JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

Norovirus Capsid Proteins Self-Assemble through Biphasic Kinetics via Long-Lived Stave-like Intermediates

Guillaume Tresset,*[,†] Clémence Le Coeur,[‡] Jean-François Bryche,[†] Mouna Tatou,[†] Mehdi Zeghal,[†] Annie Charpilienne,[§] Didier Poncet,[§] Doru Constantin,[†] and Stéphane Bressanelli*[,§]

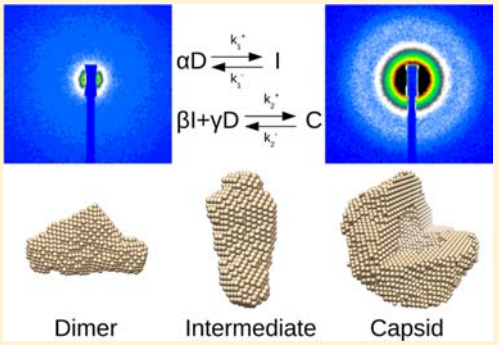
[†]Laboratoire de Physique des Solides, Université Paris-Sud, CNRS, 91405 Orsay, France

[‡]Institut de Chimie et des Matériaux Paris-Est, Université Paris-Est, CNRS, 94320 Thiais, France

[§]Laboratoire de Virologie Moléculaire et Structurale, CNRS, INRA, 91198 Gif-sur-Yvette, France

Supporting Information

ABSTRACT: The self-assembly kinetics for a norovirus capsid protein were probed by time-resolved small-angle X-ray scattering and then analyzed by singular value decomposition and global fitting. Only three species contribute to the total scattering intensities: dimers, intermediates comprising some 11 dimers, and icosahedral $T = 3$ capsids made up of 90 dimers. Three-dimensional reconstructions of the intermediate robustly show a stave-like shape consistent with an arrangement of two pentameric units connected by an interstitial dimer. Upon triggering of self-assembly, the biphasic kinetics consist of a fast step in which dimers are assembled into intermediates, followed by a slow step in which intermediates interlock into capsids. This simple kinetic model reproduces experimental data with an excellent agreement over 6 decades in time and with nanometer resolution. The extracted form factors are robust against changes in experimental conditions. These findings challenge and complement currently accepted models for the assembly of norovirus capsids.



INTRODUCTION

Noroviruses are small, nonenveloped, positive-strand RNA viruses. They are the first cause of nonbacterial sporadic gastroenteritis in humans and animals, thereby raising serious health and economic concerns worldwide.[1] The Norwalk virus (NV), the prototype norovirus targeting humans, has shown an average probability of infection for a single particle close to 0.5, exceeding the infectivity reported for any other virus studied to date.[2] Recombinant icosahedral $T = 3$ virus-like particles (VLPs) can be produced in the baculovirus system, and the crystal structure of NV-VLP is the only one reported so far for noroviruses.[3] NV-VLP is made up of 180 copies of a single structural protein, NV-VP1. Analyses of intermediates by native mass spectrometry along the disassembly pathway revealed a complex behavior of NV-VP1 as a function of solution pH, ionic strength, and protein concentration: albeit fully dissociated in dimers at pH 9.0 and weak salinity, NV-VLP remained intact at physiological pH and high salinity.[4] In the present work, we focus on a close relative of NV, the prototype genotype III/2 bovine norovirus strain Bo/Newbury2/1976/UK (NB2), the structural protein of which shares 50% sequence identity with NV-VP1. More importantly, the NB2-VP1 shell domain, which harbors the assembly determinants, holds 67% sequence identity with that of NV-VP1. As a consequence, NB2-VP1 self-assembles like NV-VP1 into $T = 3$ VLPs following a pseudo phase diagram parametrized by ionic conditions.[5] Although recombinant NB2-VP1 produced in the baculovirus system was shown to be suitable for antigenic and receptor-binding studies,[6,7] no crystal structure is available, and only a few biophysical studies have been carried out to date.

The precise and robust self-assembly of VLPs remains an elusive process, and the nature of the intermediate species is much debated. Yet the self-assembly of engineered VLPs has gained increasing attention for the production of biofunctionalized viral nanoparticles[8−10] and might help elucidate the principles of the *in vivo* replication of viruses. Although the number of possible intermediates grows combinatorially with the number of subunits, a graph-theory-inspired study performed on the assembly of a model icosahedron made of 30 dimeric subunits suggested a limited number of stable intermediates.[11] The authors enumerated all possible intermediates and estimated their probability of formation on the basis of the number of paths to them from their predecessors. Their simulations revealed very small sets of intermediates, notable for their compact form, that contribute substantially to assembly. From an experimental point of view, one intermediate at most has been identified to play a pivotal role in the assembly, and its nature varies with the viral system under consideration. For example, spherical NV-VLPs, as well as VLPs obtained from other viruses such as the cowpea chlorotic mottle virus, presumably self-assemble from dimers

Received: April 9, 2013
Published: July 3, 2013

into $T = 3$ capsids through pentamers of dimers.[3,12,13] For these capsids, the final assembly may be viewed as 12 pentamers of dimers connected by 30 interstitial dimers. On the other hand, hepatitis B VLPs are thought to rather proceed with trimers of dimers to complete either $T = 3$ or $T = 4$ capsids.[14] Additionally, there has been experimental evidence for the importance of protein conformational dynamics during assembly; more specifically, several studies emphasized the role of allostery in capsid construction by noticing that the protein adopts different conformations at different positions on the capsid lattice,[15,16] as well as the fact that the genome behaves as an allosteric effector to favor the completion of infectious virions.[17,18]

Self-assembly kinetics is often probed through the mass increase of assembling species, either by time-resolved static light scattering or by turbidity.[12,14,19,20] These techniques provide temporal traces from which characteristic time scales can be inferred, but no structural information can be extracted. In contrast, small-angle X-ray and neutron scattering (SAXS and SANS) techniques, thanks to the considerable progress realized on detectors, electronics, and efficient mixing devices, permit us now to determine the shape and the organization of particles with a spatial scale in the range of a few angstroms to thousands of angstroms, and with time resolution below 100 ms.[21] Recent kinetic studies on biological and soft matter systems by time-resolved small-angle scattering include protein[22] and RNA folding,[23] amyloid formation,[24] bacteriophage P22 procapsid assembly,[25] RNA encapsidation by virus-derived nanoparticles,[26] growth of gold nanoparticles,[27] formation of polymeric micelles,[28] and microemulsion phase transition[29] to name just a few. However, the coexistence of multiple species in solution renders data analysis difficult. In particular, it hinders the identification of the intermediates that can play a key role in the assembly process. Singular value decomposition (SVD) is a popular spectroscopic method which extracts the basis spectra contributing the most to the data set.[30,31] It allows an efficient data reduction, but it is difficult to assign a precise physical significance to the resulting basis spectra.

Here, we elucidate the nature of a long-lived intermediate involved in the self-assembly of NB2-VP1. The kinetics were probed by time-resolved small-angle X-ray scattering (TR-SAXS) with a high-brilliance synchrotron light source. The data showed contributions of only three long-lived species, namely the dissociated dimer, the final capsid, and an intermediate. The scattering pattern of the intermediate was extracted by global fitting of the data set with a three-state high-order kinetic model in such a way that the mass and volume of the intermediate were consistent with the stoichiometric coefficients. A low-resolution shape for the intermediate was numerically reconstructed *ab initio* and yielded an unexpected arrangement of proteins.

MATERIALS AND METHODS

Protein Expression and Purification. NB2-VP1 proteins were produced and purified as described elsewhere.[6] Briefly, monolayers of Sf9 cells were infected with recombinant baculovirus at a multiplicity of infection of 5 before being incubated in Hinks medium supplemented with 1% fetal calf serum for 5 days at 26 °C. After extraction of VLPs from freeze−thawed cells and culture medium by Freon X100, VLPs were purified by isopycnic centrifugation in cesium chloride gradients. Purified VLPs were stored at 4 °C in a slightly acidic buffer (20 mM MES pH 6.0 and 100 mM NaCl) and remained intact for several months. VLPs were dissociated into dimers by overnight dialysis at 4 °C against dissociation buffer (10 mM CHES pH 9.0). The formation of VLPs was triggered by rapidly mixing dimers in dissociation buffer with an equal volume of assembly buffer (100 mM MES pH 6.0 with 0 or 100 mM NaCl).

X-ray Scattering. TR-SAXS patterns were collected at the ID02 beamline of the European Synchrotron Radiation Facility (Grenoble, France). The sample temperature was maintained at 22 °C. Two-dimensional patterns were recorded using a high-sensitivity CCD detector (FReLoN). The sample-to-detector distance was set to 3.0 m, which provided wavenumbers q ($q = (4\pi/\lambda)\sin(\theta/2)$), where θ is the scattering angle and $\lambda = 0.996$ Å the X-ray wavelength) ranging from 3.3×10^{-3} to 0.18 Å^{-1}. A BioLogic SFM-400 stopped-flow apparatus (Grenoble, France) equipped with three independent syringes was used for mixing dimers with the assembly buffer. The mixed solution was then injected into a scattering cell consisting of a quartz capillary with a diameter of 1.8 mm, and the flow was stabilized with a solenoid hard stop. The delay from the mixer to the capillary was estimated to be about 4 ms. Beam exposure times were 20 ms for short time scales and 50 ms for long time scales. Intensities were converted into absolute units after subtraction of the contribution of buffer solutions. Data were often resampled on a logarithmic time scale for the fitting procedure so as to attribute a similar weight to data arising from short and long time scales.

Global Fitting. Collected TR-SAXS intensities $\mathbf{I}$ can be arranged in a matrix form:

$$\mathbf{I} = \begin{pmatrix} I(t_1, q_1) & \cdots & I(t_n, q_1) \\ \vdots & \vdots & \vdots \\ I(t_1, q_m) & \cdots & I(t_n, q_m) \end{pmatrix} \tag{1}$$

where t_n and q_m denote the sampled time and wavenumber, respectively. When several species coexist in solution with no long-range interactions between them, the total scattered intensity is the sum of the contributions arising from each species. If all the species have the same scattering properties, for example by being made up of copies of a single subunit, the total scattered intensity $\mathbf{I}$ can be factorized as follows:

$$\mathbf{I} = K(\bar{v}, \Delta b) \begin{pmatrix} M_1^2 P_1(q_1) & \cdots & M_L^2 P_L(q_1) \\ \vdots & \vdots & \vdots \\ M_1^2 P_1(q_m) & \cdots & M_L^2 P_L(q_m) \end{pmatrix} \begin{pmatrix} c_1(t_1) & \cdots & c_1(t_n) \\ \vdots & \vdots & \vdots \\ c_L(t_1) & \cdots & c_L(t_n) \end{pmatrix} = K\mathbf{P}'c$$

with $K(\bar{v},\Delta b)$ a known constant depending upon the specific volume $\bar{v}$ and the scattering length density Δb of the subunit, M_l, P_l, and c_l the molecular weight, form factor, and molar concentration, respectively, of species l, and L the number of species of interest. $P_l(q_m)$ reflects the molecular structure of species l, while $c_l(t_n)$ gives the variations of its concentration over time. The matrix of modified form factors $\mathbf{P}'$ contains the form factor of each species multiplied by the square of its molecular weight, so that $P_l'(0) = M_l^2$. The knowledge of $\mathbf{P}'$ gives access to both the structure and the molecular weight of the species contributing to the total scattered intensity. $\mathbf{c}$ is the matrix of concentrations; it reflects the kinetics of self-assembly and depends on the way subunits are exchanged between species over time. The modeling problem therefore consists of varying $\mathbf{P}'$ as well as some adjustable parameters $\{\xi_i\}$ of a model embedded into $\mathbf{c}(t,\{\xi_i\})$—typically, the reaction rates—to reproduce the variations over time of the measured intensities. In other words, we must determine sets of model parameter $\{\xi_i\}$ and a matrix $\mathbf{P}'$ of species patterns such that the relation $\mathbf{I} \approx \mathbf{I}^{\mathrm{fit}} = K\mathbf{P}'\mathbf{c}(\{\xi_i\})$ is optimal.[30] A straightforward least-squares interpretation requires that the residual,

$$\chi^2 = \frac{1}{MN - 1} \sum_{m,n=1}^{M,N} \left(\frac{I_{mn} - I_{mn}^{\mathrm{fit}}}{\sigma_{mn}} \right)^2$$

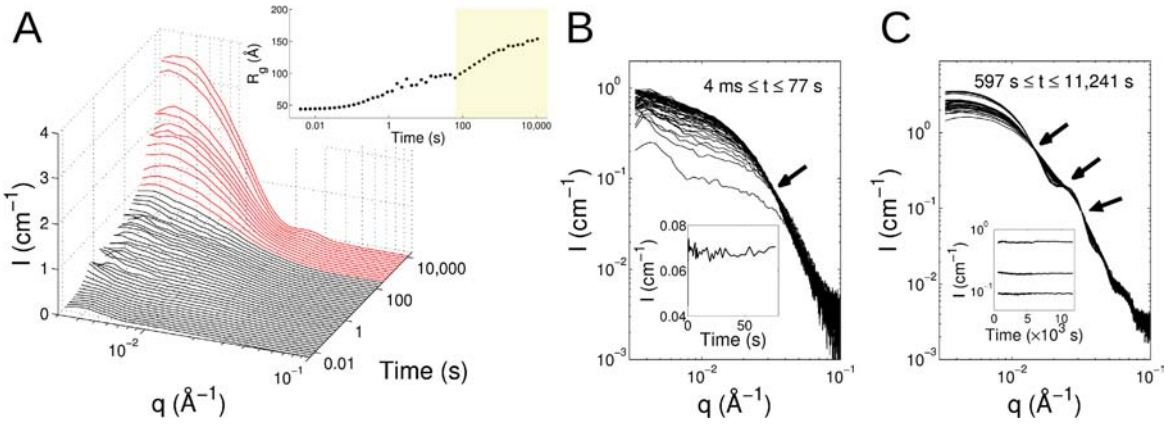


Figure 1. TR-SAXS patterns of self-assembling NB2-VP1. (A) Three-dimensional representation of the intensities on a logarithmic time scale from 4 ms to 11 721 s. The patterns in black and red correspond to short and long time scales respectively. The inset shows the radius of gyration R_g versus time. The yellow area emphasizes the second phase of these biphasic kinetics. (B,C)Superimposed patterns collected at early and late times, respectively. The intensities are plotted on a logarithmic scale for better clarity. The arrows indicate isosbestic points. Insets show the time evolution of intensities at the isosbestic points. Initial dimer concentration was 12 μM and ionic strength 55 mM.

be a minimum with respect to $\mathbf{P}'$ and model parameters $\{\xi_i\}$. M and N are the numbers of rows (q values) and columns (t values) of $\mathbf{I}$, σ the experimental uncertainties, and $\mathbf{I}^{\mathrm{fit}}$ the matrix of fitted intensities. In practice, χ^2 was computed over $q \leq 0.1$ Å^{-1}, where the signal-to-noise ratio was high. Provided that the parameters $\{\xi_i\}$ are nonlinearly related to $\mathbf{c}$, a first non-linear-least-squares algorithm running over $\{\xi_i\}$ is required to minimize χ^2. Within this loop, a second non-linear-least-squares algorithm is run over $\mathbf{P}'$, with the constraint that all its elements are strictly positive, i.e., $P_{ml}' > 0$.

The kinetic model is embedded as a set of reactions governed by reaction rates. The number of species L and the reactions connecting them are defined prior to minimization. The parameters to vary are the forward and backward reaction rates, $\{k_i^+\}$ and $\{k_i^-\}$, respectively. The kinetic model then translates into a set of coupled ordinary differential equations on the species concentrations:

$$\begin{cases} \dfrac{dc_1}{dt} = f_1(c_1, \cdots, c_L; \{k_i^+, k_i^-\}) \\ \vdots \quad \vdots \qquad\qquad \vdots \\ \dfrac{dc_L}{dt} = f_L(c_1, \cdots, c_L; \{k_i^+, k_i^-\}) \end{cases}$$

where $f_1,\ldots,f_L$ are multivariate functions depending on the kinetic model. The initial condition is always $(c_1,c_2,\ldots,c_L) = (c_0,0,\ldots,0)$, with c_0 being the initial concentration of subunits which must be conserved over time.

To assess the quality of the fitting over $q \leq 0.1$ Å^{-1}, we used χ^2 defined above and the R-factor, a common figure of merit in crystallography:[32]

$$R = \sum_{m,n=1}^{M,N} |I_{mn} - I_{mn}^{\mathrm{fit}}| \Big(\sum_{m,n=1}^{M,N} |I_{mn}| \Big)^{-1}$$

In some cases, the form factors of a few species may be known *a priori* or determined by separate measurements. Imposing these known data to the global fitting improves the figures of merit and provides the extracted data with a physical meaning. For implementing known form factors, we divide the L species of the kinetic model into sets of L_k known species with modified form factors $\mathbf{P_k}'$ and $L - L_k$ unknown species with modified form factors $\mathbf{P_u}'$. The associated $\mathbf{c}(\{\xi_i\})$ matrices are constructed, namely $\mathbf{c_k}(\{\xi_i\})$ and $\mathbf{c_u}(\{\xi_i\})$, which are the concentrations of known and unknown species, respectively. The equation to optimize for determining $\mathbf{P_u}'$ then becomes $\mathbf{I_u} = \mathbf{I} - K\mathbf{P_k}'\mathbf{c_k}(\{\xi_i\}) \approx K\mathbf{P_u}'\mathbf{c_u}(\{\xi_i\})$, which is carried out by least-squares minimization as before. The form factor of dimers was taken as the first scattering pattern of the kinetics ($t \approx 4$ ms in Figure 1). The modified form factor was obtained from a GNOM reconstruction and normalized in such a way that $P_D'(0) = M_D^2$. As for capsids, a SVD (see Supporting Information (SI)) was performed on the last patterns shown in Figure 1. Since the basis spectra given by SVD are purely mathematical quantities, rotation was used as a fast way to perform a linear combination yielding a meaningful pattern of capsids. It was carried out by multiplying a matrix whose columns were the basis spectra with a classical two-dimensional rotation matrix of angle to optimize. The two first basis spectra were thus rotated until sharp oscillations appeared, consistent with the form factor of a spherically symmetric capsid. We verified that the resulting form factor was close to the one obtained separately from a purified solution of VLPs.[5] Again, the modified form factor arose from a GNOM reconstruction normalized such that $P_C'(0) = M_C^2 = 90^2 M_D^2$.

Global Parameters and Shape Determination. For the selection of the most consistent model, we defined two global parameters, ρ_M and ρ_V, computed from the extracted modified form factor of the intermediate. Both of them provided the number of dimers making up the intermediate in terms of mass for ρ_M and volume for ρ_V. ρ_M was directly computed as

$$\rho_M \equiv \frac{M_I}{M_D} = \sqrt{\frac{P_I'(0)}{P_D'(0)}}$$

where the indexes D and I stand for dimer and intermediate. ρ_V was calculated from the Porod invariants; the volume V of a species is related to its modified form factor $P'(q)$ by the Porod invariant Q:

$$V = \frac{2\pi^2}{Q} = 2\pi^2 P'(0) \Big(\int_0^\infty q^2 P'(q)\, dq \Big)^{-1}$$

and thus, ρ_V was simply given by V_I/V_D. Since ρ_M and ρ_V designated the same quantity, i.e., the number of dimers in an intermediate, ideally, they had to be equal, and therefore, they were used to check the viability of the extracted modified form factor of the intermediate.

The three-dimensional shapes of species were determined *ab initio* by using the program DAMMIN,[33] which represents the species as a collection of densely packed beads and employs a simulated annealing procedure to search for a compact model. The program minimized discrepancy between a GNOM reconstruction and calculated curves at low resolution (up to $q \approx 0.1$–0.2 Å^{-1}). The search volume for all the species was a sphere, the diameter of which was set to D_{max}. No symmetry was assumed for dimers and intermediates, whereas the icosahedral symmetry was forced for capsids. Twelve independent models were averaged using the program DAMAVER[34] for each species. The resulting model was used as input for subsequent runs of 12 individual models, for a total of at least two refinements. The three-dimensional shapes were rendered with the program CHIMERA.[35]

dx.doi.org/10.1021/ja403550f | J. Am. Chem. Soc. 2013, 135, 15373−15381

RESULTS

First Analysis of TR-SAXS Patterns. Static scattering measurements of NB2-VP1 proteins in dissociation buffer (pH 9.0 and ionic strength 10 mM) revealed objects with a radius of gyration R_g of 42 Å and a maximum dimension D_{max} of 150 Å. The molecular weight calculated from $I(q = 0)$ in absolute units was 113.9 kDa (see SI and Figure S1), which corresponds to dimers, given that NB2-VP1 has a molecular weight of 57 kDa.

At acidic pH and increased ionic strength, NB2-VP1 dimers self-assemble into $T = 3$ VLPs each comprising 90 dimers.[5] By using a stopped-flow device, we rapidly mixed dimers with a buffer, lowering the pH to 6.0 and increasing the ionic strength to 55 mM. The self-assembly kinetics were followed by TR-SAXS from ∼4 ms, corresponding to the mixing time plus the injection delay into a capillary, up to more than 3 h when the scattering patterns showed only a marginal evolution. Data were collected in two separate experiments over two different time scales in order to avoid sample degradation due to repeated exposures. The short time scale spanned milliseconds to seconds (in black in Figure 1A), and the long time scale started from several minutes up to a few hours (in red). The kinetics were performed twice, and the scattering patterns were well reproduced.

Figure 1A shows a 35-fold increase of the low-q intensity over time, which indicated the formation of large objects. The inset gives the radius of gyration estimated by Guinier approximation during the self-assembly. The kinetics were biphasic, with a fast step starting from the first milliseconds and plateauing up to ∼80 s, followed by a very slow step which was not yet finished after 3 h. From an initial value of 45 Å corresponding to dimers, R_g increased to ∼95 Å at the end of the first step, and reached ∼160 Å when data collection was stopped.

Number of Species Contributing to the Scattered Intensities. Superimposed patterns at the early and late stages of the kinetics exhibited isosbestic points (see arrows in Figure 1B,C), which suggested the presence of only two contributing species[26] over each time scale. Therefore, among the numerous possible intermediates, only a few were long-lived enough to make a significant contribution in the scattered intensities. This finding qualitatively supports the conclusions drawn from a graph-theory-inspired study in which a very small set of species are stable and probable enough to emerge from all possible intermediates whose number grows combinatorially with the number of dimers in a capsid.[11]

In order to estimate the number of contributing species, we performed a SVD analysis. As expected from the presence of isosbestic points, the matrix of intensities **I** (see eq 1) could be approximated by a matrix of rank 3 within experimental uncertainties (see Figure S2 and corresponding text). In other words, only three species contributed to the scattered intensities over the whole kinetics.

Selection of a Consistent Kinetic Model. We carried out a global fitting of the scattering patterns with various kinetic models. Global fitting is a powerful technique used in the modeling of spectroscopic data and has been proven efficient in extracting the spectra of intermediate species as well as the reaction rates governing biochemical reactions for instance.[30] The form factor of dimers and that of capsids were imposed in the procedure (see Materials and Methods). We first considered a classical nucleation−elongation model in which a nucleus is slowly formed by aggregation of dimers and then grows rapidly by successive addition of free dimers until completion of the capsid.[36,37] Strikingly, this model, which was successful for the self-assembly of a number of viral capsids,[36] failed to reproduce our data (see SI). Whatever the number of dimers α constituting the nucleus, the figures of merit χ^2 and R were above 5.0 and 20%, respectively. Therefore, NB2 capsids are not formed by polymerization of individual nuclei.

Prasad and co-workers proposed that the NV-VP1 dimer is the building block of the NV capsid and that a major assembly intermediate is a pentamer of dimers.[3] Twelve copies of this intermediate would then be stitched together by 30 intervening dimers to form the final capsid, with a possible second intermediate consisting of a five-fold-symmetric decamer of dimers (a pentamer of dimers surrounded by five interstitial dimers[3]). To extract from our data the actual stoichiometry α of the intermediate I for the NB2 capsid, we tested a range of kinetic models assuming that β of these intermediates combine with γ dimers to form the capsid (Figure 2A). Thus, $\alpha\beta + \gamma = 90$. The reaction orders were equal to the stoichiometric coefficients, except for the dimers of the second reaction; for these, which did not contribute appreciably at this stage, a first order was assumed for the sake of simplicity and numerical convergence. Such a high cooperativity in the self-assembly was reasonable, given that kinetic studies on the hepatitis B virus

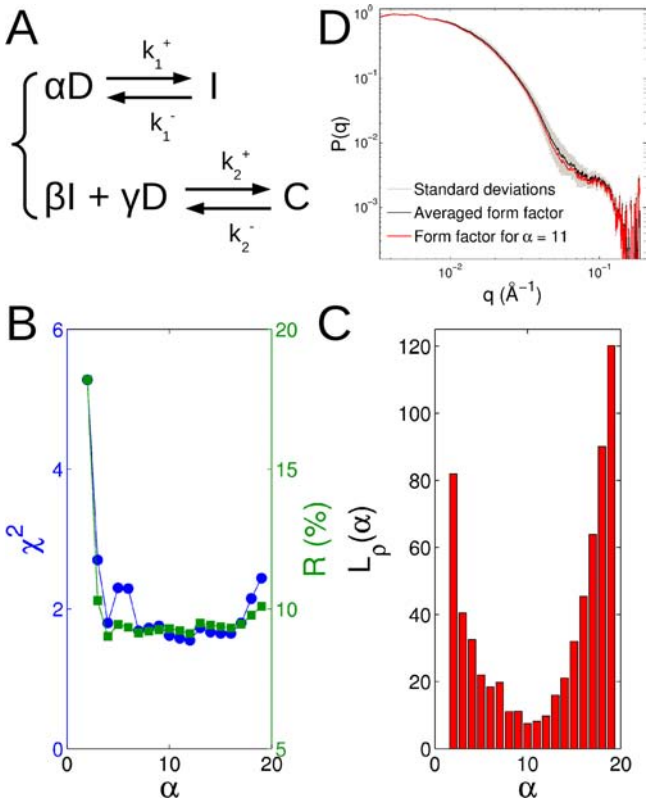


Figure 2. Global fitting with a three-state high-order kinetic model. (A) Set of reactions describing the model. D, I, and C denote dimers, intermediates, and capsids, respectively, while k^+ and k^- are forward and backward reaction rates in both reactions. α, β, and γ are stoichiometric coefficients. (B) Figures of merit χ^2 and R as a function of α. (C) Consistency function $L_\rho(\alpha)$ versus α. (D) Statistics on the form factors of extracted intermediates with α varying from 2 to 19. The averaged form factor is plotted in black, with the standard deviations in gray. The form factor extracted for $\alpha = 11$ is given in red for comparison.

dx.doi.org/10.1021/ja403550f | J. Am. Chem. Soc. 2013, 135, 15373−15381

Table 1. Parameters of the Kinetic Models and Results of Global Fitting

α	β	γ	k_1^+ ($mM^{1-\alpha}$ s^{-1})	k_1^- (s^{-1})	k_2^+ ($mM^{-\beta}$ s^{-1})	k_2^- (s^{-1})	χ^2	R (%)	ρ_M	ρ_V	$R_g{}^a$ (Å)
2	24	42	3.3×10^{-3}	1.2×10^{-4}	2.9×10^{-22}	2.0×10^{-6}	5.28	18.2	4.35	9.24	113
5	12	30	1.6×10^{-5}	1.8×10^{-4}	1.3×10^{-8}	1.9×10^{-6}	2.3	9.45	7.44	8.77	99
7	6	48	7.0×10^{-8}	1.7×10^{-4}	2.8×10^{-6}	2.2×10^{-6}	1.69	9.15	9.76	10.42	103
10	6	30	2.1×10^{-11}	1.6×10^{-5}	3.7×10^{-5}	7.5×10^{-6}	1.62	9.30	12.2	11.5	110
11	6	24	1.5×10^{-12}	8.6×10^{-6}	8.9×10^{-5}	1.1×10^{-6}	1.58	9.23	13.3	11.8	108
15	4	30	3.8×10^{-17}	7.9×10^{-6}	1.9×10^{-4}	6.7×10^{-6}	1.65	9.37	17.7	13.1	111
18	4	18	2.7×10^{-20}	8.7×10^{-6}	9.8×10^{-4}	2.5×10^{-6}	2.15	9.78	19.4	12.1	105

[a]Radius of gyration of the extracted intermediate calculated by GNOM.

capsid and on the cowpea chlorotic mottle virus capsid reported the assembly of an intermediate through a single step, the reaction order of which was measured as the number of dimers making up the intermediate.[13,14]

We considered the (α,β,γ) triplets consistent with the formation of pentameric axes in intermediate species (Figure S4). Since there are 12 five-fold axes in the capsid, $\beta = 24$ for $\alpha = 2$; $\beta = 12$ for $2 < \alpha \leq 6$ (the model of Prasad and co-workers for NV corresponds to $\alpha = 5$, $\beta = 12$, and $\gamma = 30$); $\beta = 6$ for $7 \leq \alpha \leq 12$; and $\beta = 4$ for $\alpha > 12$. Figure 2B shows the agreement between the experimental and fitted data as a function of α, and it demonstrates that the models reproduced the data well for $3 \leq \alpha \leq 18$ ($\chi^2 < 2$ and $R < 10\%$).

In a consistent model, the structure of the extracted intermediate species must be in accord with the reaction scheme in consideration. Let ρ_M and ρ_V be the intermediate-to-dimer mass and volume ratios. Both of them represent the number of dimers making up the intermediate in terms of mass for ρ_M and volume for ρ_V (see Materials and Methods). An ideally consistent model must verify $\alpha = \rho_M = \rho_V$; i.e., the shape of the intermediate—through ρ_M and ρ_V—must be consistent with the number of dimers α entering the kinetic model that permitted its extraction. The ideal consistency condition was here translated into a classical least-squares interpretation that required a consistency function $L_\rho(\alpha)$, defined as

$$L_\rho(\alpha) \equiv (\rho_M - \alpha)^2 + (\rho_V - \alpha)^2 + (\rho_M - \rho_V)^2$$

be minimal for the most probable value of α.

Figure 2C depicts $L_\rho(\alpha)$ with a minimum for $\alpha = 10-11$, which means that the most probable pathway within the general scheme of Figure 2A is the one that converts 10 or 11 dimers into an intermediate. Table 1 lists some of the parameters and fitting results obtained in the process of selection. We can see that ρ_V did not vary much around 10−11, unlike ρ_M which was more correlated to α. In fact, the extracted form factor of intermediates $P_I(q)$ was weakly dependent on α (Figure 2D), which is also evidenced by the small variability of R_g in Table 1; thus, the structure of intermediates was remarkably stable against the choice of stoichiometric coefficients. The intermediate species might consist of a distribution of oligomers made of an average number of α dimers. Nonetheless, provided that α, ρ_M, and ρ_V were close, the distribution was necessarily narrow, reducing to a single oligomer in the case of strict equality ($\alpha = \rho_M = \rho_V$).

In summary, the form factors of intermediates were robust against the choice of stoichiometric coefficients. Given that $L_\rho(\alpha)$ was minimum for $\alpha \approx 10-11$, the variability of ρ_V was weak around 11, and at last, because of structural considerations detailed below, we conclude that the intermediate species comprises an average number of $\alpha = 11$ dimers.

Form Factors and Time Evolution of Extracted Species. Figure 3A depicts a comparison between experimental data and reconstructed patterns from the kinetic model with $\alpha = 11$ (see Table 1 for the reaction rates). The agreement was excellent over 6 temporal decades and for q values covering a wide range of spatial scale (3.3×10^{-3} Å$^{-1} \leq q \leq 0.18$ Å^{-1}). Figure 3B shows the form factors of the three species of interest. The form factor of dimers was taken from the first scattering pattern collected at ~4 ms. Its R_g was 45 Å, slightly larger than that of dimers in dissociation buffer, suggesting a conformational change prior to self-assembly. The form factor of capsids was extracted from the end of the kinetics (see Materials and Methods): R_g = 173 Å and D_{max} = 435 Å. NB2 capsids were therefore larger than NV capsids (R_g = 162 Å and D_{max} = 395 Å), in agreement with previous observations.[5] As for

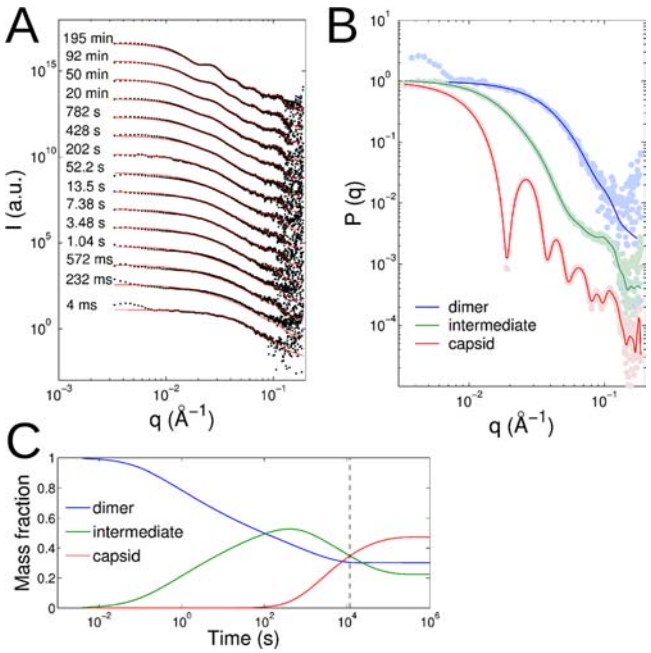


Figure 3. Results of the global fitting for $\alpha = 11$, $\beta = 6$, and $\gamma = 24$. (A) Comparison between the experimental scattering intensities (black dots) and those reconstructed from the global fitting (red line) at various times during the kinetics. (B) Form factors $P(q)$ of the three species in solution: dimers in blue, intermediates in green, and capsids in red. The symbols are experimental (dimers) or extracted (intermediates and capsids) data, while the full lines are reconstructions obtained by the program GNOM. (C) Mass fractions of the three species as a function of time estimated by the kinetic model. The vertical dashed line marks the end of the experimental data.

the intermediates, R_g = 108 Å and D_{max} = 350 Å, which indicated somewhat large and elongated objects.

Figure 3C shows the time evolution of mass fractions for each species deduced from the kinetic model. The majority of intermediates was assembled during the first minute while the fraction of capsids remained negligible. After 100 s, intermediates were slowly converted into capsids with a projected process completion time of 28 h (~10^5 s). Beyond 3 h (approximately indicated by the vertical dashed line in Figure 3C), there would be no more variation of the fraction of dimers, so some intermediates would be dissociated to provide the dimers necessary to the formation of capsids through γ in the kinetic model. Thus, because intermediates were formed rapidly and converted slowly, they had a long lifetime which translated into a significant contribution in the scattering intensities. Note that in terms of molecularity, dimers would still be by far the predominant species: at equilibrium, concentrations for dimers, intermediates, and capsids are projected as 3620, 245, and 63.2 nM, respectively.

Three-Dimensional Reconstructions. We carried out the *ab initio* shape determination of species with a model of densely packed beads that fitted as closely as possible the form factors extracted from global fitting. Figure 4A gives a low-resolution shape for a dimer just prior to assembly. With a two-fold symmetry, it exhibits a protruding domain (pointing upward in Figure 4A) and its height is ~90 Å, which corresponds to the thickness of a capsid. This overall shape is in agreement with the crystal structure of the NV-VP1 dimer in the VLP but suggests a greater extension for the isolated NB2-VP1 dimer.

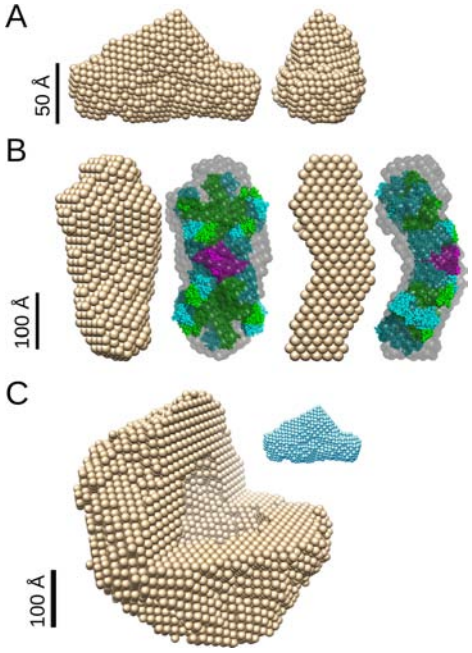


Figure 4. *Ab initio* reconstructions of the species from their extracted form factors. (A) Front (left) and side (right) views of a dimer. (B) Front (two leftmost objects) and side (two rightmost objects) views of an intermediate. The superimposed atomic model is from the crystal structure of NV capsids (PDB reference 1IHM). It consists of two pentamers of A–B dimers (green and cyan) connected by a C–C dimer (magenta). (C) Cutaway view of a capsid. A dimer is shown in cyan at the same scale as in (B) and (C) for size comparison.

A structure for the intermediate species is proposed in Figure 4B. As expected from R_g and D_{max}, intermediates were elongated (~340 Å in height for ~170 Å in width) with a thickness of ~90 Å, corresponding well to the height of dimers. The side view reveals a slight curvature and an apparent top-bottom symmetry. We extracted from the NV-VLP crystal structure a contiguous fragment made of 2 pentamers of dimers connected by a single interstitial dimer. This atomic model, although slightly smaller than the intermediate, matches its shape remarkably well.

A cutaway view of NB2 capsid is shown in Figure 4C. The outer diameter was ~410 Å on average and the inner one ~220 Å, that is, slighly bigger than the NV capsid as already suggested by R_g and D_{max}.

Robustness of Extracted Form Factors. We verified that the kinetic model and more particularly the extracted form factors remained valid when the experimental conditions were varied. Figure 5A is an example in which the final salt concentration was increased to 105 mM. A higher salinity induced a faster self-assembly, one obvious mechanism being the screening of electrostatic repulsions. As early as 157 s, oscillations on scattering patterns were visible, which was not the case for the kinetics depicted in Figure 3A. As time increased, the oscillations became more pronounced, but the kinetics were already close to equilibrium. Global fitting with the kinetic model for α = 11 and form factors given in Figure 3B readily reproduced the experimental data (χ^2 = 2.4 and R = 3.7%). In particular, the fine features of the scattering patterns were nicely fitted. It gave us confidence in the reliability of the extracted form factors, especially that of intermediates. Moreover, the level of intensities was also well reproduced. These results supported the choice of α = 11 for the stoichiometry of intermediates. This analysis also shows that the same kinetic pathway was followed even though the effective interaction energy between species was increased by the higher salinity. There is no evidence for branched or competitive kinetic pathways taking the place of the previous one.

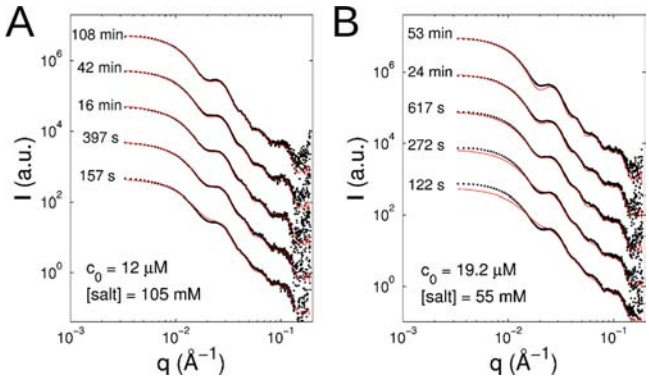


Figure 5. Comparison between experimental scattering intensities (black dots) and patterns reconstructed from the α = 11 kinetic model (red line). The form factors of all the species were fixed and are given in Figure 3B. (A) Global fitting was applied with only the reaction rates to adjust. Initial dimer concentration was 12 μM and ionic strength 105 mM. (B) There were no adjustable parameters in that case. The reaction rates were those given in Table 1. Initial dimer concentration was 19.2 μM and ionic strength 55 mM.

At last, we tested the predictive ability of the kinetic model. Using parameters given in Table 1 for α = 11 and again the form factors shown in Figure 3B, we confronted the model with data collected at an initial dimer concentration of 19.2 μM (12

dx.doi.org/10.1021/ja403550f | J. Am. Chem. Soc. 2013, 135, 15373–15381

μM previously). All the parameters were fixed since the interactions between species should be the same and the reaction rates were not expected to be affected. Figure 5B shows that the model had some difficulty predicting the early appearance of capsids at t = 122 s. Nevertheless, from t = 272 s and beyond, scattering patterns were fairly reconstructed (χ^2 = 4.8 and R = 12%). Fine features and levels of intensities were once more reproduced. Both results depicted in Figure 5 demonstrate the robustness of the extracted form factors against a change of experimental conditions and validate the stoichiometry entering the kinetic model.

DISCUSSION

The spontaneous, accurate, and fast self-assembly of icosahedral viral capsids requires well-defined intermediates that may be very limited in number.[11] However, there is little experimental data on the intermediates due to the inherent difficulties in characterizing transient molecular species in an assembly pathway. Three routes have been used to derive self-assembly pathways for icosahedral viruses: equilibrium, kinetics, and crystal-structure-derived studies. The latter approach has been the only one so far allowing prediction of not only stoichiometries but also high-resolution shapes for intermediates. There are two underlying assumptions in these analyses, as explicitly laid out by, e.g., Janin and co-workers:[38] first, that the contact surfaces in an assembly intermediate are subsets of the interfaces in the final capsid, as seen in atomic detail in crystal structures, and second, that the interfaces form in order of decreasing area.

Following this line of reasoning, Prasad and co-workers proposed that the NV-VP1 dimer is the building block of the NV capsid (since the dimer interface has the largest area) and that a major assembly intermediate would be a pentamer of dimers (since the interfaces between shell domains within a pentamer of dimers are the most prominent), i.e., a 10-mer, with a possible second intermediate consisting of another pentamer of dimers surrounded by five interstitial dimers, i.e., a 20-mer.[3] The "dimer as a building block" hypothesis was fully verified *in vitro* for the NV-VLP in several equilibrium studies, including native mass spectrometry of partially or totally dissociated NV-VP1.[4,39] In addition, data were consistent with a sheet-like intermediate species. However, the presence of 10-mers coexisting with 20-mers was not verified.

We show that, for the related NB2, the capsid building block is also a dimer[5] that assembles into capsids, apparently via a single intermediate oligomer. After a fast (millisecond) conformational change of the dimer, the intermediate starts appearing within the first second. We could extract from the kinetic data not only the stoichiometry of this intermediate but also a nanometer-resolution form factor that is robust against a range of kinetic models (Figure 2D). Despite a stoichiometry close to 10 dimers, this form factor is inconsistent with a five-fold-symmetric 20-mer (Figure S5). Indeed, the reconstructed intermediate shape clearly has no five- or three-fold symmetry but is a stave-like object. Examination of the NV-VLP structure shows that this shape most closely resembles a 22-mer made of two pentamers of dimers connected by a single interstitial dimer. The residual discrepancy between the intermediate and the model extracted from the crystal structure of NV-VLP (Figure 4B) may be due to two factors: first, the NB2-VLP is slightly larger than the NV-VLP, and second, a free intermediate may be less curved than the same assembly in the capsid. Whatever the actual arrangement of subunits, such an interlocking shape is well suited to precise self-assembly into capsids (Figure 6) with the addition of free dimers that are still present (albeit no longer detectable) at this stage (Figure 3C).

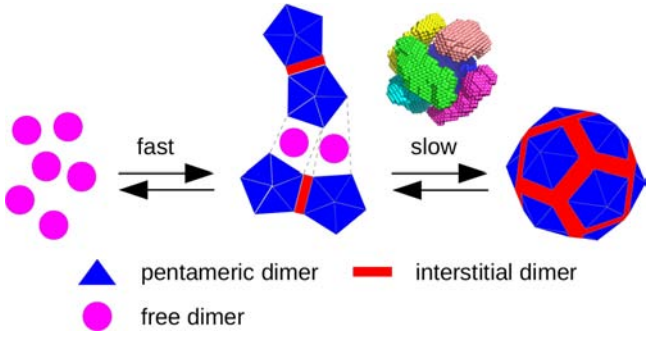


Figure 6. Kinetic scheme of norovirus capsid assembly. Free dimers are represented in magenta, dimers related by five-fold symmetry in the final capsid in blue, and interstitial dimers in red. Above the last assembly step, a representation of interlocking intermediates is given as a possible mechanism. Six intermediates, each in a different color, have been positioned above the six contiguous fragments made of two pentamers of dimers connected by an interstitial dimer in the NV capsid (not shown).

The apparent collapse of the intermediate into capsids may reflect the actual mechanism by which norovirus capsids are formed, although it is probably through a more complex pathway than the single highly cooperative step implemented here. For instance, we cannot rule out that this apparently single intermediate is actually a narrow distribution of intermediate species of similar stoichiometries and shapes. Nonetheless, our data are clearly inconsistent with a nucleation–elongation model with successive addition of dimers to the intermediate (Figure S3). In such a model, nuclei are converted as soon as they are created, and their concentration remains below detection thresholds. Here the intermediates are created rapidly and converted through a step intrinsically slower because these bulky objects need time to reorient themselves and to capture free dimers. As a result of this kinetic bottleneck, intermediates have a long lifetime and accumulate in the solution. It should confer to the self-assembly a certain robustness against kinetic traps. Interestingly, Zlotnick and co-workers reported the existence of two branched pathways in the *in vitro* assembly of CCMV capsids:[13] at moderate protein concentrations, T = 3 VLPs were formed via a nucleation–elongation process, while at high protein concentrations, a large fraction of pseudo-T = 2 VLPs resulted from the direct association of overabundant pentamers of dimers. There was no evidence for such different reaction mechanisms with NB2, as we always obtained T = 3 VLPs whatever the protein concentration and ionic strength, provided that the assembly was initiated from free dimers. However, we cannot exclude the existence of competitive pathways which might be detected by exploring other experimental conditions.

More generally, we conclude that, although reasoning by symmetry—namely, by anticipating that the building blocks have a five-fold symmetry—is a powerful guide, it also has limitations: The most stable and only detectable intermediate is here stave-like with at most a two-fold symmetry (Figure 4B). The formation of such a long-lived and interlocking intermediate on the way to an icosahedral capsid may make

the system less prone to misassembly. The molecular mechanisms by which the norovirus capsid actually achieves assembly of such an intermediate are unknown at present. In particular, specific anticooperativity is required to prevent further accretion of free dimers to the stave-like intermediate. Likely candidates for such an anticooperativity are the known points of flexibility allowing allostery in the norovirus capsid. These are the hinge between the shell and protruding domains on one hand and the N-terminal arm interconnecting the inner capsid surface on the other hand. The N-terminal arm would be a good candidate for allosteric stabilization of the intermediate, as it assumes several conformations in the NV capsid[3] and harbors (de)protonable residues that may be involved in the pH-dependence of NB2 and NV (dis)assembly *in vitro*.[5] Relocation of the arm in the intermediate may thus hinder further capture of isolated, extended dimers. An analogous allosteric, symmetry-breaking mechanism has been reported for a mutant of the unrelated HBV capsid and proposed to reflect a key capability of this capsid to form alternative interactions across polygonal intermediates, preventing assembly until the right time and place.[40] An inner N-terminal arm allowing quasi-equivalence switches is a common if not universal feature of $T = 3$ capsid proteins with the jellyroll fold, such as the norovirus VP1 and the CCMV capsid protein.[12,13,41] Unlike most such proteins, the VP1 norovirus N-terminal arm is not positively charged, and expression of VP1 outside the proper infectious context leads to assembly of empty VLP. Because of this and the difficulty in propagating noroviruses in cell culture, the mechanisms by which these highly infectious viruses[2] efficiently encapsidate the viral genome *in vivo* are unknown. The ability to form a stable intermediate capable of waiting for the proper trigger to encapsidate the viral genome could be such a mechanism. The architecture of the intermediate reported here is unexpected, as it is not simply a subassembly based on one of the rotational symmetries in the final capsid, as usually assumed and previously proposed for other viruses, e.g., three-fold (HBV) or five-fold (CCMV, NV) symmetry. This may point to a particular cooperative assembly mechanism in noroviruses.

■ ASSOCIATED CONTENT

Supporting Information

Additional data analyses and details on SVD. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

guillaume.tresset@u-psud.fr; stephane.bressanelli@vms.cnrs-gif.fr

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank Dominique Durand for fruitful advices on protein scattering. They also gratefully acknowledge the European Synchrotron Radiation Facility (Grenoble, France) and the SOLEIL facility (Saint-Aubin, France) for allocation of synchrotron beam time, as well as Jérémie Gummel and Javier Pérez for their assistance on ID02 and SWING beamlines, respectively. The research was partly funded by the Agence Nationale pour la Recherche (contract ANR-06 No. 07/2.210 076/F), the Triangle de la Physique (contract 2009-073T), and CNRS through the "Interface Physique, Chimie, Biologie" program.

■ REFERENCES

(1) Rohayem, J.; Bergmann, M.; Gebhardt, J.; Gould, E.; Tucker, P.; Mattevi, A.; Unge, T.; Hilgenfeld, R.; Neyts, J. *Antiviral Res.* **2010**, *87*, 162−178.

(2) Teunis, P. F. M.; Moe, C. L.; Liu, P.; Miller, S. E.; Lindesmith, L.; Baric, R. S.; Le Pendu, J.; Calderon, R. L. *J. Med. Virol.* **2008**, *80*, 1468−1476.

(3) Prasad, B. V.; Hardy, M. E.; Dokland, T.; Bella, J.; Rossmann, M. G.; Estes, M. K. *Science* **1999**, *286*, 287−290.

(4) Shoemaker, G. K.; van Duijn, E.; Crawford, S. E.; Uetrecht, C.; Baclayon, M.; Roos, W. H.; Wuite, G. J. L.; Estes, M. K.; Prasad, B. V. V; Heck, A. J. R. *Mol. Cell. Proteomics* **2010**, *9*, 1742−1751.

(5) Tresset, G.; Decouche, V.; Bryche, J.-F.; Charpilienne, A.; Le Coeur, C.; Barbier, C.; Squires, G.; Zeghal, M.; Poncet, D.; Bressanelli, S. *Arch. Biochem. Biophys.* **2013**, accepted for publication.

(6) Zakhour, M.; Ruvoën-Clouet, N.; Charpilienne, A.; Langpap, B.; Poncet, D.; Peters, T.; Bovin, N.; Le Pendu, J. *PLoS Pathol.* **2009**, *5*, e1000504.

(7) Vildevall, M.; Grahn, A.; Oliver, S. L.; Bridger, J. C.; Charpilienne, A.; Poncet, D.; Larson, G.; Svensson, L. *J. Med. Virol.* **2010**, *82*, 1241−1246.

(8) Douglas, T.; Young, M. *Science* **2006**, *312*, 873−875.

(9) *Viruses and nanotechnology*; Manchester, M., Steinmetz, N. F., Eds.; Springer: Heidelberg, Germany, 2009.

(10) Aniagyei, S. E.; Kennedy, C. J.; Stein, B.; Willits, D. A.; Douglas, T.; Young, M. J.; De, M.; Rotello, V. M.; Srisathiyanarayanan, D.; Kao, C. C.; Dragnea, B. *Nano Lett.* **2009**, *9*, 393−398.

(11) Moisant, P.; Neeman, H.; Zlotnick, A. *Biophys. J.* **2010**, *99*, 1350−1357.

(12) Johnson, J. M.; Tang, J.; Nyame, Y.; Willits, D.; Young, M. J.; Zlotnick, A. *Nano Lett.* **2005**, *5*, 765−770.

(13) Zlotnick, A.; Aldrich, R.; Johnson, J. M.; Ceres, P.; Young, M. J. *Virology* **2000**, *277*, 450−456.

(14) Zlotnick, A.; Johnson, J. M.; Wingfield, P. W.; Stahl, S. J.; Endres, D. *Biochemistry* **1999**, *38*, 14644−14652.

(15) Elrad, O. M.; Hagan, M. F. *Nano Lett.* **2008**, *8*, 3850−3857.

(16) Zlotnick, A.; Mukhopadhyay, S. *Trends Microbiol.* **2011**, *19*, 14−23.

(17) Basnak, G.; Morton, V. L.; Rolfsson, O.; Stonehouse, N. J.; Ashcroft, A. E.; Stockley, P. G. *J. Mol. Biol.* **2010**, *395*, 924−936.

(18) Dykeman, E. C.; Grayson, N. E.; Toropova, K.; Ranson, N. A.; Stockley, P. G.; Twarock, R. *J. Mol. Biol.* **2011**, *408*, 399−407.

(19) Casini, G. L.; Graham, D.; Heine, D.; Garcea, R. L.; Wu, D. T. *Virology* **2004**, *325*, 320−327.

(20) Chen, C.; Kao, C. C.; Dragnea, B. *J. Phys. Chem. A* **2008**, *112*, 9405−9412.

(21) Grillo, I. *Curr. Opin. Colloid Interface Sci.* **2009**, *14*, 402−408.

(22) Doniach, S. *Chem. Rev.* **2001**, *101*, 1763−1778.

(23) Russell, R.; Millett, I. S.; Tate, M. W.; Kwok, L. W.; Nakatani, B.; Gruner, S. M.; Mochrie, S. G. J.; Pande, V.; Doniach, S.; Herschlag, D.; Pollack, L. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 4266−4271.

(24) Ortore, M. G.; Spinozzi, F.; Vilasi, S.; Sirangelo, I.; Irace, G.; Shukla, A.; Narayanan, T.; Sinibaldi, R.; Mariani, P. *Phys. Rev. E* **2011**, *84*, 061904.

(25) Tuma, R.; Tsuruta, H.; French, K. H.; Prevelige, P. E. *J. Mol. Biol.* **2008**, *381*, 1395−1406.

(26) Kler, S.; Asor, R.; Li, C.; Ginsburg, A.; Harries, D.; Oppenheim, A.; Zlotnick, A.; Raviv, U. *J. Am. Chem. Soc.* **2012**, *134*, 8823−8830.

(27) Abécassis, B.; Testard, F.; Spalla, O.; Barboux, P. *Nano Lett.* **2007**, *7*, 1723−1727.

(28) Lund, R.; Willner, L.; Monkenbusch, M.; Panine, P.; Narayanan, T.; Colmenero, J.; Richter, D. *Phys. Rev. Lett.* **2009**, *102*, 188301.

(29) Tabor, R. F.; Eastoe, J.; Grillo, I. *Soft Matter* **2009**, *5*, 2125−2129.

(30) Henry, E. R. *Biophys. J.* **1997**, *72*, 652−673.

(31) Schmidt, M.; Rajagopal, S.; Ren, Z.; Moffat, K. *Biophys. J.* **2003**, *84*, 2112−2129.
(32) Putnam, C. D.; Hammel, M.; Hura, G. L.; Tainer, J. A. *Q. Rev. Biophys.* **2007**, *40*, 191−285.
(33) Svergun, D. I. *Biophys. J.* **1999**, *76*, 2879−2886.
(34) Volkov, V. V.; Svergun, D. I. *J. Appl. Crystallogr.* **2003**, *36*, 860−864.
(35) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J. Comput. Chem.* **2004**, *25*, 1605−1612.
(36) Zlotnick, A. *J. Mol. Recognit.* **2005**, *18*, 479−490.
(37) Hagan, M. F.; Elrad, O. M. *Biophys. J.* **2010**, *98*, 1065−1074.
(38) Janin, J.; Bahadur, R. P.; Chakrabarti, P. *Q. Rev. Biophys.* **2008**, *41*, 133−180.
(39) Uetrecht, C.; Barbu, I. M.; Shoemaker, G. K.; van Duijn, E.; Heck, A. J. R. *Nat. Chem.* **2011**, *3*, 126−132.
(40) Packianathan, C.; Katen, S. P.; Dann, C. E., III; Zlotnick, A. *J. Virol.* **2010**, *84*, 1607−1615.
(41) Tang, J.; Johnson, J. M.; Dryden, K. A.; Young, M. J.; Zlotnick, A.; Johnson, J. E. *J. Struct. Biol.* **2006**, *154*, 59−67.