A Theoretical Model Successfully Identifies Features of Hepatitis B Virus Capsid Assembly†

Adam Zlotnick,*,‡ Jennifer M. Johnson,‡ Paul W. Wingfield,§ Stephen J. Stahl,§ and Dan Endres||

Department of Biochemistry and Molecular Biology, Oklahoma University Health Sciences Center, Oklahoma City, Oklahoma 73190, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, Maryland 20892-2717, and Department of Mathematics and Statistics, University of Central Oklahoma, Edmond, Oklahoma 73074

Received July 13, 1999; Revised Manuscript Received September 13, 1999

ABSTRACT: The capsids of most spherical viruses are icosahedral, an arrangement of multiples of 60 subunits. Though it is a salient point in the life cycle of any virus, the physical chemistry of virus capsid assembly is poorly understood. We have developed general models of capsid assembly that describe the process in terms of a cascade of low order association reactions. The models predict sigmoidal assembly kinetics, where intermediates approach a low steady state concentration for the greater part of the reaction. Features of the overall reaction can be identified on the basis of the concentration dependence of assembly. In simulations, and on the basis of our understanding of the models, we find that nucleus size and the order of subsequent "elongation" reactions are reflected in the concentration dependence of the extent of the reaction and the rate of the fast phase, respectively. The reaction kinetics deduced for our models of virus assembly can be related to the assembly of any "spherical" polymer. Using light scattering and size exclusion chromatography, we observed polymerization of assembly domain dimers of hepatitis B virus (HBV) capsid protein. Empty capsids assemble at a rate that is a function of protein concentration and ionic strength. The kinetics of capsid formation were sigmoidal, where the rate of the fast phase had second-power concentration dependence. The extent of assembly had third-power concentration dependence. Simulations based on the models recapitulated the concentration dependences observed for HBV capsid assembly. These results strongly suggest that in vitro HBV assembly is nucleated by a trimer of dimers and proceeds by the addition of individual dimeric subunits. On the basis of this mechanism, we suggest that HBV capsid assembly could be an important target for antiviral therapeutics.

The question of capsid assembly mechanism is general. Despite their complexity, virus capsids must be able to assemble quickly and with high fidelity in vivo. Spontaneous assembly in vitro has been demonstrated for many viruses and bacteriophages (*1*).

Hepatitis B virus (HBV)[1] is an enveloped DNA virus with an icosahedral nucleocapsid (*2*). In vivo, the HBV nucleocapsid assembles around an RNA−reverse transcriptase complex. The RNA "pregenome" is then reverse transcribed within the assembled capsid, after which the capsid may acquire its protein−lipid envelope and be exported from the cell (*3*). Failure of capsids to assemble properly can interfere with retrotranscription and hence the virus life cycle (*4*, *5*).

The capsid, or core, protein of HBV is a small, 183-amino acid protein with two domains. The N-terminal 149 residues comprise the assembly domain which spontaneously forms capsids when expressed in heterologous systems (*6*, *7*). The assembly domain can be purified as a 34 kDa homodimer and assembled in vitro (*8*, *9*). Dimers assemble at a rate that is a function of protein concentration, ionic strength, and pH (*9*, *10*). Purified capsid protein dimers, which are stable in 0.1 M bicarbonate buffer (pH 9.5), spontaneously assemble when the ionic strength of the buffer is increased to >0.6 M. Heterologously expressed and in vitro-assembled capsids of HBV and related hepdnaviruses are indistinguishable from capsids isolated from infected tissue (*11*). Like infectious virus, the heterologously expressed and in vitro-assembled capsids form a mixture of $T = 3$ and $T = 4$ capsids, composed of 90 and 120 capsid protein dimers, respectively (*9*, *12*−*14*). The full-length assembly domain yields ∼90% $T = 4$ capsids, and C-terminal truncations favor progressively more of the $T = 3$ form (*15*).

Several considerations of icosahedral capsid assembly are extant, though these models are untested and do not provide tools for analyzing experimental observation of assembly (*16*−*18*). These focus on assembly of single particles and do not consider the behavior of populations of viruses, or the problem of regulating assembly. Also, Caspar proposed that capsid assembly could be regulated by "self-control" or autostery, accomplished by unbound protein fluctuating between assembly competent and incompetent states (*19*).

† Supported by Research Project Grant RPG-99-339-01 from the American Cancer Society.

* To whom correspondence should be addressed: Department of Biochemistry and Molecular Biology, University of Oklahoma Health Sciences Center, P.O. Box 26901 BRC 464, Oklahoma City, OK 73190. E-mail: adam-zlotnick@ouhsc.edu. Phone: (405) 271-9030. Fax: (405) 271-3910.

‡ Oklahoma University Health Sciences Center.

§ National Institutes of Health.

|| University of Central Oklahoma.

[1] Abbreviations: HBV, hepatitis B virus; EQ, equilibrium; KL, kinetically limited; LC, liquid chromatography; DTT, dithiothreitol.

10.1021/bi991611a CCC: $18.00
Published on Web 10/14/1999

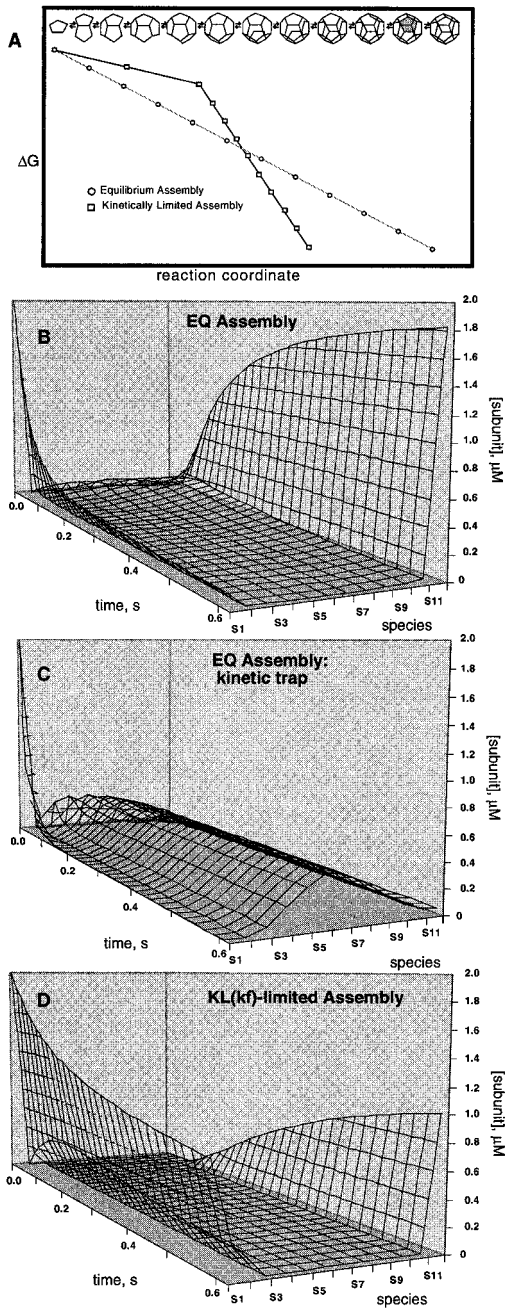

FIGURE 1: Assembling a dodecahedron from pentagonal subunits. (A) Energy diagrams of EQ and KL assembly showing the most stable species along the reaction path. In EQ assembly, all reactions occur at the same rate. In KL assembly, there are rate differences between the nucleation and elongation phases of the reaction. The points reflect the energy levels of intermediates. The energy spacing that is shown is for conceptual simplicity and is not to scale. (B) EQ assembly is successful with an initial subunit concentration of 2 μM and a microscopic association constant of 10^3 M^{-1} $contact^{-1}$. The concentration of each species, monomer through dodecahedron (S1−S12), is shown as a function of time. (C) EQ assembly is stalled by a kinetic trap at 2 μM subunit and an association constant of 10^4 M^{-1} $contact^{-1}$. (D) Incorporation of kinetically limiting steps leading to trimer formation prevents kinetic traps from forming. All reactions in the KL simulation are based on a microscopic association constant of 10^4 M^{-1} $contact^{-1}$. The limiting reactions have a slower forward rate (10^6 M^{-1} s^{-1}, compared to 10^8 M^{-1} s^{-1} for all other reactions).

Previously, we developed the equilibrium (EQ) model (Figure 1A) which describes a stochastic equilibrium between subunits, intermediates, and capsids (*20*). The EQ results emphasize that virus capsids could be assembled using modest association energies, that capsid closure was an important driving force for the reaction, and that assembly by means of a series of bimolecular reactions can be fast and efficient.

In this paper, we develop a more general model of capsid assembly, the kinetically limiting, or KL, model, and apply it to experimental observation of HBV capsid assembly. The KL model is derived from the EQ model except that it allows high association energy and nonuniform steps in the assembly process, in particular a rate-limiting "nucleation" step. This simple modification results in a much more robust model and better understanding of the properties of macromolecular assembly. Analyses based on the model allow interpretation of HBV assembly kinetics.

EXPERIMENTAL PROCEDURES

Assembly of HBV Capsids. The 149-residue assembly domain of the HBV capsid protein (Cp149) was expressed and purified as described previously (*9*, *15*). Stock solutions of Cp149 dimers [1−2 mg/mL in 0.1 M sodium bicarbonate (pH 9.5) and 5 mM DTT] were stored at −70 °C. Protein concentrations were determined using an ϵ_{280} of 60 900 M^{-1} cm^{-1}.

Light scattering experiments were performed using a SPEX Fluoromax 2 fluorometer (Instruments SA, Inc., Edison, NJ) with excitation and emission monochromators set to 320 nm with a 3 nm bandpass, using a black masked micro fluorescence cuvette (Hellma, Jamaica, NY). At this wavelength, there is no protein absorbance of incident or scattered light. Light scattering was recorded at 20 °C for 1 s integration times, usually over the course of 600 s. In a typical experiment, light scattering of protein diluted with storage buffer (60 μL total volume) was measured for 40 s. Assembly was initiated at 50 s when the sample was diluted with an equal volume of NaCl in 0.1 M sodium bicarbonate (pH 9.5) to the desired final NaCl concentration. The preassembly scattering measurement was used to generate a standard curve for calculating protein concentration and for normalizing the relative light scattering. Most reactions had a final volume of 100 or 120 μL, but identical results were obtained with 0.5 mL volumes.

Chromatography experiments were performed at 20 °C using a SMART-FPLC system and a 0.6 cm × 10 cm column (2.5 mL) packed with Sephacryl S-300 HR (all from Amersham-Pharmacia Biotech, Piscataway, NJ). Samples were mixed as described for the light scattering experiments. At 300 s after mixing, 50 μL of the sample was loaded onto the column and eluted at a rate of 90 μL/min with 0.1 M bicarbonate and the same NaCl concentration used for assembly. The elution profiles of chromatographs were recorded at 280 nm. Integrated peaks were quantified on the basis of absorbance at 280 nm.

Relative proportions of $T = 4$ and $T = 3$ capsids were determined by sedimentation through a 5 to 30% sucrose gradient (0.1 M sodium bicarbonate and 0.4 M NaCl) prepared with a Gradient Master (Biocomp, Fredericton, NB) (*21*). Samples with final volumes of 0.3−0.5 mL and final protein concentrations of 0.2−1 mg/mL were assembled as described above, by mixing protein in storage buffer with a buffered salt solution (final NaCl concentration of >0.6 M).

The lower salt concentration of the gradient prevented further assembly.

Simulations. Simulations of capsid assembly are based on the formation of a dodecahedron from 12 pentagonal subunits (Figure 1A). The association constant for the addition of one subunit to an intermediate with n subunits ($K_a n$) was calculated from the microscopic association constant ($K_a$contact) for edge−edge contacts, the number of contacts (j), and a statistical term (Sn) dependent on the degeneracy of the reactants and products (eq 1) (*20*).

$$K_a n = K_a\text{contact}^j Sn \text{ or } \Delta Gn = j \times \Delta G\text{contact} + RT \ln Sn \quad (1)$$

Kinetic simulations were based on the addition of one subunit at a time to the growing polymer. To simplify the calculations, all reactions were given the same forward rate constant, unless otherwise stated. The back reaction rate constants, k_b, were calculated from the association constant, $K_a n$, and k_f ($k_b n = k_f/K_a n$). Both forward and back reactions are considered for all steps (eq 2).

$$\text{d[polymer}_n\text{]/d}t = k_f([\text{polymer}_{n-1}] - [\text{polymer}_n])[\text{free subunit}] - k_b[\text{polymer}_n] \quad (2)$$

Simulations for the formation of a dodecahedron were developed as a cascade of 11 reactions represented by 12 differential equations of the form of eq 2. Kinetically limiting (KL) assembly simulations incorporated microscopic constants for the association and forward rate for the nucleation step. Numerical integrations were calculated with a fourth-order Runge−Kutta approximation using STELLA (High Performance Systems, Inc., Hanover, NH). The relative light scattering was calculated as the weight average molecular weight and normalized to a maximum of 1.0, for the case where all subunits are in capsids.

We excluded the effect of association of oligomers in these simulations. The concentration of intermediates, and hence their likelihood of participating in a reaction, is low. The accumulation of intermediates in some reactions results in formation of kinetic traps (see the text). Many viruses assemble from preformed oligomers (*1*); in this formalism, such oligomers are treated as subunits.

The reactants and products shown in Figure 1A are the most stable and the most likely species. For clarity, a single reaction path is shown; however, the entropy term (*20*) takes into account the degeneracy of the reactions. The effect of alternative intermediates on assembly was investigated. Because the larger part of the reaction is unidirectional, these alternative intermediates have an inconsequential effect on kinetics and no effect on the final concentrations of species.

RESULTS

Comparison of Assembly Models: Equilibrium (EQ) Assembly Revisited. In EQ assembly (*20*), all steps in the assembly reaction (Figure 1A) have the same on-rate and all intersubunit contacts have the same relatively weak binding energy. An example of EQ assembly under optimal conditions is shown in Figure 1B. For an initial subunit concentration of 2 μM, a microscopic association energy $\Delta G°$ of −4.2 kcal mol^{-1} contact^{-1} (corresponding to a K_a of 10^3 M^{-1}), and a diffusion-limited k_f of 10^8 M^{-1} s^{-1}, >80% of the material is assembled into capsids in less than 300 ms. During an assembly reaction, appreciable (experimentally measurable) concentrations of intermediates accumulate and are consumed to produce capsids. The oligomers associate and dissociate, making and breaking the various intermediates and maintaining a supply of free subunits that allow the reaction to proceed; the EQ model is named after this equilibration. The diffusion-limited rates were chosen to illustrate how fast this reaction can proceed.

EQ assembly is susceptible to kinetic traps, a situation where assembly is initiated too many times, leaving an insufficient concentration of free subunits to allow the reaction to proceed to completion (Figure 1C) (*20*, *22*). Kinetic traps can result from a protein concentration that is too high and/or an association constant that is too strong. As the K_a is made stronger, the protein concentration must be proportionally decreased. In Figure 1C, the kinetic trap results from increasing the K_a by 1 order of magnitude to 10^4 M^{-1} contact^{-1}. After 300 ms, most subunits are found as oligomers with four to nine subunits. The concentration of dodecahedra increases, but very slowly (50 nM after 2 s). Assembly is not necessarily prevented by a kinetic trap, but it is impeded. For EQ assembly with a given K_d, there is an optimal subunit concentration range that will yield the fastest capsid formation. Both higher and lower concentrations will result in slower reactions.

EQ assembly is a viable mechanism for capsid assembly, but constraints must be used to avoid kinetic traps. A low subunit−subunit association energy will yield fragile capsids. Fragile provirions or procapsids, which mature to a more robust form, are found in many virus families, including picornaviridae (*23*), nodaviridae (*24*), and herpesviridae (*25*). In these examples, maturation is triggered by a proteolytic modification of the capsid protein.

Kinetically Limited (KL) Assembly. Kinetic traps resulting from depletion of free subunits can be avoided by limiting the association energy and/or subunit concentration. Another strategy for avoiding traps is to prevent overinitiating capsid assembly, i.e., incorporate a kinetically limiting step into the assembly reaction. This can be accomplished by postulating that early steps in the reaction are relatively slow and/or weak and that assembly must proceed through these early intermediates (Figure 1A). The kinetically limiting step may require addition of an exogenous nucleus such as a specific RNA (*26*) or an RNA−protein complex (*27*). The case where the nucleus may be a specific oligomer of capsid protein will be discussed in depth in this paper.

KL assembly simulations have been calculated with nuclei made of trimers and tetramers of the pentagonal subunits (Figures 1, 2, and 4). Nucleation reactions can be separated from elongation reactions by two physical distinctions. The steps leading to the nucleation of assembly can be slower than the subsequent elongation steps (k_f-limiting), or the intermediates can be less stable. The effect of either of these mechanisms is nearly indistinguishable, so only the k_f-limited mechanism will be discussed in detail.

KL assembly (Figure 1D) has many of the same features observed in EQ assembly, a lag phase followed by rapid growth. This reaction is somewhat slower than EQ assembly under optimal conditions, but is extremely robust with respect to protein concentration (Figure 4D). Kinetic traps do occur but only in the most extreme conditions (Figure 4B). In a

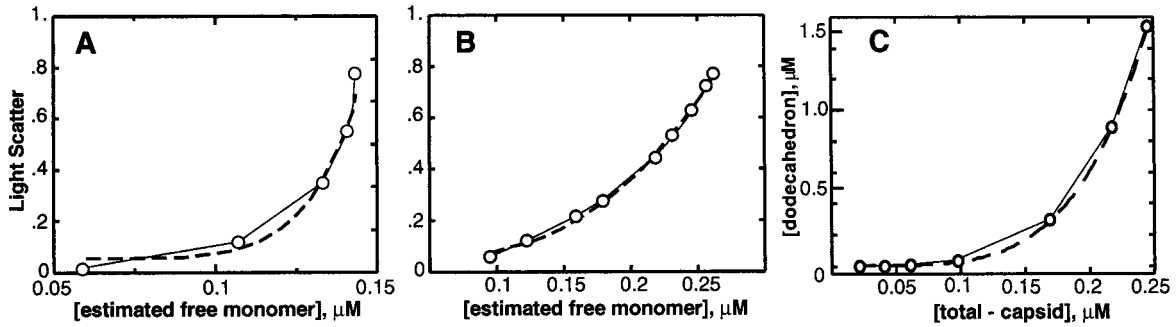

FIGURE 2: Concentration dependence of the extent of EQ and KL assembly. (A) The extent of EQ assembly estimated from light scatter is best fit to a power of 9.4, close to the size of the capsid. The extents of KL assembly (B) with a trimeric nucleus and (C) with a tetrameric nucleus are best fit with concentration dependences of 3.0 and 3.9, respectively. For KL with a tetrameric nucleus, calculated concentrations of species were used for the plot. For the other two curves (A and B), the concentrations were estimated by light scattering using the equation [free subunit] = [initial subunit](1 − relative light scatter). The data are fit (dashed lines) to the equation [capsid] = $(c_1[\text{free subunit}])^{c_2}$, where c_1 is a proportionality and c_2 is the power dependence. Correlation coefficients for these curves are all >0.998. The same power laws may be derived from a log−log plot.

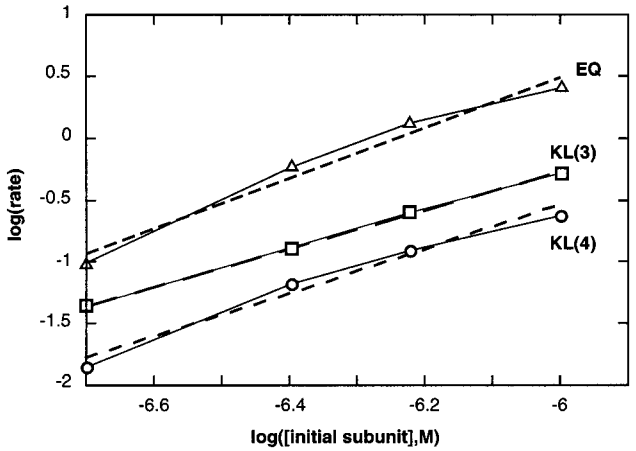

FIGURE 3: Concentration dependence of the rate of assembly. Rates were extracted from a linear fit to the fast phase of assembly simulations for families of curves of calculated light scatter. The resulting power laws were 2.0 for the EQ model (△), 1.6 for the KL assembly with a trimeric nucleus (□), and 1.8 for the KL assembly with a tetrameric nucleus (○).

productive reaction, dimers assemble slowly but accumulate (ca. 10^{-7} M) because they are consumed slowly by trimer formation. On the other hand, trimers and later intermediates react rapidly. Except at the initiation of the reaction, the concentrations of trimer through undecamer approximate a steady state of $\sim 10^{-11}$ M (in Figure 1D) until free subunit concentrations become depleted. In this simulation, the net on-rates for dimerization and trimerization are 10^6 M^{-1} s^{-1}, 2 orders of magnitude slower than the nine elongation reactions, but still very fast. This slower rate may be attributed to the greater conformational flexibility of the early reactants.

The classical explanation for multiphase kinetics is in terms of nucleated assembly, where the lag phase is related to the kinetics of nucleus formation (panels B and D of Figures 1 and 4). From simulations of KL assembly, we arrive at a different explanation. The lag phase is the time required to generate a steady state concentration of intermediates. Once intermediates achieve steady state, they have nearly identical concentrations which are held nearly constant through most of the reaction. This steady state pipeline can be thought of as a capsid assembly line. At the earliest time points, the concentrations of intermediates are nearly 0 M; the assembly rate becomes maximal once the steady state "pipeline" is established.

Table 1: Comparison of EQ and KL Assembly

	EQ assembly	KL assembly
nucleation	not required	necessary
kinetic traps	susceptible	resistant
association energy	low	not constrained
capsid stability	relatively unstable, provirion-like	not constrained
concentration dependence		
rate	reaction order	reaction order
extent	capsid size	nucleus size
thermodynamic editing	yes	no
sensitive to defective or modified subunits	no	yes

KL assembly is a robust model of capsid assembly. The slow rate of formation of reactive intermediates due to the impeded initial step has many consequences. Kinetic traps can be avoided even when the association energy of the elongation steps is very great (more than -14 kcal mol^{-1} $contact^{-1}$). The kinetics of assembly and susceptibility to kinetic traps are largely a function of the limiting step. This implies that the formation and conformation of the nucleus are the directing events of KL assembly.

Another effect of KL assembly is its potential for retarding capsid assembly. In vivo, capsids may assemble around nucleic acid and/or specific protein(s). In KL assembly, the concentration of capsid protein can accumulate far beyond the level required for efficient elongation of a pre-existing nucleus. The addition of an exogenous nucleus or nucleating complex to a supersaturated solution will result in rapid assembly of as many capsids as there are nuclei.

Identification and Comparison of Assembly Models. Under a reasonable range of conditions, the predicted behavior for EQ and KL models is similar (Figure 1). There are features that allow determination of the appropriate assembly model (see Table 1). (i) EQ assembly requires a low intersubunit association energy. (ii) EQ assembly is particularly sensitive to generating kinetic traps. (iii) The concentration dependence of the extent of assembly identifies the size of the limiting species in the assembly process in EQ assembly, or nucleus for the KL mechanism.

The concentration dependence of the extent of assembly can be extracted by plotting the concentration of capsid

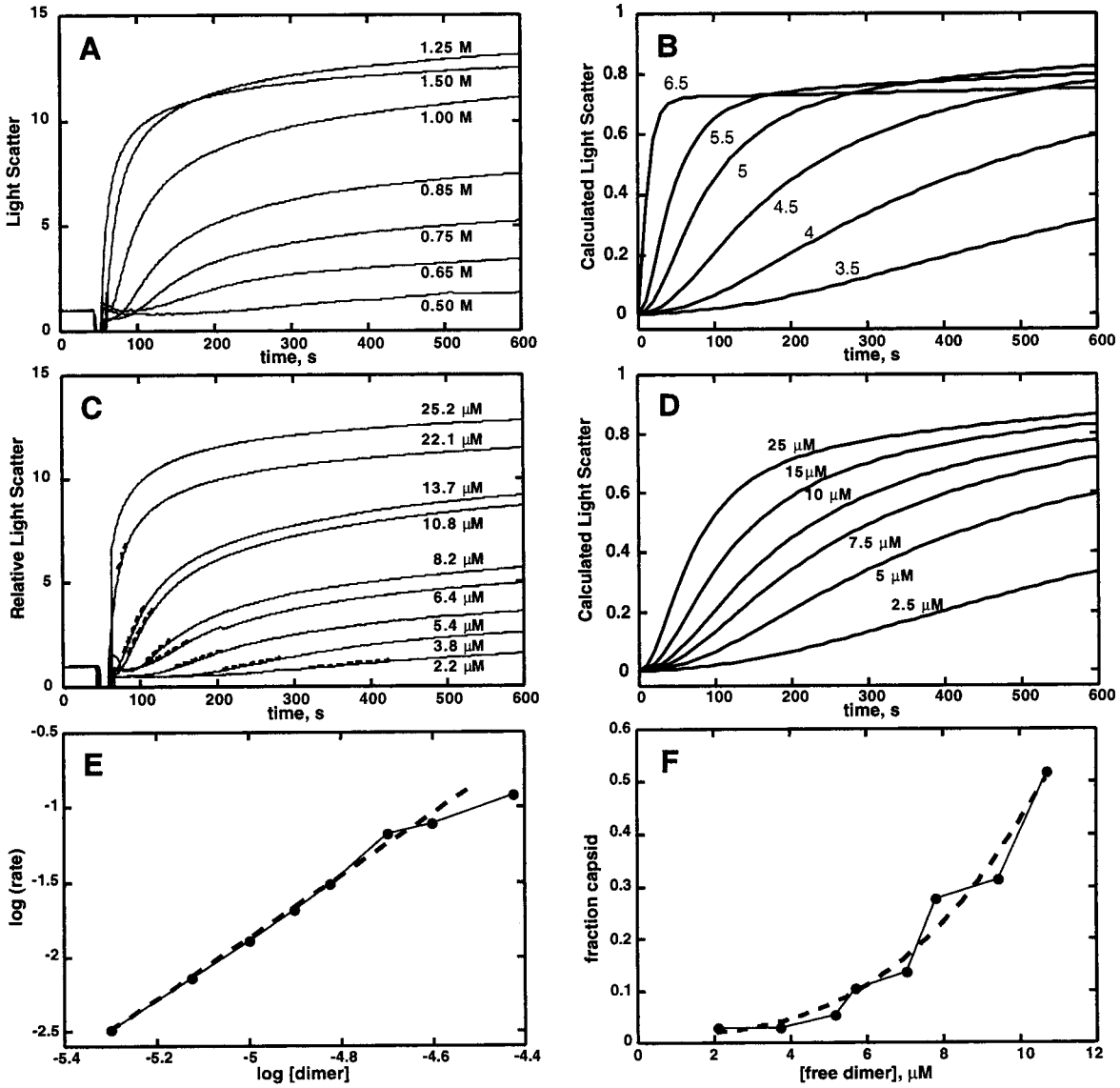

FIGURE 4: Comparison of HBV assembly with simulations of capsid assembly. (A) Different salt concentrations were used to induce the assembly of 5 μM HBV. (B) Simulation of capsid assembly, based on a dodecahedral model with a trimeric nucleus, with varying association energies (in kilocalories per mole per contact). The simulation predicts kinetic trap formation at high association energy with a resulting attenuation of the light scattering. (C) Assembly of varying HBV concentrations by 0.6 M NaCl. The rates of the fast phase were estimated from the dashed lines. These data were used in panels E and F. (D) Simulation of the concentration dependence of assembly using an association energy of $10^{4.5}$ M^{-1}, a nucleation forward rate of $10^{2.5}$, and an elongation rate of $10^{4.5}$ M^{-1} s^{-1}. (E) The rate of assembly exhibits a 2.1-power concentration dependence (correlation coefficient = 0.997). (F) The concentration dependence of the extent of assembly exhibits a 2.9-power concentration dependence (correlation coefficient = 0.997; $\chi^2 = 0.007$). For the assembly reactions, HBV was in 0.1 M sodium bicarbonate (pH 9.5); assembly was induced by the addition of buffered NaCl at 50 s. Changes in light scattering were observed at 320 nm using a fluorometer.

against the free subunit concentration (Figure 2). For the results from these plots to be interpretable, the reaction must be in the "pipeline phase", not too near completion, and not impeded by a kinetic trap. For EQ assembly, at low initial subunit concentrations, assembly exhibits a 9.4-power dependence (Figure 2A). At higher initial concentrations, more material is kinetically trapped. For KL assembly with a trimeric nucleus, the curve exhibits a 3.0-power dependence (Figure 2B). For KL assembly with a tetrameric nucleus, the concentration dependence has a power of 3.9 (Figure 2C). This concentration dependence relationship breaks down as a reaction approaches completion.

This behavior is an effect of the steady state pipeline where there are few nuclei and intermediates. Though nucleus formation may actually be a cascade of low order reactions, its rate of formation can be approximated as a single high order reaction, where n is the number of subunits in the nucleus and k_{fn} is the net rate constant for nucleus assembly (eq 3).

$$\mathrm{d[nucleus]}/\mathrm{d}t \approx k_{fn}[\mathrm{subunit}]^n \qquad (3)$$

Because each nucleus entering the pipeline leads to a capsid, when the amount of capsid is observed at a time point much later than the time required to form a single particle, the amount of capsid observed is proportional to the amount of nuclei formed. On the basis of eq 3, this relationship can be expressed as a definite integral between t_0 and some

time t within the pipeline phase of the reaction, where k'_{fn} is a constant related to the rate of nucleation, k_{fn}, and subsequent elongation (eq 4). This is demonstrated in plots of capsid concentration versus free subunit concentration (Figure 2). This approximation also holds for EQ assembly at low protein concentrations, where there are relatively low concentrations of intermediates (Figure 2A).

$$[\text{capsid}] \approx \int_{t_0}^{t} k'_{fn}[\text{subunit}]^n \, dt \qquad (4)$$

The concentration dependence of the estimated rate of reaction reflects the order of the elongation reaction. The rates of assembly curves were estimated as the slope of the light scattering, following the lag phase. The concentration dependence of these rates was measured in a log−log plot of rate versus the initial subunit concentration (*28*). For EQ, KL(trimer), and KL(tetramer) models, the concentration dependence estimated from light scattering was 2.1, 1.6, and 1.8, respectively (Figure 3). A higher-order concentration dependence for the rate would reflect either higher-order reactions or the cooperative binding of several subunits at a time.

HBV Assembly Is Salt- and Concentration-Dependent. HBV assembly at pH 9.5 was observed using light scattering. We observe that the kinetics of high-pH assembly are slow and require at least 0.5 M NaCl, even at dimer concentrations of 1 mg/mL (30 μM).

In Figure 4A, we show a family of curves in which assembly of 5 μM capsid protein dimer, induced by different NaCl concentrations, was observed by light scattering. Kinetics were sigmoidal and can be divided into lag, fast, and asymptotic phases. As the concentration of sodium chloride used to induce assembly is increased, the lag phase becomes shorter and the slope of the fast phase becomes steeper. In general, the maximum light scatter observed during the course of an experiment also increased as the ionic strength is increased. The exception was observed in assembly at the highest NaCl concentration, 1500 mM, where the slope of the asymptotic phase is almost horizontal and the magnitude of the light scattering increase is reproducibly smaller than that of the signal observed with slightly lower (1250 mM) salt concentrations.

HBV capsid assembly is also protein concentration-dependent. In the family of curves shown in Figure 4C, assembly was initiated by adding NaCl to a final concentration of 600 mM and the protein concentration was varied. The light scatter is normalized for protein concentration. As with the previously described experiments, at higher protein concentrations the lag phase becomes shorter and the fast phase becomes steeper. The extent of the reaction is also progressively greater, even at the highest protein concentrations that were tested.

Simulations for comparison to HBV assembly were calculated for the formation of a dodecahedron whose assembly is initiated by formation of a trimeric nucleus followed by a cascade of second-order reactions, where the forward rate of nucleation is $^1/_{100}$ of the rate of the subsequent elongation reactions. The great difference between assembly of HBV (a 120-mer) and the model (a 12-mer) precludes direct comparison. However, the model provides a means for understanding the HBV reaction.

The family of simulations that resemble the observed HBV NaCl dependence were generated by varying the energy of association between subunits (Figure 4B). This change in energy is evenly divided between increasing the k_f and increasing the microscopic K_a. At the highest association energy, the simulation resembles the assembly of 10 μM HBV by 1500 mM NaCl, where the initial rate is very steep but the maximum signal is attenuated and the slope of the final phase of the reaction is nearly horizontal. In the simulation, we find that the source of this behavior can be attributed to formation of a kinetic trap.

The family of simulations that mirror the HBV concentration dependence were calculated using a single association constant ($10^{4.5}$ M^{-1} $contact^{-1}$) and forward rate constants for nucleation reactions ($10^{2.5}$ M^{-1} s^{-1}) and elongation reactions ($10^{4.5}$ M^{-1} s^{-1}) (Figure 4B,D). The simulations are qualitatively similar to the observed kinetics and, most importantly, reproduce the concentration dependence of the reaction rate and the extent of reaction (Figure 4E,F).

The extent of HBV capsid assembly exhibits a third-order concentration dependence (Figure 4E). A third-order concentration dependence in the extent of assembly is also seen in plots derived from assembly simulations where there was a trimeric nucleus (Figure 2B). On the basis of the approximation in eq 4, this result suggests that HBV assembly, like the simulation, is initiated by the formation of a trimeric nucleus. A classical means of determining nucleus size is by determining the concentration dependence of the lag phase (*29*, *30*). In capsid assembly, the steady state concentration of intermediates is a governing feature of the overall reaction. The lag phase is mainly a function of the time required for generating the pipeline of intermediates. Formation of capsids is dependent on the entry of protein into the passage. Therefore, because the flux of nascent capsids through the pipeline is dependent on nucleation, the nucleus size is reflected by the concentration dependence of the extent of assembly.

The rate of HBV capsid assembly exhibits a second-order concentration dependence, which is generally diagnostic for a second-order reaction. The rate of assembly was estimated on the basis of tangent lines to the steepest slope of the observed assembly kinetics (note the dashed line in Figure 4C). Using these data derived from Figure 4C, a log−log plot of the rate of light scatter increase versus the initial protein concentration has a slope of 1.9 (Figure 4F). An equivalent plot derived from the simulation in Figure 4D (see Figure 1D), where the elongation reactions are known to be second-order, has a slope of 1.6 (Figure 3). In simple association reactions, one expects that the order of the reaction will be reflected in the concentration dependence of the rate; this generalization also appears to hold true as an approximation for cascades of association reactions. The concentration dependence of the rate suggests that assembly of HBV capsids proceeds by a series of second-order reactions; i.e., subunits are added to the growing polymer one at a time. As with the estimation of nucleus size, the basis of this interpretation lies in the steady state concentration of intermediates in the pipeline. The steady state concentration is a function of the initial protein concentration. However, for HBV and the simulations, the association reactions of intermediates in the steady state pipeline are all second-order with respect to intermediate and free subunit

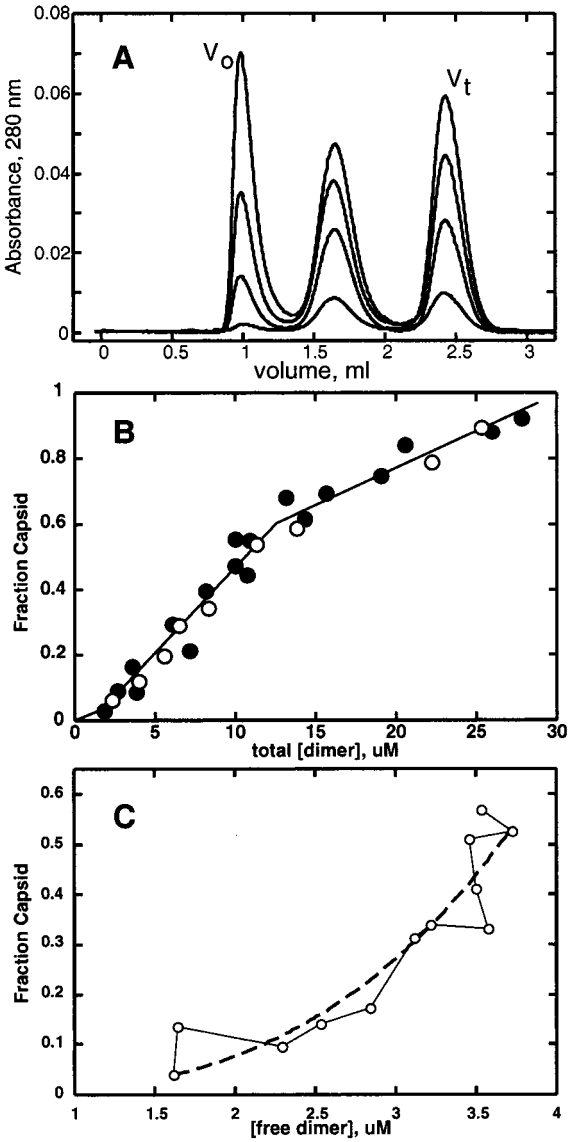

FIGURE 5: Determination of the extent of capsid assembly with size exclusion chromatography. (A) Sample chromatographs of reactions with 2, 5, 8, and 13 μM total dimer. Capsid elutes in the void volume (V_o) at 1 mL; dimer elutes at 1.7 mL, and DTT from storage buffer elutes at the total volume (V_t) of 2.5 mL. (B) Comparison of chromatographic (●) and light scattering (○) estimates of assembly shows excellent agreement. (C) The extent of assembly exhibits a 3.1-power concentration dependence (correlation coefficient = 0.92, $\chi^2 = 0.054$).

concentrations. A second-power concentration dependence does not necessarily follow from this treatment of capsid assembly data; in a similar plot, a fifth-power concentration dependence was observed with bacteriophage P22 (*28*).

Similar results for the extent of assembly can be derived from liquid chromatography (LC) studies of HBV assembly (Figure 5). Where the light scattering experiments identify relative changes in the average molecular weight of solute, size exclusion LC provides information about the molecular weight range of species and, using an absorbance detector, quantifies the amount of material in each peak. As with the light scattering experiments, assembly was induced by 0.6 M NaCl. After 5 min, samples were resolved using a short Sephacryl S300 column. According to the manufacturer, the S-300 resin has a useful fractionation range of about 10^4−10^6 Da (HBV dimers and capsids have masses of 3.4×10^4 and 4×10^6 Da, respectively). Two peaks were observed on typical column chromatograms monitored simultaneously at 280 and 320 nm, appearing at the void volume and midway through the included volume (Figure 5A). A peak that included residual DTT from the capsid protein storage buffer was seen at the end of some chromatograms. The elution of the first peak in the void volume and its light scattering properties are consistent with identifying it as capsids or nearly complete capsids. The second peak appears midway through the column and elutes at the same volume as dissociated capsid protein in low-ionic strength buffer (not shown), identifying it as free dimer. In longer columns, we achieve baseline separation of capsid and dimer, suggesting that, at most, very low concentrations of stable intermediates are present.

The extent of assembly determined by LC correlates well with the light scattering results. Using an estimated 32-fold difference in scattering as the maximum possible scattering signal, we normalized LC and light scattering data to the mole fraction capsid (Figure 5B). The extent of assembly, as determined by LC, exhibits a third-power concentration dependence (Figure 5C). This is not surprising considering the agreement observed between the extent of assembly in light scattering and LC results. However, the two methods are completely independent of one another.

Assembly products were also analyzed to show that they were capsids and not amorphous aggregates. Cp149 assembled at neutral pH is typically >90% $T = 4$ capsids as determined by analytical ultracentrifugation, sucrose gradient sedimentation, and cryoelectron microscopy (*15*). Using sucrose gradient sedimentation, we found that samples assembled at pH 9.5 had high yields of capsid but somewhat lower proportions of $T = 4$ capsids (ca. 80%). LC of assembly reactions through Sepharose CL-4B resolves capsids from the void volume. Only capsids and dimers were observed on this column (dimers were poorly resolved from V_t). There was no evidence of aggregation in fresh assembly reactions. Aggregates can be observed in the void volume of Sepharose CL-4B columns after the sample has been stored several days.

DISCUSSION

We have examined HBV assembly at pH 9.5, where association is much weaker than at pH 7.5 (*9*, *31*). We find that the extent of assembly exhibits a third-power concentration dependence, indicating a trimeric nucleus. The rate of assembly exhibits a second-power concentration dependence, indicating elongation proceeds via second-order reactions. On the basis of these features and considering assembly-related details of the capsid structure (*14*, *15*, *32*−*34*), we suggest that assembly of empty HBV capsids begins with formation of a triangular nucleus (Figure 6). This is the smallest possible closed structure that can be formed during capsid assembly. Once the rate-limiting nucleus has formed, subsequent dimers add to the growing polymer one at a time until it is complete. This suggests that HBV is an example of KL assembly because (i) assembly is nucleated, (ii) assembled particles are stable to thermal denaturation and do not readily disassemble even at pH 9.5 (*9*, *15*, *34*), and (iii) at pH 7.5 the association energy will be too strong to be consistent with an EQ assembly mechanism.

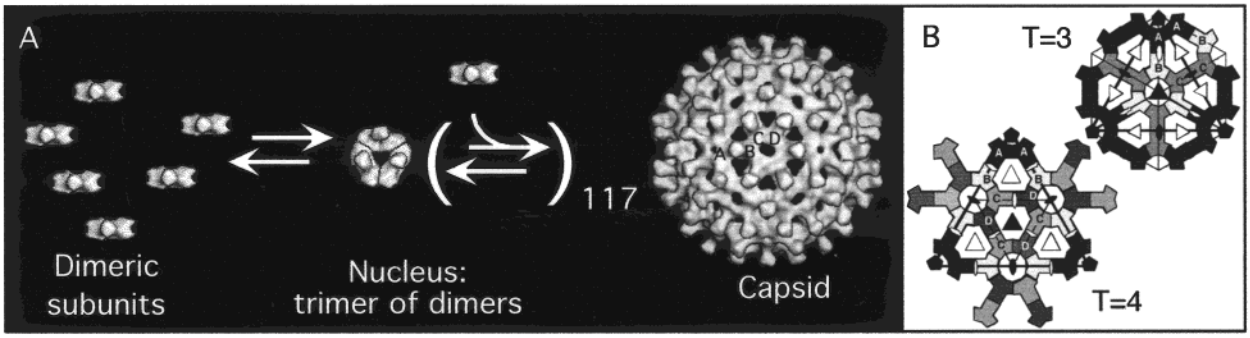

FIGURE 6: Schematic of HBV capsid assembly. (A) Capsid assembly begins with formation of a trimeric nucleus; subsequent elongation reactions add one dimeric subunit at a time. The product of the reaction shown here is a $T = 4$ capsid at 17 Å resolution (*15*). (B) Schematic representation of $T = 3$ and $T = 4$ icosahedral facets, emphasizing the dimeric structure of the subunit. Putative nuclei correspond to $T = 3\{AB-AB-CC\}$, $T = 4\{AB-AB-CD\}$, and $T = 4\{3CD\}$. Icosahedral symmetry operators are shown as solid pentagons (5-fold axes), triangles (3-folds), and ovals (2-folds). Quasi-symmetry axes are shown as white symbols.

This study provides evidence that assembly of empty HBV capsids is regulated by nucleus formation, which allows a supersaturated solution of capsid protein to accumulate before capsid assembly begins. Thus, assembly can be initiated by an exogenous nucleus before the homopolymeric nucleus is formed. In vivo assembly of HBV may benefit from a lag between the availability of sufficient capsid protein and the initiation of assembly.

The assembly path proposed in this paper also has implications for the morphological switch between the $T = 3$ and $T = 4$ capsid forms of HBV. Previously, we proposed that switching was regulated on a kinetic basis (*15*). The C-termini of the capsid protein, an important determinant of capsid morphology, are localized to 5-fold and quasi-6-fold vertices (*34*). These vertices are also the vertices of the proposed triangular nucleus. The major difference in the quaternary structure of $T = 3$ and $T = 4$ capsids is localized to the quasi-6-fold vertices; the 5-fold vertices, comprised of AB subunits, are identical (*15*). In a nascent $T = 4$ capsid, there are two possible nuclei, which are constructed of either two AB dimers and a CD dimer or three CD dimers (see Figure 6). For a nascent $T = 3$ capsid, only one nucleus is possible, which is composed of two AB dimers and one CC dimer. The difference between the 2(AB)+CC and the 3(CD) conformations may be the true conformational switch. When first formed, these two conformations may be interconvertable. Once assembly begins, this switch becomes locked into a single conformation by the addition of other subunits.

We have considered two general models for the assembly of a protein polymer that are based on (i) an equilibrium association reaction or (ii) a unidirectional reaction with a kinetically limiting early step. The models make testable predictions for the behavior of assembly reactions, e.g., concentration dependence, association energy, and susceptibility to kinetic traps. Both models are compatible with other levels of regulation such as Caspar's autostery mechanism (*19*).

There are advantages and disadvantages to both model assembly mechanisms (Table 1). KL assembly is resistant to kinetic traps and can yield a robust capsid. EQ assembly, on the other hand, necessarily yields a relatively fragile, provirion-like capsid. However, if defective subunits are incorporated (or misincorporated) into a growing polymer during EQ assembly, they will be displaced by equilibration with competent subunits to achieve the global energy minimum. In KL assembly, there is no such "thermodynamic editing". For this reason, we suggest that capsid assembly is a viable target for antiviral intervention only in viruses that assemble by a KL mechanism.

Assembly of a virus capsid typically involves tens to hundreds of subunits. Though the complexity of the reactions may be daunting, simple models of the reactions and an understanding of their behavior provide guidelines for the experimental dissection of in vitro capsid assembly.

ACKNOWLEDGMENT

We thank Gillian Air and Stephen Stray for their many helpful comments. A.Z. dedicates this paper to the memories of his father, Martin Zlotnick, and first mentor, Dr. Caroline T. Holloway, both of whom taught the importance of inquiry and humanity.

REFERENCES

1. Chiu, W., Burnett, R. M., and Garcea, R. L. (1997) *Structural Biology of Viruses* p 484, Oxford University Press, New York.
2. Ganem, D. (1996) in *Fields Virology* (Fields, B. N., Knipe, D. M., Howley, P. M., Chanock, R. M., Melnick, J. L., Monath, T. P., Roizman, B., and Straus, S. E., Eds.) pp 2703−2737, Lippincott-Raven Publishers, Philadelphia, PA.
3. Gerelsaikhan, T., Tavis, J. E., and Bruss, V. (1996) *J. Virol. 70*, 4269−4274.
4. Lamberts, C., Nassal, M., Velhagen, I., Zentgraf, H., and Schroder, C. H. (1993) *J. Virol. 67*, 3756−3762.
5. Guidotti, L. G., Matzke, B., Pasquinelli, C., Schoenberger, J. M., Rogler, C. E., and Chisari, F. V. (1996) *J. Virol. 70*, 7056−7061.
6. Stahl, S., MacKay, P., Magazin, M., Bruce, S. A., and Murray, K. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 1606−1610.
7. Cohen, B. J., and Richmond, J. E. (1982) *Nature 296*, 677−679.
8. Zhou, S., and Standring, D. N. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 10046−10050.
9. Wingfield, P. T., Stahl, S. J., Williams, R. W., and Steven, A. C. (1995) *Biochemistry 34*, 4919−4932.
10. Seifer, M., Zhou, S., and Standring, D. N. (1993) *J. Virol. 67*, 249−257.
11. Kenney, J. M., von Bonsdorff, C. H., Nassal, M., and Fuller, S. D. (1995) *Structure 3*, 1009−1019.
12. Caspar, D. L. D., and Klug, A. (1962) in *Cold Spring Harbor Symp. Quant. Biol. 27*, 1−24.
13. Stannard, L. M., and Hodgkiss, M. (1979) *J. Gen. Virol. 45*, 509−514.
14. Crowther, R. A., Kiselev, N. A., Bottcher, B., Berriman, J. A., Borisova, G. P., Ose, V., and Pumpens, P. (1994) *Cell 77*, 943−950.

15. Zlotnick, A., Cheng, N., Conway, J. F., Booy, F. P., Steven, A. C., Stahl, S. J., and Wingfield, P. T. (1996) *Biochemistry 35*, 7412−7421.
16. Horton, N., and Lewis, M. (1992) *Protein Sci. 1*, 169−181.
17. Berger, B., Shor, P. W., Tucker-Kellogg, L., and King, J. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 7732−7736.
18. Reddy, V. S., Giesing, H. A., Morton, R. T., Kumar, A., Post, C. B., Brooks, C. L., III, and Johnson, J. E. (1998) *Biophys. J. 74*, 546−558.
19. Caspar, D. L. (1980) *Biophys. J. 32*, 103−138.
20. Zlotnick, A. (1994) *J. Mol. Biol. 241*, 59−67.
21. Zlotnick, A., Palmer, I., Stahl, S. J., Steven, A. C., and Wingfield, P. T. (1999) *Acta Crystallogr., Sect. D*, 717−720.
22. Schwartz, R., Shor, P. W., Prevelige, P. E. J., and Berger, B. (1998) *Biophys. J. 75*, 2626−2636.
23. Rueckert, R. R. (1996) in *Fields Virology* (Fields, B. N., Knipe, D. M., Howley, P. M., Chanock, R. M., Melnick, J. L., Monath, T. P., Roizman, B., and Straus, S. E., Eds.) pp 609−654, Lippincott-Raven, Philadelphia, PA.
24. Gallagher, T. M., and Rueckert, R. R. (1988) *J. Virol. 62*, 3399−3406.
25. Newcomb, W. W., Homa, F. L., Thomsen, D. R., Booy, F. P., Trus, B. L., Steven, A. C., Spencer, J. V., and Brown, J. C. (1996) *J. Mol. Biol. 263*, 432−446.
26. Witherell, G. W., Wu, H. N., and Uhlenbeck, O. C. (1990) *Biochemistry 29*, 11051−11057.
27. Bartenschlager, R., and Schaller, H. (1992) *EMBO J. 11*, 3413−3420.
28. Prevelige, P. E. J., Thomas, D., and King, J. (1993) *Biophys. J. 64*, 824−835.
29. Hofrichter, J., Ross, P. D., and Eaton, W. A. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 3035−3039.
30. Eaton, W. A., and Hofrichter, J. (1995) *Science 268*, 1142−1143.
31. Zlotnick, A., Stahl, S. J., Wingfield, P. T., Conway, J. F., Cheng, N., and Steven, A. C. (1998) *FEBS Lett. 431*, 301−304.
32. Bottcher, B., Wynne, S. A., and Crowther, R. A. (1997) *Nature 386*, 88−91.
33. Conway, J. F., Cheng, N., Zlotnick, A., Wingfield, P. T., Stahl, S. J., and Steven, A. C. (1997) *Nature 386*, 91−94.
34. Zlotnick, A., Cheng, N., Stahl, S. J., Conway, J. F., Steven, A. C., and Wingfield, P. T. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 9556−9561.

BI991611A