. Cerritelli, M. E., N. Cheng, A. H. Rosenberg, C. E. McPherson, F. P. Booy, and A. C. Steven. 1997. Encapsidated conformation of bacteriophage T7 DNA. *Cell.* 91:271–280.
6. Serwer, P., S. A. Khan, S. J. Hayes, R. H. Watson, and G. A. Griess. 1997. The conformation of packaged bacteriophage T7 DNA: Informative images of negatively stained T7. *J. Struct. Biol.* 120:32–43.
7. Richards, K. E., R. C. Williams, and R. Calendar. 1973. Mode of DNA packing within bacteriophage heads. *J. Mol. Biol.* 78:255–259.
8. Earnshaw, W. C., and S. C. Harrison. 1977. DNA arrangement in isometric phage heads. *Nature.* 268:598–602.
9. Jiang, W., J. Chang, J. Jakana, P. Weigele, J. King, and W. Chiu. 2006. Structure of epsilon15 bacteriophage reveals genome organization and DNA packaging/injection apparatus. *Nature.* 439:612–616.
10. Lander, G. C., L. Tang, S. R. Casjens, E. B. Gilcrease, P. Prevelige, A. Poliakov, C. S. Potter, B. Carragher, and J. E. Johnson. 2006. The structure of an infectious p22 virion shows the signal for headful DNA packaging. *Science.* 312:1791–1795.
11. LaMarque, J. C., T. L. Le, and S. C. Harvey. 2004. Packaging double-helical DNA into viral capsids. *Biopolymers.* 73:348–355.
12. Tao, Y., N. H. Olson, W. Xu, D. L. Anderson, M. G. Rossmann, and T. S. Baker. 1998. Assembly of a tailed bacterial virus and its genome release studied in three dimensions. *Cell.* 95:431–437.
13. Riemer, S. C., and V. A. Bloomfield. 1978. Packaging of DNA in bacteriophage heads: Some considerations on energetics. *Biopolymers.* 17:785–794.
14. Tzlil, S., J. T. Kindt, W. M. Gelbart, and A. Ben-Shaul. 2003. Forces and pressures in DNA packaging and release from viral capsids. *Biophys. J.* 84:1616–1627.
15. Purohit, P. K., J. Kondev, and R. Phillips. 2003. Mechanics of DNA packaging in viruses. *Proc. Natl. Acad. Sci. USA.* 100:3173–3178.
16. Purohit, P. K., M. M. Inamdar, P. D. Grayson, T. M. Squires, J. Kondev, and R. Phillips. 2005. Forces during bacteriophage DNA packaging and ejection. *Biophys. J.* 88:851–866.
17. Rau, D. C., B. Lee, and V. A. Parsegian. 1984. Measurement of the repulsive force between polyelectrolyte molecules in ionic solution: hydration forces between parallel DNA double helices. *Proc. Natl. Acad. Sci. USA.* 81:2621–2625.
18. Rau, D. C., and V. A. Parsegian. 1992. Direct measurements of the intermolecular forces between counterion-condensed DNA double helices. *Biophys. J.* 61:246–259.
19. Smith, D. E., S. J. Tans, S. B. Smith, S. Grimes, D. L. Anderson, and C. Bustamante. 2001. The bacteriophage ϕ29 portal motor can package DNA against a large internal force. *Nature.* 413:748–752.
20. Kindt, J. T., S. Tzlil, A. Ben-Shaul, and W. M. Gelbart. 2001. DNA packaging and ejection forces in bacteriophage. *Proc. Natl. Acad. Sci. USA.* 98:13671–13674.
21. Spakowitz, A. J., and Z.-G. Wang. 2005. DNA packaging in bacteriophage: Is twist important? *Biophys. J.* 88:3912–3923.
22. Locker, C. R., and S. C. Harvey. 2006. A model for viral genome packing. *Multiscale Model. Simul.* 5:1264–1279.
23. Forrey, C., and M. Muthukumar. 2006. Langevin dynamics simulations of genome packing in bacteriophage. *Biophys. J.* 91:25–41.
24. Hafner, E. W., C. W. Tabor, and H. Tabor. 1979. Mutants of *Escherichia coli* that do not contain 1,4-diaminobutane (putrescine) or spermidine. *J. Biol. Chem.* 254:12419–12426.
25. Odijk, T. 1993. Undulation-enhanced electrostatic forces in hexagonal polyelectrolyte gels. *Biophys. Chem.* 46:69–75.
26. Odijk, T. 1998. Hexagonally packed DNA within bacteriophage T7 stabilized by curvature stress. *Biophys. J.* 75:1223–1227.

27. Arsuaga, J., R. K.-Z. Tan, M. Vazquez, D. Sumners, and S. C. Harvey. 2002. Investigation of viral DNA packaging using molecular mechanics models. *Biophys. Chem.* 101–102:475–484.

28. Lu, X.-J., and W. Olson. 1999. Resolving the discrepancies among nucleic acid conformational analyses. *J. Mol. Biol.* 285:1563–1575.

29. Schellman, J. A. 1974. Flexibility of DNA. *Biopolymers.* 13:217–226.

30. Hagerman, P. J. 1988. Flexibility of DNA. *Annu. Rev. Biophys. Biol.* 17:265–286.

31. Kratky, O., and G. Porod. 1949. X-ray investigations on fibrous molecules. *Rec. Trav. Chim. Pays-Bas.* 68:1106–1122.

32. Conway, J. F., W. R. Wikoff, N. Cheng, R. L. Duda, R. W. Hendrix, J. E. Johnson, and A. C. Steven. 2001. Virus maturation involving large subunit rotations and local refolding. *Science.* 292:744–748.

33. Cerritelli, M. E., B. E. Trus, C. S. Smith, N. Cheng, J. F. Conway, and A. C. Steven. 2003. A second symmetry mismatch at the portal vertex of bacteriophage T7: 8-fold symmetry in the procapsid core. *J. Mol. Biol.* 327:1–6.

34. Tan, R. K.-Z., and S. C. Harvey. 1993. YAMMP: Development of a molecular mechanics program using the modular programming method. *J. Comput. Chem.* 14:455–470.

35. Berendsen, H. J. C., J. P. M. Postma, W. F. Vangunsteren, A. Dinola, and J. R. Haak. 1984. Molecular dynamics with coupling to an external bath. *J. Comp. Phys.* 81:3684–3690.

36. Frank, J., M. Radermacher, P. Penczek, J. Zhu, Y. Li, M. Ladjadj, and A. Leith. 1996. Spider and web: processing and visualization of images in 3d electron microscopy and related fields. *J. Struct. Biol.* 116:190–199.

37. Klug, W. S., and M. Ortiz. 2003. A director-field model of DNA packaging in viral capsids. *J. Mech. Phys. Solids.* 51:1815–1847.

38. Klug, W. S., M. T. Feldmann, and M. Ortiz. 2004. Three-dimensional director-field predictions of viral DNA packing arrangements. *Comput. Mech.* 35:146–152.

39. Butcher, S. J., D. H. Bamford, and S. D. Fuller. 1995. DNA packaging orders the membrane of bacteriophage PRD1. *EMBO J.* 14:6078–6086.

40. Abrescia, N. G., J. J. Cockburn, J. M. Grimes, G. C. Sutton, J. M. Diprose, S. J. Butcher, S. D. Fuller, C. S. Martin, R. M. Burnett, D. I. Stuart, D. H. Bamford, and J. K. Bamford. 2004. Insights into assembly from structural analysis of bacteriophage PRD1. *Nature.* 432:68–74.

41. Newcomb, W. W., F. L. Homa, and J. C. Brown. 2005. Involvement of the portal at an early step in herpes simplex virus capsid assembly. *J. Virol.* 79:10540–10546.