To Build a Virus on a Nucleic Acid Substrate

Adam Zlotnick,[†‡*] J. Zachary Porterfield,[†‡] and Joseph Che-Yen Wang[†]
[†]Department of Molecular and Cellular Biochemistry, Indiana University, Bloomington, Indiana; and [‡]Department of Biochemistry and Molecular Biology, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma

ABSTRACT Many viruses package their genomes concomitant with assembly. Here, we show that this reaction can be described by three coefficients: association of capsid protein (CP) to nucleic acid (NA), K_{NA}; CP-CP interaction, ω; and α, proportional to the work required to package NA. The value of α can vary as NA is packaged. A phase diagram of average $\ln\alpha$ versus $\ln\omega$ identifies conditions where assembly is likely to fail or succeed. NA morphology can favor ($\ln\alpha > 0$) or impede ($\ln\alpha < 0$) assembly. As $\ln\omega$ becomes larger, capsids become more stable and assembly becomes more cooperative. Where $(\ln\alpha + \ln\omega) < 0$, the CP is unable to contain the NA, so that assembly results in aberrant particles. This phase diagram is consistent with quantitative studies of cowpea chlorotic mottle virus, hepatitis B virus, and simian virus 40 assembling on ssRNA and dsDNA substrates. Thus, the formalism we develop is suitable for describing and predicting behavior of experimental studies of CP assembly on NA.

INTRODUCTION

Viruses are self-replicating molecular machines and obligate parasites that package a genome and deliver it to a host. In many cases, the shell (or capsid) of a spherical virus self-assembles around the viral genome. This is the case for spherical viruses such as hepatitis B virus (HBV, a member of the Hepadnaviridae family), cowpea chlorotic mottle virus (CCMV, family Bromoviridae), and simian virus 40 (SV40, polyomaviridae), three viruses that are unrelated but have well characterized assembly systems. Like HBV, CCMV, and SV40, most spherical virions that package their nucleic acids concomitant with assembly have capsid proteins (CPs) with conspicuously basic nucleic acid (NA) binding domains. Of course, for some spherical viruses, the genome is pumped into an empty procapsid through a special portal (e.g., bacteriophage phi 29 and herpes simplex virus). For a few virus families (e.g., Picornaviridae and Comoviridae), which have neither a discernible NA binding domain nor a portal, the mechanism(s) by which NA is packaged is unknown.

NA plays a structural role in the assembly of many viruses, but there is surprisingly little information about this role, nor is there a quantitative description of its assembly-related functions (1). Substantial amounts of NA are seen in only a few (e.g., Gouet et al. (2), Tang et al. (3), Larson et al. (4), and Chen et al. (5)) of the many virus crystal structures. Cryo-electron microscopy (cryo-EM) image reconstructions, which include all low-resolution terms, frequently show NA density, which often has considerable fine structure (e.g., Tihova et al. (6)). In an x-ray structure of Pariacoto virus, the viral ssRNA is clearly visualized as a dodecahedral cage with dsRNA segments running through protein-defined grooves to generate a dodecahedral cage (3), even though there is no evidence for a periodic organization of complementary segments intrinsic to the viral genome. Even more striking is that when packaging essentially random RNA in a closely related nodavirus, flock house virus, capsid structure still induced the packaged RNA to conform to a dodecahedral conformation (6). An unrelated virus, satellite tobacco necrosis virus, has a similar RNA organization (4); chemical analysis of its viral RNA suggests that it may have a fold that favors dodecahedral organization (7). A crystal structure of beanpod mottle virus showed a trefoil of ~30 ssRNA nucleotides visible on the 20 icosahedral threefolds (5), where the viral RNA had evidence of patches of pyrimidine-rich sequences (8). In bluetongue virus (family Reoviridae), with a multipartite dsRNA genome, the genome appears to be a liquid-crystal phase; however, the outermost layer is sufficiently ordered by contact with the capsid that the paths (but not the sequence) of the RNA can be traced (2). In HBV, packaged random ssRNA and genomic dsDNA (9,10) can be seen in cryo-EM image reconstructions; NA density is concentrated near basic C termini (localized to fivefold and quasi-sixfold vertices). There are two fundamental points to be made from this very short catalog. 1), There is interplay between the viral capsid and packaged NA that appears to be driven by the capsid structure. 2), The majority of packaged NA shows much less order than the protein shell. Theoretical simulations of encapsidated dsDNA also underscore the importance of interaction with CPs in NA organization (11–14).

We (and many others) have hypothesized that NA organization before and during capsid assembly can affect the capsid-assembly reaction itself (12,15–21). Using coarse-grained simulations, Hagan and Elrad posited two extremes for self-assembly: stepwise addition of CPs where

Submitted October 27, 2012, and accepted for publication February 8, 2013.

*Correspondence: azlotnic@indiana.edu

J. Zachary Porterfield's present address is Department of Internal Medicine, 330 Cedar St., Boardman 110, P.O. Box 208056, Yale University, New Haven, CT 06520-8056.

Editor: Bertrand Garcia-Moreno.

© 2013 by the Biophysical Society

0006-3495/13/04/1595/10 $2.00

http://dx.doi.org/10.1016/j.bpj.2013.02.005

protein-protein interactions are strong and initial formation of a disordered nucleoprotein complex that subsequently anneals to a virus-like particle where protein-protein interactions are weak (Fig. 1 *B*) (18). Formation of a disordered micellelike structure followed by reorganization of subunits was also proposed by McPherson (22). Neutralization of charge by CP subunits has been proposed as a mechanism for collapse of such micelles (23). Elaborating on the theme of stepwise addition of subunits, assembly may be a highly cooperative reaction where the addition of the first subunit creates an environment that preferentially completes a capsid, or it may be a low-cooperativity system where subunits distribute evenly over all NAs. Solution studies of coordinated assembly and packaging have mainly relied upon electrophoretic mobility-shift assays (Fig. 2) to demonstrate assembly of CPs that have a very strong electrostatic association with NA. The three posited assembly paths (Fig. 1 *B*) have predictable and distinct electrophoretic results (Fig. 2 *C*). The disordered-complex path will show no electrophoretic shift if subunits dissociate from the NA during electrophoresis, whereas it will result in a gradual shift of the NA band if the CP remains NA-associated (15). The low-cooperativity stepwise path will result in a gradual shift of NA migration (15), whereas the high-cooperativity path,

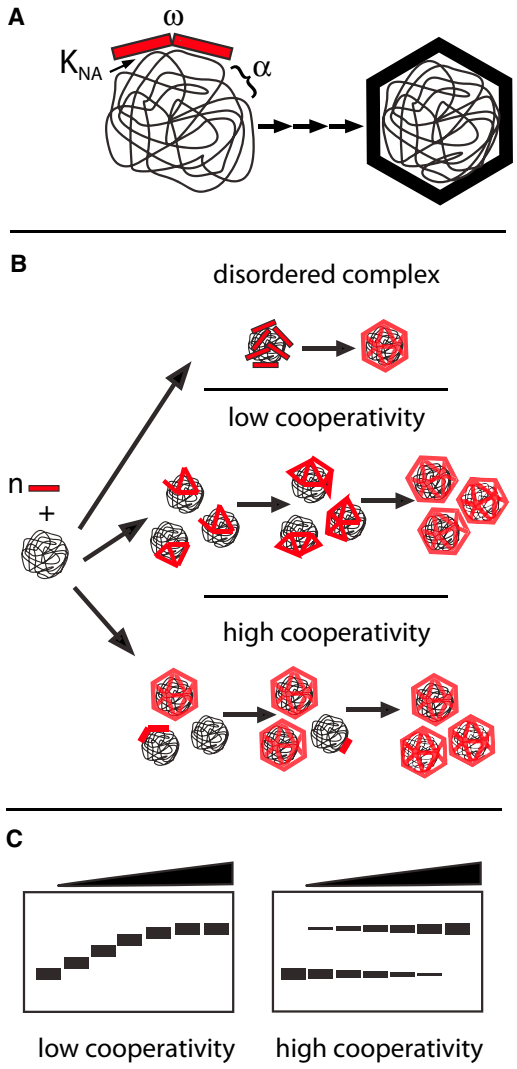

FIGURE 1 Schematic of the assembly of an NA-filled virus-like particle. (*A*) Capsid protein (CP) interaction with nucleic acid (NA) is defined by the CP-NA association constant, K_{NA}, the CP-CP interaction, ω, and the work of fitting the NA into the constraints of the growing capsid, α. (*B*) Three general paths of assembly. In collapse of disordered complexes, subunits adsorb to the NA to form a micellelike structure and then rearrange to form a capsid; this path has been associated with weak CP-CP interaction (19). A low-cooperativity assembly path has a more ordered addition of subunits, where the subunits are uniformly distributed over the many NA polymers. A high-cooperativity path, also based on ordered addition of subunits, results from preferential addition of subunits to NA where growth is already nucleated. (*C*) Hypothetical observations of different assembly reactions in which products are electrophoretically separated on an agarose gel, after which bound and free NA are visualized by ethidium bromide staining. Disordered complex formation and low-cooperativity assembly lead to a gradual shift of NA migration through a gel, so long as the CP remains NA-associated. High-cooperativity assembly leads to a bimodal distribution of NA into free and encapsidated fractions.

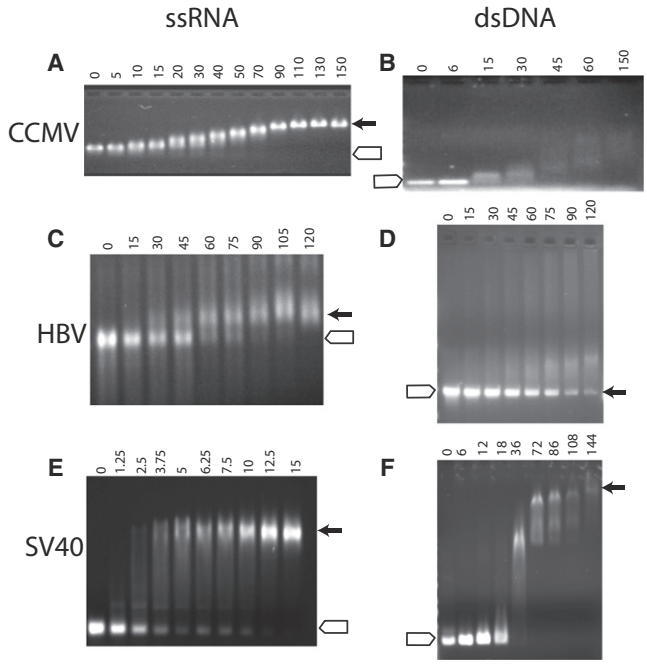

FIGURE 2 Titrations of ssRNA or dsDNA by different CPs showing the ratio of CP oligomer per NA polymer. Agarose gels run at neutral show electrophoretic migration of NA in the presence of CP at the stated ratio per NA polymer. (*A* and *B*) CCMV dimeric CP titrating 3100-nucleotide viral RNA1 (*A*) (15) or a dsDNA 600-mer (*B*) (49). (*C* and *D*) HBV CP (Cp183-EEE) titrating a 3200-nucleotide viral RNA (*C*) (21) or an 1800-base pair dsDNA (*D*) (12). Note that HBV capsids run almost at the same position as the dsDNA but are not evident. (*E and* F) SV40 VP1 pentamers titrating an 850-nucleotide ssRNA (*E*) (33) or a 600-mer dsDNA that is ~1/10 the length of the SV40 genome (*F*). In all cases (*A–F*), the expected position for free NA is denoted by the open arrow and the capsid (where known) by the closed arrow. Of these different CPs, only SV40 consistently forms dsDNA-filled capsids (16). Gel were 0.8% or 1% agarose run using a Tris-acetate buffer system.

where one NA is coated with CP before the next NA begins to get encapsidated, resulting in a bimodal distribution (21). Implicit in these simplified models is that the nature of the NA will make a defining contribution to the path of capsid assembly.

A formalism for describing protein assembly on NAs remains elusive. von Hippel and co-workers described a model of nonspecific binding of proteins to a linear NA lattice (24,25), which has been qualitatively applied to virus assembly (15,21) but does not explicitly account for polyvalent proteins and the more complex patterns of protein-NA interaction possible in a virus. Physical models with simplified representations of NA have provided insights into the process of assembly, particularly kinetics. Indeed, the relationship between charge and volume of packaged NA has been considered a determinant of particle size (26). RNA has also been considered an antenna that both attracts protein to an assembly site and facilitates one-dimensional diffusion along the RNA (27), analogous to suppressor diffusion on DNA (28). In previous works, we described a formalism for assembly of empty spherical capsids (29,30). Here, we extend it to assembly of NA-filled particles. We observe that the synergism between NA packaging and protein-protein interaction leads to testable predictions consistent with our observations.

RESULTS AND DISCUSSION

Development of an assembly model

A basic scheme has been used to describe the equilibrium of a series of concurrent reactions in which one subunit at a time is assembled to form an empty capsid, and, by analogy, an NA-filled capsid (Fig. 1 *A*). In both cases, each subunit is multivalent, often tetravalent dimers (e.g., HBV and CCMV) or penta- and hexavalent pentamers (e.g., SV40). For the empty capsid (30,31), each intersubunit contact has a microscopic association constant of K_{contact}. The first two subunits interact with K_{contact}; the third subunit may then interact with both subunits of the growing complex with ${K_{\text{contact}}}^2$. The microscopic association constants must be modified by a statistical term, s, particular to the geometry of each intermediate, to account for reaction degeneracy. This is sufficient information to describe the equilibrium concentration of an N-subunit capsid and each n-subunit intermediate. It is important to note that this model of empty-capsid formation predicts a steeply descending energy surface (29) such that free subunits ($n = 1$) and complete N-subunit capsids are the dominant species with infinitesimal concentrations of intermediates (29,30).

The mass-action laws describing assembly of an empty capsid and n-subunit intermediates (n-mers) are

$$K_{\text{capsid}} = \frac{[\text{capsid}]}{[\text{Cp}]^N} \text{ or } [\text{capsid}] = K_{\text{capsid}}[\text{Cp}]^N \quad (1)$$

$$K_{n\text{-mer}} = \frac{[n\text{-mer}]}{[\text{Cp}]^n} \text{ or } [n\text{-mer}] = K_{n\text{-mer}}[\text{Cp}]^n. \quad (2)$$

The equilibrium constants can be broken down to components

$$K_{\text{capsid}} = \prod s_j K_{\text{contact}}^{Nc/2} \quad (3)$$

$$K_{\text{capsid}} = \prod s_j K_{\text{contact}}^{nc/2}, \text{ where } nc/2 < Nc/2. \quad (4)$$

In Eqs. 3 and 4, for a capsid of N subunits or an intermediate of n subunits, $\prod s_j$ is the product of the j degeneracy terms peculiar to each intermediate, K_{contact} is the microscopic association constant common to every protein-protein contact, and the exponent (the superscript) is the total number of contacts made, where each multivalent subunit makes up to c contacts. There are critical simplifications explicit in these equations, namely, that off-path complexes make no contribution to the reaction and that all assembly reactions eventually reach equilibrium. However, assembly can go off path, stop short of equilibrium, or otherwise run into kinetic barriers (29,31,32). Kinetic simulations consistent with the above formalism do a remarkable job of matching kinetic and equilibrium data (see, e.g., Kler et al. (33) and Zlotnick and Mukhopadhyay (34)).

Adding NA to the assembly reaction requires recasting of the capsid-assembly formalism (Fig. 1). Association of CPs to form an NA-filled virus needs to include terms for protein-protein interaction, protein-NA interaction, and the energy of packaging the NA. The approach taken by McGhee and von Hippel (24,25) to describe assembly of protein on a linear NA substrate is instructive: K_{NA} (in units of M^{-1}) was used to describe the association of one protein to substrate and the interaction between adjacent proteins was defined by the unitless association coefficient ω. McGhee and von Hippel also made the assumption that proteins would bind several consecutive nucleotides, and they rigorously considered the effects of bound protein occluding several sites. In low-cooperativity reactions (low-ω reactions), the frequent gaps between proteins led to apparent negative cooperativity of association, a barrier that arose because the gaps could not fit an integral number of subunits. However, such exclusion effects are not appropriate for virus CPs, where a protein may bind a discontinuous segment of NA and the nucleoprotein contacts may be much more fluid.

We consider that assembly of a virus must include terms for 1), binding NA, 2), protein-protein interaction, and 3), constraining NA in a shell. Following McGhee and von Hippel, we define the association constant for NA, K_{NA}, in units of M^{-1}. The protein-protein interaction is closely related to K_{contact} but is now unit-less, because it describes the intramolecular interaction between two or more NA-bound proteins; it is thus appropriate to use the

McGhee-von Hippel term ω. The ω term is not identical to K_{contact} as it does not include the entropic costs of diffusion and loss of rotational freedom inherent in K_{contact}. These costs are now built into K_{NA}. For NA-based assembly, the first protein binds to NA with an association of K_{NA}. When a second protein is added to the NA, if it is adjacent to the first, the binding constant is ωK_{NA}. If the third protein binds to a doubly contiguous site, its association constant will be based on $\omega^2 K_{\text{NA}}$. Of course, these products of microscopic association constants must also have a degeneracy component, s.

Unlike the one-dimensional McGhee-von Hippel DNA lattice, viral CPs must also constrain NA that would normally occupy a volume much larger than the capsid interior (14,17,18). In the case of dsDNA, the substrate favors a volume substantially greater than typical for a virus capsid (13,35). To account for the effect of NA conformation on assembly, we introduce the coefficient α as a component of the assembly association constant. As the growing capsid will progressively change the conformational space available to the NA, there must be a unique α for each added subunit. For simplicity of expression, we will only singly subscript α with the number of subunits in the intermediate, though ideally it may be appropriate to acknowledge the many possible arrangements of subunits and conformations of NA. When the NA impedes assembly, α will have a value between 0 and 1, attenuating association; when the NA promotes assembly, α will be ≥ 1.

Considering an assembly reaction at equilibrium, the mass-action expressions for a virus-like particle consisting of a single NA molecule packaged by a capsid of N subunits (VLP) and for n-subunit intermediates are

$$K_{VLP} = \frac{[\text{capsid}]}{[\text{Cp}]^N[\text{NA}]} \quad \text{or} \quad [\text{capsid}] = K_{VLP}[\text{Cp}]^N[\text{NA}] \tag{5}$$

$$K_{n\text{-}mer} = \frac{[n\text{-mer}]}{[\text{Cp}]^n[\text{NA}]} \quad \text{or} \quad [n\text{-mer}] = K_{n\text{-}mer}[\text{Cp}]^n[\text{NA}]. \tag{6}$$

The corresponding evaluations of the association constant include α and ω terms. Since both degeneracy and α terms are unique to each intermediate, they are included within the product. The contribution of ω, the same for all contacts, is readily described with an exponent of the number contacts.

$$K_{\text{VLP}} = \prod s_j \alpha_j \omega^{Nc/2} K_{\text{NA}}^N \tag{7}$$

$$K_{n\text{-mer}} = \prod s_j \alpha_j \omega^{nc/2} K_{\text{NA}}^n \tag{8}$$

These simple equations describe complex behavior. Most of the tradeoff is between the conflicting forces described by α and ω (and ω^{contacts}). When α is ~1, the resulting binding isotherm shows cooperativity based entirely on ω; because a large exponent is applied to ω, binding curves under conditions of $\alpha \sim 1$ will always show a strong positive cooperativity. To generate low-cooperativity, as seen with CCMV CP binding to RNA (15), the overall value of α must be <1. Low cooperativity will be accentuated when α decreases as n approaches N, so that the product of α and ω^{contacts} is nearly constant throughout the binding isotherm. This is reminiscent of the closure difficulties observed in coarse-grained simulations where the encapsidated NA was too large for the particle (19). The hypothesized counterpoint between α and ω in CCMV disguises the effect of multivalent interactions between CP subunits during capsid closure and results in a binding curve that appears hyperbolic.

Now consider reactions where the substrate hinders initial steps but supports subsequent assembly. When $1/\alpha$ is similar to but smaller than ω, at low CP concentrations there will be weak CP-CP interaction and little evidence of capsid formation (the concentration dependence of binding to NA is defined only by K_{NA}). However, as the CP concentration increases, the exponential effect defined in Eqs. 5 and 6 will become dominant, resulting in a steeply sigmoidal binding curve, as seen with SV40 binding to dsDNA (16).

Because virus-assembly reactions are prone to kinetic traps, it is necessary to expand Eqs. 7 and 8 to allow consideration of kinetic effects. Assembly kinetics of a capsid population can be simulated with a series of ordinary differential equations for each intermediate assuming addition of one subunit at a time. The mathematical approach is consistent with high and low cooperativity models (Fig. 1); the result reasonably mimics formation of empty capsids (36,37) and capsids assembled on RNA (33). For empty capsids, such ordinary differential equation simulations closely match coarse-grained dynamic simulations (38,39). Equation 9 describes assembly of an n-mer with only one possible predecessor and scion (a simplified form of Eq. 9, lacking an explicit α, was shown to fit observed assembly kinetics for SV40 CP binding to a short RNA (33)):

$$\begin{aligned}\frac{\text{d}[n\text{-mer}]}{\text{d}t} &= f s_{fn} \alpha_n^{\ddagger} [(n-1)\text{-mer}][\text{CP}] \\ &\quad - f s_{fn+1} \alpha_{n+1}^{\ddagger} [n\text{-mer}][\text{CP}] \\ &\quad + b_{n+1}[(n+1)\text{-mer}] - b_n[n\text{-mer}].\end{aligned} \tag{9}$$

In the first two, forward, terms of this equation, f is the microscopic forward rate that is modified by the forward degeneracy term s_f and the time-dependent NA conformational coefficient $\alpha^{\ddagger}$; although $\alpha^{\ddagger}$ is different from the equilibrium α, the values for the two terms are likely to be similar. The backward rate terms, b, are based on the forward rate modified by the energetic cost of dissociation, that is, the number of CP-CP contacts broken (ω to an

appropriate exponent) and breaking the CP-NA interaction (K_{NA}). It is clear that a small number of CP-CP contacts or apparent-novel contacts with weak ω values will lead to dissociation of fragile complexes. The multiplicity of possible paths has a modest effect for assembly of empty capsids (40). However, for assembly on NA, each incoming subunit has unique $\alpha^{\ddagger}$ that may vary substantially depending on the particular assembly path and thus has a unique ability to redirect assembly. For simplicity of exposition, we use a single indexing scheme in Eq. 9. The important point is that assembly on NA may be particularly sensitive to local energy minima. As with empty particles, kinetic traps fall into categories of 1), overnucleation leading to accumulation of partial capsids with insufficient CP to allow completion; 2), high association energy, which leads to stalling reactions short of equilibration due to paucity of intermediates; or 3), failure of subunits incorporated with incorrect geometry to dissociate. Strong CP-CP association makes assembly more prone to the first two forms of kinetic trap. Due to antagonism between α and ω, the last sort of failure is of particular importance for NA-directed assembly.

A phase diagram of assembly

A phase diagram is a natural way to view correlated parameters. Siber and co-workers developed phase diagrams to investigate the effects of RNA and protein charge on capsid size but did not explicitly consider the CP-CP interaction or the work required to constrain the NA in a small volume (see 62). Equations 5–8 indicate tradeoffs and synergism between the suitability of an NA for packaging, α, and the strength of protein-protein interaction, ω; these are globally illustrated by a phase diagram (Fig. 3). The major effect of ${K_{NA}}^{N}$ is to shift assembly to lower subunit concentration. As α and ω are unitless association constants, they are easily scaled and understood in terms of their natural logs, proportional to their free energies. Because ω is constant throughout assembly, whereas α changes discontinuously, we will construct the phase diagram comparing $\ln\omega$ to the average term $\ln\langle\alpha\rangle$, acknowledging that particularly in the earliest steps of assembly α and $\langle\alpha\rangle$ may differ substantially. The free energy for assembly of an N-subunit VLP would thus be

$$\Delta G_{VLP} = -RTN \ln\left(\langle\alpha\rangle + \omega^{\text{contacts}/2} + K_{NA}\right) \qquad (10)$$

Considering ω independently, strong protein-protein interactions are expected to result in highly cooperative assembly (where high cooperativity is a term relative to other examples of cooperative capsid assembly). The weakest reasonable value for ω is 1 ($\ln\omega = 0$), indicating negligible interaction. The value of ω is analogous to $K_{contact}$. Because subunits are multivalent, making several equivalent independent contacts, the pseudocritical concentration of assembly of an empty capsid is approximately ${K_{contact}}^{-(\text{contacts}/2)}$, in the micromolar range for HBV and the nanomolar range for SV40. Assembly of empty capsids is inhibited when $K_{contact}$ is too strong due to kinetic traps that lead to errors in assembly and overnucleation. At modest values, assembly simulations indicate that reversibility is critical for releasing improperly bound subunits (29,39). ω is expected to behave in much the same way as $K_{contact}$, correlating with increased cooperativity until, at high ω values, irreversibly incorporated defects and overnucleation lead to defective assembly.

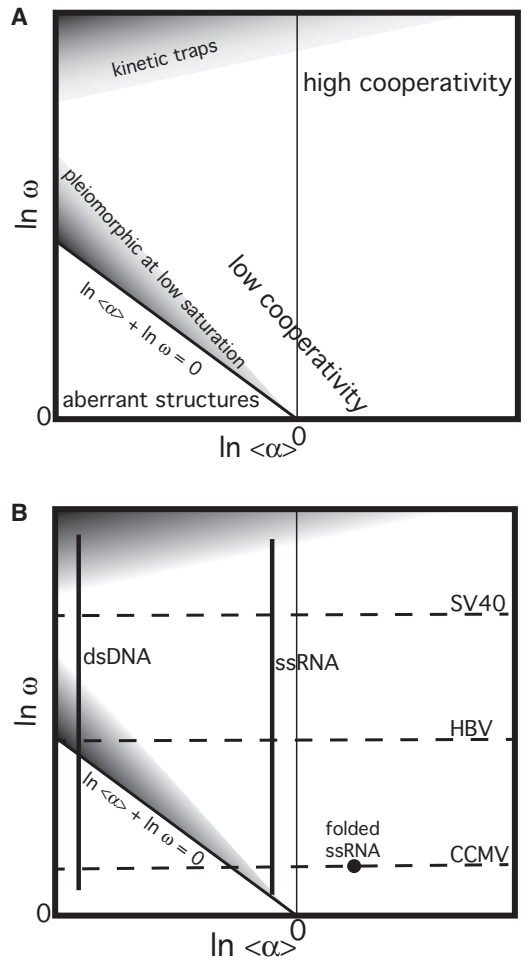

FIGURE 3 Phase diagrams of virus assembly. (*A*) Diagram highlighting stable and metastable phases. Low cooperativity may arise from weak CP-CP interaction, NA that is an unsuitable substrate for assembly, or (as described by McGhee and von Hippel (25)) the inability of CP to redistribute on NA to optimize interaction. High cooperativity is driven by high $\ln\omega$ and high $\ln\langle\alpha\rangle$. (*B*) Overlay of assembly properties of three viruses and two NAs on the virus-assembly phase diagram. The deduction that ssRNA has a negative $\ln\langle\alpha\rangle$ is based on the volume normally occupied by a viral RNA (48,61) and observations that CCMV assembly on RNA normally shows low cooperativity, and that substoichiometric concentrations of CCMV CP cause RNA to compact to a form that is encapsidated with high cooperativity (15).

The effects of $\langle\alpha\rangle$ on assembly, and the structural basis for effect, are more complex. When the NA (or non-NA)

substrate favors assembly, $\ln<\alpha>$ will be greater than zero. A spherical nanoparticle may be an ideal embodiment of a high $\ln<\alpha>$ substrate (41,42). When the substrate disfavors assembly, as stiff linear dsDNA would, $\ln<\alpha>$ will be less than zero. The evaluation of $<\alpha>$ is possible empirically. We suggest that the underlying physics for α is related to the work done on NA by capsid protein. The work done on the NA may include folding, restraining movement, and compaction. In general, viral ssRNA is flexible, highly branched, and relatively compact, but when unconstrained, it is still larger than the interior of its cognate capsid (17,43). dsDNA, on the other hand, is a stiff linear polymer (with a persistence length greater than or similar to the capsids diameter of most viruses (13,35)) that must be bent to conform to the capsid of a small virus, which may involve substantial work distributed over a small number of subunits. Work done by the NA includes restraining and orienting the capsid proteins and also possibly folding their RNA-binding domains. In addition, α may also include the energetic expense of distorting protein-protein interactions (i.e., ω) from ideality. Thus, at each step, α contributes to the free energy of assembly:

$$-RT\ \ln\ \alpha\ =\ \Delta E_{\mathrm{NA\_folding}} + \Delta E_{\mathrm{CP\_binding}} + \Delta E_{\mathrm{CP\text{-}CP\ distortion}}. \tag{11}$$

The phase diagram of virus assembly, $\ln\omega$ versus $\ln<\alpha>$, has several distinct zones (Fig. 3 and Table 1). Where $\ln<\alpha>$ is positive, increased $\ln\omega$ corresponds to increasingly cooperative assembly. As $\ln<\alpha>$ increases, the K_{VLP} product increases, also enhancing cooperativity. There is a low-cooperativity region of the phase diagram; low cooperativity can result from a combination of low $\ln\omega$ (i.e., from weak CP-CP interaction) and negative $\ln<\alpha>$ (NA that is a poor substrate for assembly). Not on this phase diagram, the inability of CP to redistribute on NA to optimize interaction will result in unfillable gaps giving rise to the appearance of negative cooperativity (25).

Where $\ln<\alpha>$ is less than zero, the phase diagram is more complex. The line described by $(\ln<\alpha> + \ln\omega) = 0$ is a demarcation between VLP assembly and aberrant particles. Below this line it is energetically more favorable either to have subunits bind NA independent of one another or to find an alternative, noncapsid protein-protein interaction. Even above the $(\ln<\alpha> + \ln\omega) = 0$ line, assembly may not yield VLPs.

TABLE 1 Interplay of α and ω

Condition	Response
$\alpha \geq 1$ and $\omega \gg 1$	Assembly is highly cooperative
$\alpha > 1$	NA structure promotes assembly (increased cooperativity)
$\alpha < 1$	NA structure inhibits assembly
$\alpha < 1$ and $\alpha\omega > 1$	CP-CP interaction trumps NA structure
$\alpha < 1$ and $\alpha\omega < 1$	NA structure dominates assembly

Although it will not be investigated in detail in this article, the generalized description of assembly (Fig. 1) implies a third dimension of $\ln(K_{\mathrm{NA}})$. One can expect $\ln(K_{\mathrm{NA}})$ to be responsive to ionic strength or to CP charge changed by mutation. As $\ln(K_{\mathrm{NA}})$ decreases, the tendency to assemble empty capsids can be expected to increase where the sum $(\ln\omega + \ln\alpha)$ is close to or less than 0. In accordance with the predictions of CP adsorbing to the surface of a sphere (44), one can predict that weaker $\ln(K_{\mathrm{NA}})$ will tend to lead to stepwise cooperative assembly of capsids; strong $\ln(K_{\mathrm{NA}})$, particularly in the presence of weak $\ln\omega$, will favor formation of a disordered complex of CP associated with NA.

Kinetic effects on assembly make their appearance on the phase diagram in two regions due to predicted defects in assembly: where $(\ln<\alpha> + \ln\omega)$ is slightly greater than zero and where $\ln\omega$ is very high. Near the $(\ln<\alpha> + \ln\omega) = 0$ border, early intermediates in assembly may be dominated by a relatively negative $\ln\alpha$, so it is likely that they will not resemble on-path complexes. Furthermore, if subsequent subunits are not able to participate in remodeling the complex, the result will be a region where, although $(\ln<\alpha> + \ln\omega) > 0$, aberrant structures form a metastable state. The shape of this region will be dependent on assembly path but is likely to be broader at progressively more negative values of $\ln\alpha$ and larger values of $\ln\omega$.

In a similar way, at high values of $\ln\omega$, necessarily, the nature of the kinetic trap depends on the path where subunits are added to the growing polymer. At high values of ω, assembly reactions of empty capsids can be trapped, because the reaction runs out of subunits to complete capsids, because improperly bound subunits do not dissociate, and because reactions become starved for intermediates (29,32,34,39). Defects in assembly due to high $\ln\omega$ will be path-dependent and will be further stabilized by the strength of K_{NA}. Like reactions at the $(\ln<\alpha> + \ln\omega) = 0$ border, we suggest that high $\ln\omega$ defects are more likely to form and be trapped at negative values of $\ln<\alpha>$. Although they are not thermodynamic phenomena, these traps have a place in demarcating the phase space of successful assembly.

The behavior of NA or other packaged cargo will vary almost continuously during assembly. The value of α can be thought of as a complex order parameter so that α and the phase diagram will be unique for each possible intermediate. Nonetheless, the phase diagrams outlined above allow a straightforward analysis of real assembly data below.

Three examples of assembly: CCMV, HBV, and SV40

Experimental observations of virus assembly can be understood in terms of an $\alpha\omega K_{\mathrm{NA}}$ product. We will consider three examples of virus assembly: CCMV, HBV, and SV40 (summarized in Table 2). In each case, we know something about assembly of empty particles, assembly on ssRNA, and

TABLE 2 Summary of CCMV, HBV, and SV40 assembly properties

	CCMV	HBV	SV40
$K_{Dapparent}$	>75 μM (46)	~10 μM (51)	$\ll$0.1 μM (53)
$K_{contact}$*	>9 mM	~3mM	$\ll$1 mM
Substrate†			
ssRNA	Low cooperativity (15)	High cooperativity (21)	High cooperativity
dsDNA (low ν)	Condensate (49)	Condensate (12)	Condensate (16)
dsDNA (high ν)	Tubes (49)	T = 4, aberrant (12)	T = 7 (16)

*$K_{contact}$ is proportional to the $c/2$ root of $K_{Dapparent}$, where c is the number of contacts/subunit (29).
†For dsDNA, results are reported for low and high occupancy, ν.

assembly on dsDNA. Where apparent cooperativity is low, electrophoresis shows that all of the NA in a given sample has a similar amount of bound CP. Where cooperativity is high, we expect, and find, a bimodal distribution of free and bound NA (Fig. 1, *B* and *C*).

CCMV was the first spherical virus assembled in vitro, forming both empty capsids and infectious, RNA-filled particles (45). Virions are comprised of 28-nm-diameter T = 3 capsids assembled from 90 dimers. Empty particles readily assemble from purified CP in a pH-dependent reaction. At pH 5, assembly is observed at >2 μM CP dimer (36); assembly is not observed at neutral pH at concentrations up to 75 μM (45). Assembly of empty particles depends on formation of pentamers of dimers (46,47). CCMV assembly on RNA shows low cooperativity (Fig. 2 *A*). Viral RNA mixed with CP and electrophoresed through a native agarose gel 20 min after mixing shows a heterogeneous mixture of partial capsids at partial saturation, indicating low cooperativity (15). After 24 h incubation, the same reactions show a bimodal distribution of free RNA and VLP, indicative of a reorganization of CP to a lower energy state due to a larger number of CP-CP interactions, analogous to Ostwald ripening. Indeed, small amounts of CCMV CP cause RNA to fold into a more compact structure that migrates faster on a gel and supports highly cooperative assembly of capsids; that is, substoichiometric concentrations of CP are able to compact the RNA to prime further assembly (15); a similar, but sequence-specific, NA compaction was recently reported for two other RNA viruses (bacteriophage MS2 and Sstellite tobacco necrosis virus) (48). Thus, compaction may be an evolved mechanism for increasing α. Indeed, NA encoding high-affinity sites for CP, especially if those sites are strategically spaced, is a specific realization of NA with a (relatively) positive lnα. Conversely, when CCMV CP was mixed with low-α dsDNA at neutral pH, it assembled into 17-nm-diameter tubes, not capsids at all (Fig. 4) (49). When tubes are acidified, strengthening protein-protein interaction, the subunits reorganized, dissociating from the dsDNA to form empty T = 1 capsids (47).

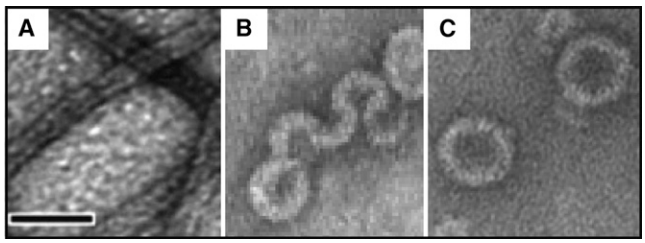

FIGURE 4 Electron micrographs showing the diversity of assembly products when different viral coat proteins associate with a 600-base pair dsDNA (ln$\alpha \ll 0$). (*A*) CCMV CP (low lnω) assembly is completely redirected to form 17-nm-diameter tubes with a core of parallel overlapping dsDNA molecules (49). (*B*) HBV CP has a moderate lnω, similar to the lnα of dsDNA, resulting in aberrant particles showing conformational influences from both macromolecules as the CP unsuccessfully tries to form 35-nm-diameter spheres (12). (*C*) The CP of SV40 (high lnω) overcomes the stiffness of dsDNA to yield virus-like 45-nm-diameter capsids (from Mukherjee et al. (16)). All micrographs were stained with uranyl acetate. Scale bar, 50 nm.

CCMV can be put in the context of the $\alpha\omega$ phase diagram (Fig. 3 *B*). The very weak association energy at neutral pH, demonstrated by failure to form empty capsids, indicates that ω is small, <80 based on extrapolating its low-pH pseudocritical concentration to neutral pH (36,46). However, when substoichiometric CP assembly refolds the ssRNA into a more compact form, assembly becomes cooperative. As even a weak ω leads to high cooperativity of assembly, and CP is the same with both unfolded and compacted RNA, this leads to the conclusion that unfolded ssRNA must have a lnα value that is <0, whereas when RNA is compact, lnα becomes >0, allowing ω to drive cooperative assembly. With either RNA state, the $\alpha\omega K_{NA}$ product is sufficient to drive capsid assembly. On the other hand, for dsDNA the lnα term must be substantial and negative, resulting in a (lnα + lnω) sum that is <0, favoring a noncapsid alternative protein-protein interaction, so that dsDNA redirects assembly.

HBV is a major cause of human disease, contributing to ~600,000 deaths each year. HBV has an unusual lifecycle in that it is a dsDNA virus with an ssRNA intermediate (50): its 120-dimer T = 4 capsid assembles on viral RNA, which is then reverse-transcribed within the capsid. HBV is also one of the best-characterized virus assembly systems (34), where ω, approximated by $K_{contact}$, under physiological conditions is ~150–300 (51). The association of the HBV CP for ssRNA and ssDNA is driven by an RNA-binding domain with a net charge of +30/dimer and is very strong (12,21). When assembled on RNA, CP binding is highly cooperative and yields capsids with virus-like morphology, ~35 nm in diameter (21). However, when dimeric HBV CP is added to dsDNA, the resulting complexes are heterogeneous (12). In electron micrographs it can be seen that the majority of HBV CP-DNA particles appear to be spirals and aggregates of capsid fragments, with possibly a few normal capsids (Fig. 4). Nonetheless, in vivo, all infectious HBV

virions contain a circular, almost entirely double-stranded DNA genome.

HBV CP crosses the $\alpha\omega$ phase diagram along a different line than CCMV. The lnω for HBV is substantially greater than lnω for CCMV, corresponding to highly cooperative assembly on ssRNA (ln$\alpha \leq 0$ based on analysis of CCMV assembly). Because of its higher lnω, because infectious HBV does contain dsDNA, and because at least some HBV assembled on dsDNA appears to form VLPs, HBV CP with a dsDNA substrate therefore has an overall ln$\alpha\omega > 0$. However, the heterogeneity of dsDNA assembly products (Fig. 4) indicates that ln$\alpha\omega < 0$ for some early intermediates in the assembly process. The structure of any given assembly product therefore depends on its stochastic starting point and assembly path.

Unlike HBV or CCMV, SV40 is a dsDNA virus that assembles on a dsDNA substrate in vivo, albeit NA is compacted by bound histones. SV40 has unusual architecture, basing a 45-nm-diameter T = 7 capsid of 72 CP pentamers. In vitro, SV40 is capable of packaging bare dsDNA, up to twice the length of its 5200-bp genome; these complexes have been used as vectors for transfection (20,52). Assembly of empty SV40 capsids can be induced by calcium, high ionic strength, or pH, but this rapid and quantitative reaction yields pleiomorphic results consistent with its very high association energy (53). Assembly can also be catalyzed by chaperones in vitro, indicating the presence of a kinetic, not thermodynamic, barrier to formation of virus-like capsids (54). Assembly of SV40 on ssRNA is highly cooperative and yields mainly 22-nm-diameter T = 1 particles; with longer RNA ($\geq$1200 nucleotides), a few 35-nm particles are also observed (33). On dsDNA, SV40 has a sigmoidal assembly curve, suggesting that there is a barrier to initiating assembly but that it proceeds cooperatively thereafter (16). Assembly products with DNA are uniformly 45 nm in diameter (Fig. 4).

The very strong association constant of SV40 CP indicates a high value for ω; the ω value of the structurally similar Human Papillomavirus 16 is on the order of 1000 (55). With an RNA substrate, assembly is aggressive and pleiomorphic, suggesting the presence of kinetic traps, errors in assembly that could not be corrected, presumably because of slow dissociation and rearrangement of subunits. In terms of the phase diagram, this places SV40 near the border high cooperativity and kinetic traps. However, even with ln$\alpha \ll 0$, the lnα + lnω sum is still substantially greater than 0, and normal capsids arise with a sufficiently high CP concentration.

An important confirmatory observation is that with a dsDNA substrate (with ln$\alpha \ll 0$), low concentrations of CPs from CCMV, HBV, and SV40 all induced formation of large cylindrical DNA condensates (12,15,16). Condensates are a stable or metastable phase of DNA, typically induced by multivalent cations, where DNA strands are approximately parallel, with a minimal spacing of typically 2.7 nm (13,35); these condensates usually are toroidal but may also be rodlike. Such complexes are incompatible with capsid-like CP-CP interactions. In these complexes, the CPs may be noncovalently cross-linked and bundle many dsDNA polymers. The assembly conditions of these condensates, low lnα and low [CP], are consistent with conditions where negative lnα and low CP concentration effectively eliminate the role of ω in assembly. Thus, the condensate assembly reaction is based on K_{NA}.

CONCLUSIONS

A substantial number of spherical viruses package their genomes during the process of capsid assembly. Besides the examples discussed in this article (Bromoviridae, Hepadnaviridae, and Polyomaviridae), families of viruses that assemble in conjunction with packaging their genomes include Flaviviridae, Togaviridae, Papillomaviridae, and possibly Retroviridae. Although the biological details are unique to each virus, the underlying physics will remain the same. The challenge we address in this work is to place the sometimes-conflicting factors governing assembly and NA packaging into a unified framework. Here, we consider assembly as a reaction where each step is the product of terms of binding NA (K_{NA}), protein-protein interaction (ω), and the work of encapsidating the NA (α). From this analysis of assembly, it can be argued that NA plays a particularly dominant role in directing at the earliest stages in the process. RNA compaction by CP (observed in CCMV and bacteriophage MS2 (15,48)) and DNA compaction by histones (in SV40) are viral mechanisms for increasing α. Later in a virion's assembly, the assembly energy related to ω becomes dominant. Where assembly is able to equilibrate, conflicts between NA and CP can resolve. The phase diagram indicates conditions where assembly can be trapped due to the stability of an early assembly product. In those regions, in vitro assembly will be frustrated and in vivo assembly will require chaperones or a multiphase assembly process to avoid energetic bottlenecks (e.g., HBV assembly on an RNA form of its genome). This logic suggests that evolutionary pressures select for coordination between ω and α. Indeed, this coordination may be useful for understanding the speciation of viruses and for learning how to manipulate them.

The description of assembly of an NA-filled capsid is directly applicable to experimental data. The value of ω for the interaction between two subunits is approximately $K_{contact}$. The two terms differ because $K_{contact}$ includes entropic loss of freedom due to protein association (56). The entropic cost of association, implicit in $K_{contact}$ and K_{NA}, but not in ω, has been experimentally estimated to be -5 ± 4 cal M^{-1} deg^{-1}, about +1.5 kcal/mol under physiological conditions (56), which is substantial compared to the −3.5 kcal/mol energy/contact determined for HBV (51).

Consider the role of ω in assembly. A series of HBV CP mutations (observed in chronic infections) that modestly

enhance ω may provide an advantage to the virus when production of the CP is suppressed by the host (29,57). Small alterations at the HBV CP-CP interface can have catastrophic results. Small molecules that enhance HBV assembly, increasing ω, disrupt the timing and organization of virion formation (34). Thus, we suggest that virus viability is very sensitive to ω. The phase diagrams of Fig. 3 provide a predictive guideline to effects of ω.

Determination of α will be difficult. The average value of α, $<\alpha>$, can be estimated from the difference between association of an empty capsid, association of capsid protein to NA in the absence of assembly, and assembly of an NA-filled capsid. Perhaps a more powerful method of evaluating α is to compare NA substrates to a standard. The variant of length has been examined in silico (18) and to some degree in vitro (20). However, the effects on α of NA length, sequence, local, and higher-order structure have not been examined systematically in solution, though presumably these factors will be critical for in vivo NA packaging and for taking advantage of the NA capacity of viruses as NA carriers in culture and gene therapy. Though $<\alpha>$ can be determined experimentally from bulk assembly experiments, understanding the basis for α and how it changes during assembly will almost certainly require single-molecule experiments and coarse-grained simulations.

Consider the role of α. Viral RNAs tend to be more compact than random RNA (17,43). The term α provides a quantifiable index of the selective pressure for different NA tertiary structures (7,48). Besides the basic biology of infection, viruses can find use in bio- and nanotechnology. This application makes the work required to package nonviral NAs a question of great importance (20,52). This question can also be applied to non-NA polymers (14) and solid substrates (42,58,59).

The role of K_{NA} is critical to assembly. Without a strong CP-NA interaction there will certainly be a strong entropic barrier to NA-packaging, as predictable from the explicit assumptions developed in Eqs. 5–10. The strength of K_{NA} will contribute to the ability of a CP to package any NA and will be absolutely necessary for NA with $\alpha < 1$. Values for K_{NA} are not generally known. An SV40 VP1 pentamer has a nanomolar dissociation constant for DNA (60). The K_{NA} of HBV and CCMV for different NAs has not been quantified, but it appears to be strong given their affinity for dsDNA despite the fact that dsDNA induces unusual CP-CP interactions (12,49).

The ability of a virus to package its own genome is central to biology. This same process can be used to package nonviral cargo and harness viruses to designed tasks. The ability of a virus to assemble on a given substrate requires either adapting the substrate to the virus, a function of α, or adapting the virus CP to the substrate, a function of the product $\alpha\omega$. Learning how to regulate α and ω will open new opportunities to develop viruses as tools.

A 2011 meeting at the Kavli Institute for Theoretical Physics provided the impetus for first description of these concepts. We acknowledge discussions with Fred Mackintosh, Robijn Bruinsma, and Boris Shklovskii on the physics of assembly, David Reguera on the effect of K_{NA} on the phase diagram, and Ariella Oppenheim on the biochemistry of SV40.

This work was supported by National Institutes of Health grant R01-AI077688 to A.Z.

REFERENCES

1. Schneemann, A. 2006. The structural and functional role of RNA in icosahedral virus assembly. *Annu. Rev. Microbiol.* 60:51–67.
2. Gouet, P., J. M. Diprose, …, P. P. Mertens. 1999. The highly ordered double-stranded RNA genome of bluetongue virus revealed by crystallography. *Cell.* 97:481–490.
3. Tang, L., K. N. Johnson, …, J. E. Johnson. 2001. The structure of pariacoto virus reveals a dodecahedral cage of duplex RNA. *Nat. Struct. Biol.* 8:77–83.
4. Larson, S. B., S. Koszelak, …, A. McPherson. 1993. Double-helical RNA in satellite tobacco mosaic virus. *Nature.* 361:179–182.
5. Chen, Z. G., C. Stauffacher, …, J. E. Johnson. 1989. Protein-RNA interactions in an icosahedral virus at 3.0 Å resolution. *Science.* 245:154–159.
6. Tihova, M., K. A. Dryden, …, A. Schneemann. 2004. Nodavirus coat protein imposes dodecahedral RNA structure independent of nucleotide sequence and length. *J. Virol.* 78:2897–2905.
7. Schroeder, S. J., J. W. Stone, …, D. M. Mathews. 2011. Ensemble of secondary structures for encapsidated satellite tobacco mosaic virus RNA consistent with chemical probing and crystallography constraints. *Biophys. J.* 101:167–175.
8. MacFarlane, S. A., M. Shanks, …, G. P. Lomonossoff. 1991. Analysis of the nucleotide sequence of bean pod mottle virus middle component RNA. *Virology.* 183:405–409.
9. Dryden, K. A., S. F. Wieland, …, M. Yeager. 2006. Native hepatitis B virions and capsids visualized by electron cryomicroscopy. *Mol. Cell.* 22:843–850.
10. Wang, J. C., M. S. Dhason, and A. Zlotnick. 2012. Structural organization of pregenomic RNA and the carboxy-terminal domain of the capsid protein of hepatitis B virus. *PLoS Pathog.* 8:e1002919.
11. Petrov, A. S., and S. C. Harvey. 2008. Packaging double-helical DNA into viral capsids: structures, forces, and energetics. *Biophys. J.* 95:497–502.
12. Dhason, M. S., J. C.-Y. Wang, …, A. Zlotnick. 2011. Differential assembly of Hepatitis B Virus core protein on single- and double-stranded DNA suggests the dsDNA core is spring-loaded. *Virology.* 430:20–29.
13. Tzlil, S., J. T. Kindt, …, A. Ben-Shaul. 2003. Forces and pressures in DNA packaging and release from viral capsids. *Biophys. J.* 84:1616–1627.
14. Hu, Y., R. Zandi, …, W. M. Gelbart. 2008. Packaging of a polymer by a viral capsid: the interplay between polymer length and capsid size. *Biophys. J.* 94:1428–1436.
15. Johnson, J. M., D. A. Willits, …, A. Zlotnick. 2004. Interaction with capsid protein alters RNA structure and the pathway for in vitro assembly of cowpea chlorotic mottle virus. *J. Mol. Biol.* 335:455–464.
16. Mukherjee, S., S. Kler, …, A. Zlotnick. 2010. Uncatalyzed assembly of spherical particles from SV40 VP1 pentamers and linear dsDNA incorporates both low and high cooperativity elements. *Virology.* 397:199–204.
17. Yoffe, A. M., P. Prinsen, …, A. Ben-Shaul. 2008. Predicting the sizes of large RNA molecules. *Proc. Natl. Acad. Sci. USA.* 105:16153–16158.
18. Elrad, O. M., and M. F. Hagan. 2010. Encapsulation of a polymer by an icosahedral virus. *Phys. Biol.* 7:045003.

19. Kivenson, A., and M. F. Hagan. 2010. Mechanisms of capsid assembly around a polymer. *Biophys. J.* 99:619–628.
20. Kimchi-Sarfaty, C., M. Arora, …, M. M. Gottesman. 2003. High cloning capacity of in vitro packaged SV40 vectors with no SV40 virus sequences. *Hum. Gene Ther.* 14:167–177.
21. Porterfield, J. Z., M. S. Dhason, …, A. Zlotnick. 2010. Full-length HBV core protein packages viral and heterologous RNA with similarly high cooperativity. *J. Virol.* 84:7174–7184.
22. McPherson, A. 2005. Micelle formation and crystallization as paradigms for virus assembly. *Bioessays.* 27:447–458.
23. Devkota, B., A. S. Petrov, …, S. C. Harvey. 2009. Structural and electrostatic characterization of pariacoto virus: implications for viral assembly. *Biopolymers.* 91:530–538.
24. Kowalczykowski, S. C., L. S. Paul, …, P. H. von Hippel. 1986. Cooperative and noncooperative binding of protein ligands to nucleic acid lattices: experimental approaches to the determination of thermodynamic parameters. *Biochemistry.* 25:1226–1240.
25. McGhee, J. D., and P. H. von Hippel. 1974. Theoretical aspects of DNA-protein interactions: co-operative and non-co-operative binding of large ligands to a one-dimensional homogeneous lattice. *J. Mol. Biol.* 86:469–489.
26. Zandi, R., and P. van der Schoot. 2009. Size regulation of ss-RNA viruses. *Biophys. J.* 96:9–20.
27. Hu, T., and B. I. Shklovskii. 2007. Kinetics of viral self-assembly: role of the single-stranded RNA antenna. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 75:051901.
28. von Hippel, P. H., and O. G. Berg. 1989. Facilitated target location in biological systems. *J. Biol. Chem.* 264:675–678.
29. Katen, S. P., and A. Zlotnick. 2009. The thermodynamics of virus capsid assembly. *Methods Enzymol.* 455:395–417.
30. Zlotnick, A. 1994. To build a virus capsid. An equilibrium model of the self assembly of polyhedral protein complexes. *J. Mol. Biol.* 241:59–67.
31. Katen, S. P., S. R. Chirapu, …, A. Zlotnick. 2010. Trapping of hepatitis B virus capsid assembly intermediates by phenylpropenamide assembly accelerators. *ACS Chem. Biol.* 5:1125–1136.
32. Morozov, A. Y., R. F. Bruinsma, and J. Rudnick. 2009. Assembly of viruses and the pseudo-law of mass action. *J. Chem. Phys.* 131:155101.
33. Kler, S., R. Asor, …, U. Raviv. 2012. RNA encapsidation by SV40-derived nanoparticles follows a rapid two-state mechanism. *J. Am. Chem. Soc.* 134:8823–8830.
34. Zlotnick, A., and S. Mukhopadhyay. 2011. Virus assembly, allostery and antivirals. *Trends Microbiol.* 19:14–23.
35. Purohit, P. K., J. Kondev, and R. Phillips. 2003. Mechanics of DNA packaging in viruses. *Proc. Natl. Acad. Sci. USA.* 100:3173–3178.
36. Johnson, J. M., J. Tang, …, A. Zlotnick. 2005. Regulating self-assembly of spherical oligomers. *Nano Lett.* 5:765–770.
37. Porterfield, J. Z., and A. Zlotnick. 2010. An overview of capsid assembly kinetics. *In* Emerging Topics in Physical Virology. P. G. Stockley and R. Twarock, editors. Imperial College Press, London. 131–158.
38. Hagan, M. F., and D. Chandler. 2006. Dynamic pathways for viral capsid assembly. *Biophys. J.* 91:42–54.
39. Rapaport, D. C. 2010. Studies of reversible capsid shell growth. *J. Phys. Condens. Matter.* 22:104115.
40. Endres, D., M. Miyahara, …, A. Zlotnick. 2005. A reaction landscape identifies the intermediates critical for self-assembly of virus capsids and other polyhedral structures. *Protein Sci.* 14:1518–1525.
41. Hagan, M. F. 2008. Controlling viral capsid assembly with templating. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 77:051904.
42. Chen, C., E.-S. Kwak, …, B. Dragnea. 2005. Packaging of gold particles in viral capsids. *J. Nanosci. Nanotechnol.* 5:2029–2033.
43. Kuznetsov, Y. G., S. Daijogo, …, A. McPherson. 2005. Atomic force microscopy analysis of icosahedral virus RNA. *J. Mol. Biol.* 347:41–52.
44. Hagan, M. F. 2009. A theory for viral capsid assembly around electrostatic cores. *J. Chem. Phys.* 130:114902.
45. Bancroft, J. B. 1970. The self-assembly of spherical plant viruses. *Adv. Virus Res.* 16:99–134.
46. Zlotnick, A., R. Aldrich, …, M. J. Young. 2000. Mechanism of capsid assembly for an icosahedral plant virus. *Virology.* 277:450–456.
47. Burns, K., S. Mukherjee, …, A. Zlotnick. 2010. Altering the energy landscape of virus self-assembly to generate kinetically trapped nanoparticles. *Biomacromolecules.* 11:439–442.
48. Borodavka, A., R. Tuma, and P. G. Stockley. 2012. Evidence that viral RNAs have evolved for efficient, two-stage packaging. *Proc. Natl. Acad. Sci. USA.* 109:15769–15774.
49. Mukherjee, S., C. M. Pfeifer, …, A. Zlotnick. 2006. Redirecting the coat protein of a spherical virus to assemble into tubular nanostructures. *J. Am. Chem. Soc.* 128:2538–2539.
50. Nassal, M. 2008. Hepatitis B viruses: reverse transcription a different way. *Virus Res.* 134:235–249.
51. Ceres, P., and A. Zlotnick. 2002. Weak protein-protein interactions are sufficient to drive assembly of hepatitis B virus capsids. *Biochemistry.* 41:11525–11531.
52. Mukherjee, S., M. Abd-El-Latif, …, A. Oppenheim. 2007. High cooperativity of the SV40 major capsid protein VP1 in virus assembly. *PLoS ONE.* 2:e765.
53. Kanesashi, S. N., K. Ishizu, …, H. Handa. 2003. Simian virus 40 VP1 capsid protein forms polymorphic assemblies in vitro. *J. Gen. Virol.* 84:1899–1905.
54. Chromy, L. R., J. M. Pipas, and R. L. Garcea. 2003. Chaperone-mediated in vitro assembly of Polyomavirus capsids. *Proc. Natl. Acad. Sci. USA.* 100:10477–10482.
55. Mukherjee, S., M. V. Thorsteinsson, …, A. Zlotnick. 2008. A quantitative description of in vitro assembly of human papillomavirus 16 virus-like particles. *J. Mol. Biol.* 381:229–237.
56. Tamura, A., and P. L. Privalov. 1997. The entropy cost of protein association. *J. Mol. Biol.* 273:1048–1060.
57. Yuan, T. T., G. K. Sahu, …, C. Shih. 1999. The mechanism of an immature secretion phenotype of a highly frequent naturally occurring missense mutation at codon 97 of human hepatitis B virus core antigen. *J. Virol.* 73:5731–5740.
58. Dixit, S. K., N. L. Goicochea, …, B. Dragnea. 2006. Quantum dot encapsulation in viral capsids. *Nano Lett.* 6:1993–1999.
59. Sun, J., C. DuFort, …, B. Dragnea. 2007. Core-controlled polymorphism in virus-like particles. *Proc. Natl. Acad. Sci. USA.* 104:1354–1359.
60. Li, P. P., A. Nakanishi, …, H. Kasamatsu. 2001. Simian virus 40 Vp1 DNA-binding domain is functionally separable from the overlapping nuclear localization signal and is required for effective virion formation and full viability. *J. Virol.* 75:7321–7329.
61. Gopal, A., Z. H. Zhou, …, W. M. Gelbart. 2012. Visualizing large RNA molecules in solution. *RNA.* 18:284–299.
62. Siber, A., R. Zandi, and R. Podgornik. 2010. Thermodynamics of nanospheres encapsulated in virus capsids. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 81:051919.