Energetics and Kinetic Assembly Pathways of Hepatitis B Virus Capsids in the Presence of Antivirals

Kalouna Kra, Siyu Li, Laetitia Gargowitsch Poncet, Jéril Degrouard, Javier Pérez, Roya Zandi, Stéphane Bressanelli, Guillaume Tresset

To cite this version:

Kalouna Kra, Siyu Li, Laetitia Gargowitsch Poncet, Jéril Degrouard, Javier Pérez, et al.. Energetics and Kinetic Assembly Pathways of Hepatitis B Virus Capsids in the Presence of Antivirals. ACS Nano, 2023, 17 (13), pp.12723-12733. 10.1021/acsnano.3c03595 . hal-04160101

HAL Id: hal-04160101
https://hal.science/hal-04160101v1

Submitted on 12 Jul 2023

Energetics and kinetic assembly pathways of Hepatitis B virus capsids in the presence of antivirals

Kalouna Kra,[1,2] Siyu Li,[3] Laetitia Gargowitsch,[1] Jéril Degrouard,[1] Javier Pérez,[4] Roya Zandi,[3]
Stéphane Bressanelli[2] and Guillaume Tresset[1*]*

[1]Université Paris-Saclay, CNRS, Laboratoire de Physique des Solides, 91405 Orsay, France.

[2]Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC),
91198 Gif-sur-Yvette, France.

[3]Department of Physics and Astronomy, University of California, Riverside, California 92521,
United States.

[4]SOLEIL Synchrotron, 91192 Gif-sur-Yvette, France.

*Corresponding authors : stephane.bressanelli@i2bc.paris-saclay.fr;
guillaume.tresset@universite-paris-saclay.fr

Abstract

Capsid assembly modulators (CAMs) are antiviral molecules that disturb the formation of icosahedral viral capsids, in particular, those of the Hepatitis B virus (HBV). We report an integrated, physics-driven study elucidating quantitatively the effects of two classes of CAMs on HBV capsid assembly. Time-resolved small-angle X-ray scattering measurements revealed accelerated self-assembly processes that implied the increase of subunit binding energy from 9- up to 18-fold the thermal energy due to CAMs. Cryotransmission electron microscopy images showed that both classes induce various changes in capsid morphology: from a slight elongation, unrecognized in previous work, to a strong deformation with a capsid size more than twice as large. The observed capsid morphologies were closely reproduced in coarse-grained simulations by varying the Föppl-von-Kármán number, thus pointing out the role of CAMs in altering the capsid elastic energy. Our results illuminate the mechanisms of action of CAMs on HBV capsid assembly at high spatiotemporal resolution, and may bring perspectives on virus-derived nanocapsules with tunable morphologies.

Keywords: Hepatitis B virus, capsid assembly, assembly modulators, kinetic modelling, capsid mechanics, time-resolved small-angle X-ray scattering, cryotransmission electron microscopy

The Hepatitis B virus (HBV) is a major health problem worldwide. Despite the existence of a safe and effective vaccine, more than 292 million people are chronically infected by HBV, mainly in Africa and South Asia, due to a poor vaccination rate. This virus causes serious pathologies such as liver cirrhosis in some patients, but also hepatocellular cancer that can lead to death.[1] The treatments currently approved use nucleos(t)ide analogues. However, the duration of the treatments is very long, usually lifelong,[2] and they only permit to control the infection but not to cure it. Therefore, it is important to develop antiviral compounds with alternative mechanisms of action that may be used more easily, in combination or not with other drugs, in order to completely treat the infection.

HBV is an enveloped virus with an icosahedral nucleocapsid that contains the viral genome in the form of a 3.2 kb partially double-stranded DNA. The HBV capsid, which plays an essential role in the HBV viral cycle, is composed of a homodimeric Core protein.[3] Core is a 183-residue-long protein composed of two domains. The 149 N-terminal residues are an alpha-helical assembly domain called N-terminal domain (NTD) referred to as Core protein 149 (Cp149).[4] The 34 remaining residues are the intrinsically disordered C-terminal domain (CTD) which interacts with RNA and is required for its encapsidation. Cp149 can self-assemble[5,6] both *in vivo* and *in vitro* mainly in $T = 4$ icosahedral capsid (120 dimers) but also in $T = 3$ icosahedral capsid (90 dimers).

Early studies suggested that Cp149 capsid assembly follows a nucleation-elongation process.[7–9] Nucleation consists of the formation of a critical nucleus made of dimers, the latter being the building blocks of the assembly and hereafter called subunits. Free subunits are sequentially added to the nucleus during the elongation phase to form the final capsid. There are at most a few long-lived intermediate species present during this step.[10] Many distinct factors, including the ionic strength,[11] the stability of the intra-subunit interface[12] and/or the contact region of the subunits,[13] influence HBV capsid assembly and its duration, but this regulation, however, occurs without the assistance of chaperone proteins.

Since the intermediate steps can occur very fast, following and solving the structures involved in the assembly are difficult. With a spatial sensitivity of a few nanometers and a temporal precision of a few milliseconds, time-resolved small-angle X-ray scattering (TR-SAXS) enables to track complex reactions,[14] notably those pertaining to virus self-assembly.[10,15–22] TR-SAXS data revealed that the assembly of Cp149 empty capsids appeared to be a two-state process involving solely dimeric subunits and $T = 4$ capsid under low salt conditions – *i.e.*, physiological pH ~7.5 and monovalent salt concentration ~150 mM.[18] Nevertheless, the development of $T = 3$ capsids and the brief accumulation of intermediates containing between 7 and 35 subunits were reported under high salt conditions[18] – *i.e.*, > 300 mM. In another study,[20] under conditions where pH was far from being physiological – *i.e.*, pH 9.0 –, neither the intermediates nor the $T = 3$ minor species were detected at high salt concentration. However, an additional assembly step, referred to as relaxation, was observed, during which misassembled capsids rearranged themselves via a disorder-order transition into well-ordered capsids.

Even though the HBV assembly to icosahedral shell is robust under many different *in vitro* and *in vivo* conditions, small molecules known as capsid assembly modulators (CAMs) have been developed to target the HBV Core protein, which has proven to be a successful antiviral target.[2] Through a number of ways, they can inhibit HBV replication and possibly even combat virus persistence:[3,23] Their main effect is to disrupt capsid assembly, but some CAMs have been shown to be capable also of blocking the recycling pathway of nucleocapsids to the nucleus. CAMs actually accelerate capsid assembly in a dose-dependent manner by binding to a hydrophobic pocket that mediates contacts between dimers of the Core protein in the final capsid.[24,25] According to their mode of action, two classes of modulators have been defined. Heteroaryldihydropyrimidines (HAP) induce the formation of aberrant capsids[26] and are the only molecules belonging to the class CAM-A (for aberrant).[27] Class CAM-E (for empty)[27] is made up of several different types of molecules, *e.g.*, sulfamoylbenzamides (SBA), phenyl propenamides derivatives (PPA),[25,28] and glyoxamide-pyrrolamide (GPA). The latter molecules, unlike HAP, are thought to induce the formation of morphologically intact HBV capsids. However, they inhibit the formation of nucleocapsids containing pregenomic (pg)RNA.[29,30] Several *in vitro* investigations were conducted about the effects of modulators on the capsid morphology, but the kinetic assembly pathways with CAMs have not been explored yet.

We report here a multidisciplinary and integrated study aimed at shedding light quantitatively on the effects of JNJ-632[25,31] and Bay 41-4109,[32,33] two modulators of type CAM-E and CAM-A, respectively, on the assembly of Cp149 capsids at both nano- and millisecond-scale resolutions. We chose physicochemical conditions – *i.e.*, salinity of 150 mM, pH 7.5 and temperature of 37 °C – that reproduced those found in the intracellular milieu. Thanks to the high spatiotemporal sensitivity of TR-SAXS, we were able to probe the assembly timescales and to investigate the capsid morphologies with different stoichiometric ratios of modulator. We further investigated capsid morphologies by cryotransmission electron microscopy (cryoTEM) to provide nanometer-scale structural details at various time points. By using a kinetic model of phase transition, we were able to measure the effective subunit binding free energy in the presence of modulator. Finally, we related the observed morphologies of final capsids to their elastic properties via coarse-grained simulations, and we propose mechanisms by which modulators might act at the molecular level.

Results

Without modulators, a majority of $T = 4$ capsids were self-assembled.

SAXS data measurements of subunits prior to assembly were consistent with Cp149 dimers, albeit slightly aggregated (Figure S1). Capsids were assembled by mixing manually subunits at concentrations of 15 μM to 30 μM with an assembly buffer, bringing the monovalent salt ammonium acetate up to a final concentration of 150 mM, at pH 7.5 and 37°C. The scattering intensities of reassembled capsids were measured after 48 hours. They displayed oscillations with a first minimum at q = 0.022 Å^{-1}, which indicated that purified subunits led to the formation of a majority of $T = 4$ capsids (Figure S2).

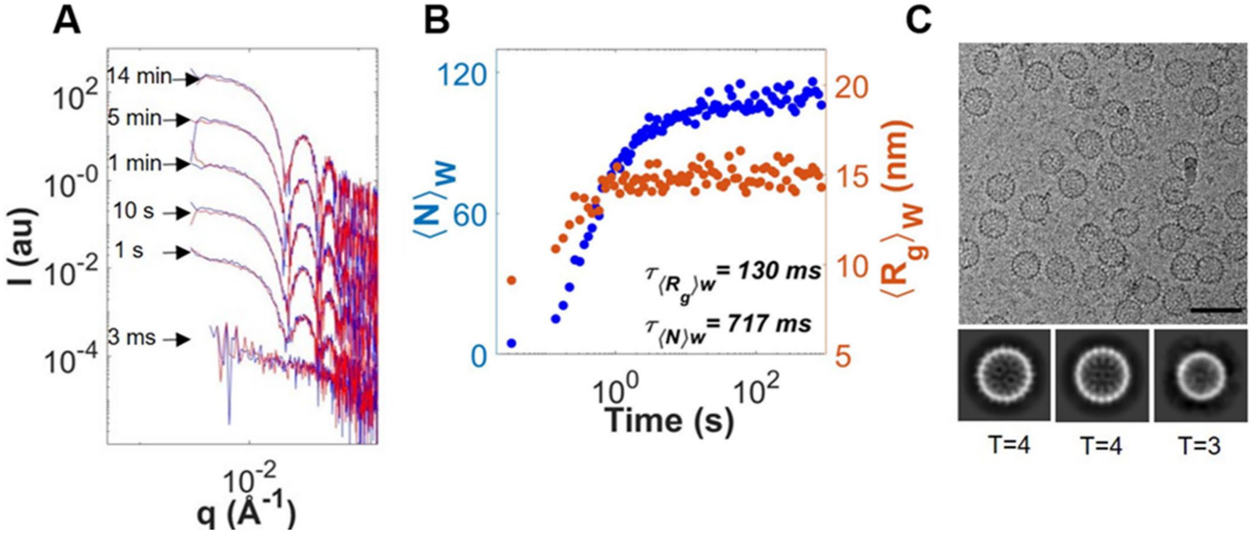

Figure 1. TR-SAXS and cryoTEM measurements for the assembly of capsids without modulator. (A) TR-SAXS patterns of two repeated assembly experiments (blue and red lines, respectively) for 30 μM of subunits. The assembly was triggered with 150 mM of ammonium acetate. The patterns at different time points are shifted for clarity. (**B)** Evolution of the mean aggregation number $\langle N\rangle_\mathrm{w}$ (blue discs) and the mean radius of gyration $\langle R_\mathrm{g}\rangle_\mathrm{w}$ (orange discs) as a function of time for the assembly of capsids with 30 μM of subunits. The timescales were obtained by fitting an exponential decay function. (**C**) CryoTEM micrograph (top) and 2D classification (bottom) of manually assembled capsids. Capsids were assembled with a subunit concentration of 30 μM then concentrated 10 times. Scale bar is 50 nm. The box size for 2D classification is 50 nm. Number of $T = 4$ particles: 3,476; and number of $T = 3$ particles: 47.

To examine the processes of capsid assembly, TR-SAXS experiments were carried out using a stopped-flow mixer. Capsid assembly was performed at 37°C by rapidly mixing a solution of subunits with a buffer solution containing a high concentration of ammonium acetate, both at pH 7.5. The final concentration of subunits was 30 μM (~1 g.L^{-1}), and the final salt concentration was

150 mM. The scattering patterns were collected at different time points for 15 minutes, and each condition was duplicated. SAXS curves are presented in Figure 1A. The SAXS pattern at the earliest time point (3 ms) exhibited a monotonic decrease of the scattering intensity with the wavenumber q, indicating that initially, mostly dissociated subunits were present in the solution. One second after triggering the assembly, oscillations appeared on the SAXS patterns, and the first minimum at q = 0.023 Å^{-1} indicated the formation of $T = 4$ capsids. After 4 s, the patterns remained essentially unchanged, suggesting that the assembly had reached equilibrium. No signs of aggregation were seen during the assembly and capsids were self-assembled with excellent repeatability (Figure 1A, compare blue and red lines). The SAXS curves obtained at 4 s and 850 s after triggering assembly with the stopped-flow mixer, and that measured at equilibrium via manual mixing showed a strong degree of similarity (Figure S3), demonstrating that the use of a stopped-flow mixer produced eventually the same $T = 4$ capsids as those assembled manually under the same conditions.

Figure 1B shows the mean radius of gyration $\langle R_\mathrm{g}\rangle_\mathrm{w}$ and the mean aggregation number $\langle N\rangle_\mathrm{w}$ extracted from each SAXS curve of the kinetics (see Methods for details). The extracted timescales for each quantity were calculated by fitting the data with single exponential decay functions. The extracted timescale for $\langle R_\mathrm{g}\rangle_\mathrm{w}$ was 130 ± 22 ms (95% confidence interval). $\langle R_\mathrm{g}\rangle_\mathrm{w}$ stopped evolving about 1 s after the start of the assembly and plateaued at a value of about 150 Å. The mean aggregation number $\langle N\rangle_\mathrm{w}$, corresponding to the weight-averaged number of subunits found in each object during assembly plateaued slightly below 120, which would have been the value obtained for a pure solution of $T = 4$ capsids. This was due to the presence of remaining free subunits at the end of assembly as expected from the law of mass action. The extracted timescale observed for the mean aggregation number was 717 ± 54 ms (95% confidence interval), which was higher than that for the mean radius of gyration, *i.e.*, 130 ms.

The capsid morphology of manually assembled capsids was further investigated by cryotransmission electron microscopy (cryoTEM). As shown in the 2D classes (Figure 1C), a vast majority of $T = 4$ capsids were formed. Furthermore, a very small number of incomplete capsids was observed and less than 5% of the capsids exhibited a size compatible with the $T = 3$ symmetry. From this 2D classification, a 3D map (Figure S4) was obtained by imposing icosahedral symmetry during 3D reconstruction, which matched very well the $T = 4$ crystal structure.

CAM-E shortened the assembly timescale while inducing structural changes on capsids.

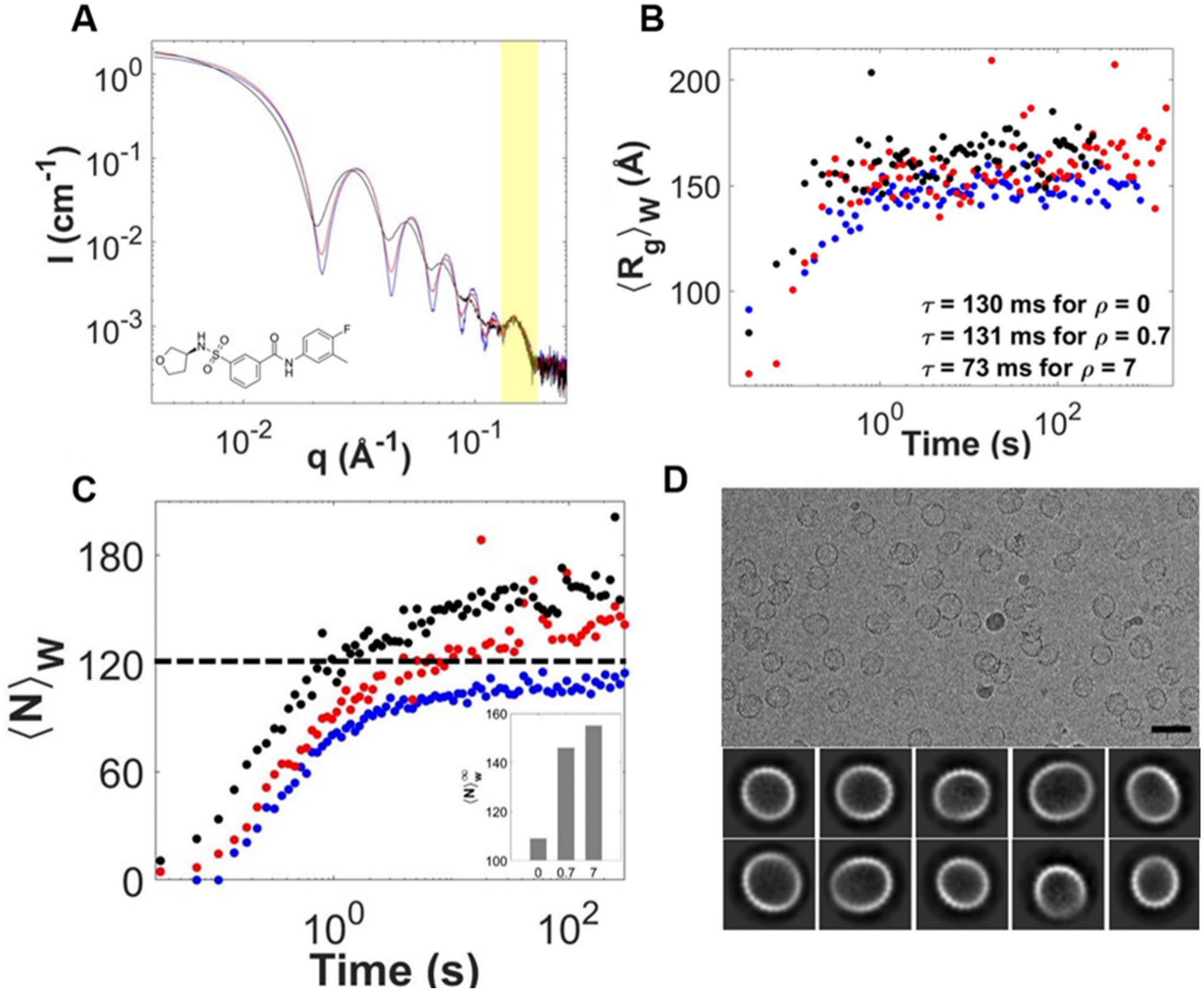

Figure 2. Capsid assembly in the presence of CAM-E. (A) SAXS patterns at equilibrium with modulator-to-subunit molar ratios ρ of 0 (blue), 0.7 (red) and 7.0 (black). The shaded area highlights the region where all the patterns were superimposed. The chemical structure of CAM-E is represented at the bottom left. (**B)** Evolution of the mean radius of gyration $\langle R_{\mathrm{g}} \rangle_{\mathrm{w}}$ as a function of time. The timescales were obtained by fitting exponential decay functions. **(C)** Plot of $\langle N \rangle_{\mathrm{w}}$ as a function of time. The inset gives the equilibrium mean aggregation numbers $\langle N \rangle_{\mathrm{w}}^{\infty}$ as a function of ρ. For **(A)**, (**B)**, and (**C)**, the subunit concentration was 30 μM and the color code for the ρ values is the same. (**D)** CryoTEM micrograph (top) and 2D classification (bottom) of capsids

assembled with ρ = 7.0 and a subunit concentration of 30 µM before being concentrated about 10 times. The scale bar and the box size for 2D classification are 50 nm.

Capsid assembly was triggered manually with 150 mM of ammonium acetate, 30 µM of subunits, and varying modulator-to-subunit molar ratios ρ of CAM-E (see chemical structure on Figure 2A). The scattering intensities measured at equilibrium (Figure 2A) indicated that the forward intensity I_0 was 10% to 23% larger upon assembly in the presence of CAM-E. Moreover, the oscillations were less pronounced, and the first minimum was slightly shifted to smaller q-values. These effects increased with ρ but gradually saturated above $\rho = 3$ (Figure S5). The differences observed on the SAXS curves show that CAM-E definitely induces structural modifications on capsids. Still, at large q-values (see the shaded area in Figure 2A), all SAXS curves were superimposed, which means that the local arrangements of subunits seem to be well preserved.

Figure 2B depicts the mean radius of gyration $\langle R_{\mathrm{g}} \rangle_{\mathrm{w}}$ extracted from TR-SAXS measurements performed for capsid assembly in the presence of CAM-E at $\rho = 0.7$ and 7.0. The equilibrium value of $\langle R_{\mathrm{g}} \rangle_{\mathrm{w}}$ increased with ρ from about 150 Å in the absence of modulator to 170 Å at $\rho =$ 7.0. Moreover, at the latter ratio, CAM-E accelerated by a factor 1.8 the capsid assembly. The extracted timescale observed for an assembly with ρ = 7.0 for the CAM-E was equal to 73 ± 22 ms (95% confidence interval) and the extracted timescale for an assembly in the absence of modulators was equal to 130 ms. Figure 2C shows that the equilibrium value $\langle N \rangle_{\mathrm{w}}^{\infty}$ was over 120 in the presence of CAM-E, which means that the final capsids were larger than the native $T = 4$ ones. Just like in the absence of modulator, $\langle N \rangle_{\mathrm{w}}$ evolved more slowly than $\langle R_{\mathrm{g}} \rangle_{\mathrm{w}}$.

2D classification from cryoTEM images at $\rho = 7.0$ (Figure 2D) obtained at equilibrium revealed slightly larger-than-native, ellipsoidal capsids (see Figure S6 for low-resolution 3D reconstructions). No spherical capsids were identified. By contrast, at $\rho = 0.7$, 75% of the particles were similar to projections of $T = 4$ capsids formed without modulator – although the lower quality of the 3D reconstruction indicated that the icosahedral symmetry was likely less perfect –, and 25% appeared to be larger and ellipsoidal (Figure S7).

In summary, under conditions where a vast majority of $T = 4$ capsids are assembled, CAM-E leads to slightly larger and elongated capsids, the proportion of which increases with the amount of modulator.

CAM-A modulator had a dramatic effect on capsid morphology.

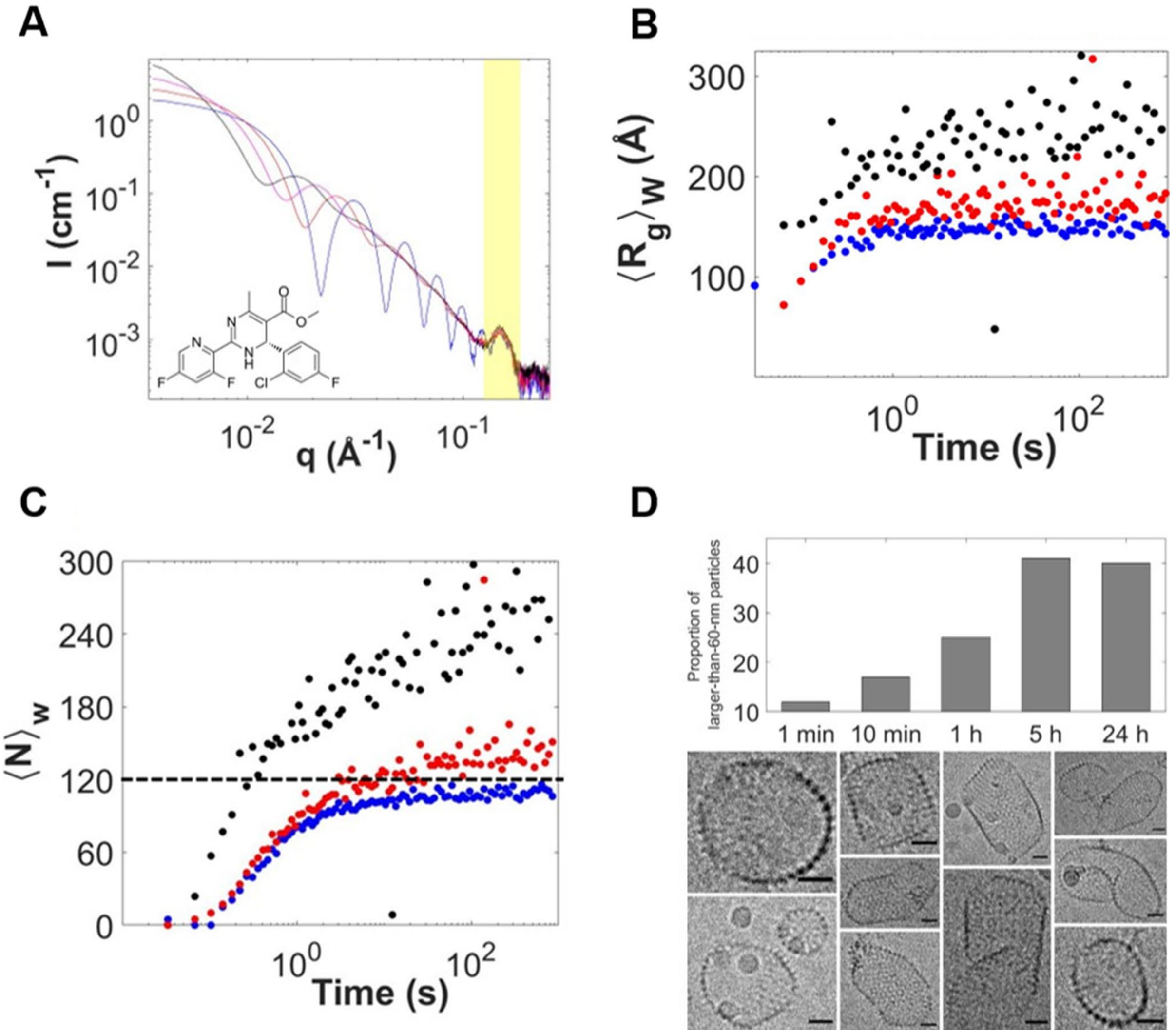

Figure 3. Capsid assembly in the presence of CAM-A. (A) SAXS patterns at equilibrium with modulator-to-subunit molar ratios ρ of 0 (blue), 0.34 (red), 1.7 (magenta) and 3.4 (black). The shaded area highlights the region where all the patterns were superimposed. The chemical structure of CAM-A is represented at the bottom left. **(B)** Evolution of the mean radius of gyration $\langle R_{\mathrm{g}} \rangle_{\mathrm{w}}$ as a function of time. **(C)** Plot of $\langle N \rangle_{\mathrm{w}}$ as a function of time. For **(A)**, **(B)**, and **(C)**, the color code for the ρ values is the same. **(D)** Proportion of larger-than-60-nm particles as a function of time estimated from cryoTEM micrographs at $\rho = 3.4$ (up). CryoTEM micrographs of capsids

assembled after 24 h (bottom). Scale bars are 50 nm. In all cases, the subunit concentration was 30 μM and $\rho = 3.4$.

Figure 3A shows SAXS patterns at equilibrium for capsid assembly in the presence of a varying amount of CAM-A. Compared to the patterns in the absence of modulator, almost all the oscillations were strongly damped, and the first minima were shifted to much smaller values of q. The curves also showed that I_0 increased with ρ, as was the case with CAM-E. Likewise, all SAXS curves were superimposed at high q-values, even in the presence of high ratios of CAM-A (Figure 3A). Thus, the local organization of subunits were conserved in the presence of any of the two modulators, regardless of the modulator-to-subunit molar ratio.

Figures 3B and 3C depict the time evolutions of $\langle R_\mathrm{g} \rangle_\mathrm{w}$ and $\langle N \rangle_\mathrm{w}$ during capsid assembly in the presence of CAM-A with $\rho = 0.34$ and 3.4. Clearly, the presence of CAM-A accelerated the assembly process. At the end of the kinetics, the mean radius of gyration $\langle R_\mathrm{g} \rangle_\mathrm{w}$ was about 176 Å and 251 Å, and the mean aggregation number $\langle N \rangle_\mathrm{w}^\infty$ was about 128 and 253, with $\rho = 0.34$ and 3.4, respectively, which confirmed that the final capsids were larger than native $T = 4$ capsids. These two quantities kept rising and did not seem to reach a plateau after about 15 min, indicating that the objects were still growing and/or self-organizing.

We investigated the morphologies of the objects by cryoTEM over the course of the assembly in the presence of CAM-A at $\rho = 3.4$. Throughout a 24-hours period, several micrographs were collected at different time points. Figure 3D gives the proportion of larger-than-60-nm particles observed at these time points. The proportion of large particles increased steadily and then stabilized after 5 h. The micrographs of Figure 3D obtained after 24 h, reveal elongated and facetted capsids, with a high shape variability, and many of them were uncapped. No spherical capsids other than a few $T = 4$ ones were seen.

Both modulators strongly increased the subunit binding energy.

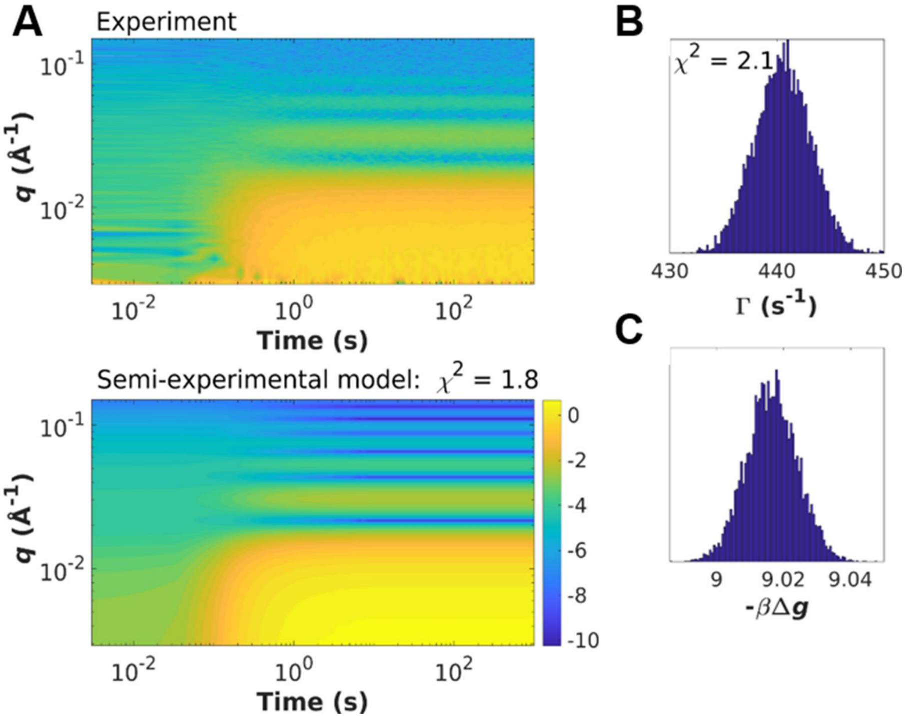

Figure 4. Kinetic modelling of TR-SAXS data without modulator. (**A**) Comparison between experimental and semi-experimental scattering intensities for capsid assembly. Intensities are in absolute units (cm^{-1}) and plotted in logarithmic scale. (**B**), (**C**) Distribution of parameters Γ and $-\beta\Delta g$ entering 'model A' (Eqs. 1-3) determined by Bayesian inference on experimental data. Subunit concentration was 30 µM.

In order to get a quantitative insight into the effects of modulators on capsid assembly, we carried out kinetic modeling of our TR-SAXS measurements. Icosahedral capsid assembly can be regarded as a spherical cap growing by sequential additions of free subunits until a shell is completed. In that simple picture, it is often reasonable to consider only free subunits and complete capsids at each time point of the process, the intermediate objects being not sufficiently long-lived to accumulate in detectable amounts. Indeed, singular value decomposition (SVD) analysis confirmed that the scattering intensities could be reconstructed with only two singular values within noise level (Figure S8). The scattering intensities $I(t,q)$ during the capsid assembly can then be reduced to:

$$I(t,q) \propto \phi(t)NP_N(q) + [1-\phi(t)]P_1(q) \tag{1}$$

where $\phi(t)$ denotes the fraction of subunits in capsids, while $P_1(q)$ and $P_N(q)$ are the form factors of subunits and assembled capsids made up of N subunits, respectively. Note that $P_1(0) = P_N(0) = 1$. Figure 4A shows a comparison between experimental and semi-experimental scattering intensities for capsid assembly without modulator. The semi-experimental intensities were computed with the two-state assumption (Eq. 1), in which the form factors were obtained from the crystal structures of a subunit and a $T = 4$ capsid. The semi-experimental fractions $\phi_{\mathrm{se}}(t)$ were determined by fitting the experimental intensities in the least-squares sense with bound constraints ($0 \leq \phi_{\mathrm{se}}(t) \leq 1$; see Methods for details). The semi-experimental model succeeded in reproducing the main features of the experimental data, which was quantitatively supported by a goodness-of-fit parameter χ^2 of 1.8.

The self-assembly of empty capsids is akin to a thermodynamic phase transition, in which free subunits are suddenly brought to a nonequilibrium state after a so-called quench, and subsequently relax towards assembled capsids. In our experiments, the quench was carried out by raising the salinity and adding modulators via rapid mixing. A kinetic theory adapted to quenched capsid assembly was proposed by van der Schoot and Zandi.[34] The so-called 'model A' is based on the relaxation of a non-conserved quantity, here ϕ, at a phenomenological rate Γ assumed constant, until an equilibrium value that minimizes the Helmholtz free energy of the system is reached. The kinetic equation reads:

$$\frac{d\phi}{dt} = -\Gamma c \left[-\ln\left(\frac{1-\phi}{1-\phi_\infty}\right) + \frac{1}{N}\ln\left(\frac{\phi}{\phi_\infty}\right)\right] \tag{2}$$

where c stands for the dimensionless concentration of subunits before quench and ϕ_∞ is the fraction of subunits in capsids at equilibrium. The latter is related to the subunit binding energy Δg through

$$\beta\Delta g = \ln[(1-\phi_\infty)c] - \frac{1}{N}\ln\left(\frac{\phi_\infty}{N}c\right) \tag{3}$$

with $1/\beta = k_\mathrm{B}T$, k_B being the Boltzmann constant and T the temperature. ‘Model A’ contained solely two free parameters, *i.e.*, the reaction rate Γ, which was considered as an invariant of time and of c, and the subunit binding free energy $-\beta\Delta g$ in $k_\mathrm{B}T$ units, which varied with the amount of modulator. Figure 4B depicts the distribution of the two parameters determined by a Bayesian approach on TR-SAXS data for capsid assembly without modulator. The form factors $P_1(q)$ and $P_N(q)$ ($N = 120$) were calculated from the crystal structures. The mean value of Γ was found to be close to 440 s^{-1}, that of $-\beta\Delta g$ was around 9.0, and the standard deviation in both cases was weak. χ^2 calculated with the mean values of Γ and $-\beta\Delta g$ was 2.1, which is a good score for so minimalistic a model.

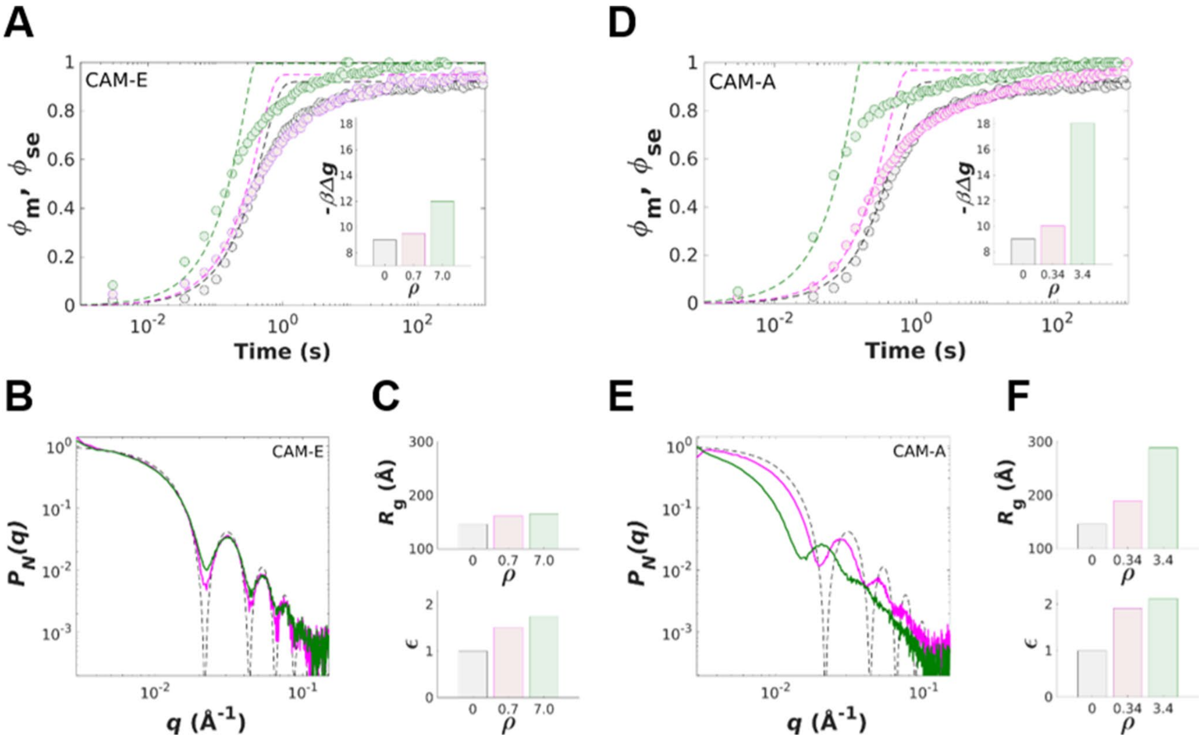

Figure 5. Kinetic modelling of TR-SAXS data in the presence of modulators. (**A)** Fraction of subunits in capsids ϕ versus time for capsid assembly in the presence of CAM-E at various modulator-to-subunit molar ratios ρ: 0 (black), 0.7 (magenta) and 7.0 (green). Discs are the fractions ϕ_{se} obtained with a semi-experimental model (Eq. 1; see Methods for details) while dashed lines represent the fractions ϕ_{m} estimated through 'model A' (Eqs. 1-3) with $\Gamma = 440$ s^{-1}. The inset gives the subunit binding energies $-\beta\Delta g$ used in 'model A'. (**B)** Form factors $P_N(q)$ used in the kinetic model for capsid assembly with CAM-E. The dashed black line was calculated from the crystal structure of a $T = 4$ capsid. (**C)** Radius of gyration R_{g} and axial ratio ϵ corresponding to the form factors shown in **(B)**. **(D)** Same as (**A**) but in the presence of CAM-A at ratios ρ of 0 (black), 0.34 (magenta) and 3.4 (green). **(E)** Form factors used in 'model A' with CAM-A. **(F)** R_{g} and ϵ corresponding to the form factors shown in **E**. In all cases, the subunit concentration was 30 μM.

In the presence of modulators, modelling was performed by fixing Γ to 440 s^{-1}, $P_1(q)$ was calculated from the crystal structure of a subunit, and only $-\beta\Delta g$ and $P_N(q)$ were fitting parameters. Figure 5A shows the fraction of subunits in capsids ϕ_{se} and ϕ_{m} obtained with either a semi-experimental model (discs) or 'model A' (dashed lines), respectively, for different ratios ρ of CAM-E. The goodness-of-fit parameter χ^2 of 'model A' was between 1.6 and 3.2. Like capsid assembly without modulator, the two-state assumption was supported by SVD analysis for capsid assembly in the presence of CAM-E (Figure S9). It can be seen that $-\beta\Delta g$ increased with ρ, starting from 9.0 in the absence of modulator up to 12 for $\rho = 7.0$. The extracted form factors of the final capsids (Figure 5B) confirmed the formation of structures slightly larger than $T = 4$ capsids as evidenced by the radius of gyration R_{g} (~165 Å versus 146 Å for a native capsid; Figure 5C), and more elongated as demonstrated by the axial ratio ϵ exceeding 1.0 (Figure 5C; see Methods for the definition of ϵ). These findings were consistent with the cryoTEM reconstructions described earlier. Likewise, we can see on Figure 5D that CAM-A dramatically strengthened the binding between subunits with $-\beta\Delta g$ reaching 18 for $\rho = 3.4$. χ^2 ranged from 2.1 to 4.4, the latter value being obtained at $\rho = 0.34$. This deviation from experimental intensities was ascribed to the coexistence of more than two long-lived species in solution as revealed by SVD analysis (Figure S10). However, because 'model A' was mainly adjusted with the early stage of the assembly, during which intermediate species were still in low concentration, the estimate of the subunit binding energy should be reliable. The final capsids were much larger ($R_{\mathrm{g}} = 289$ Å at $\rho = 3.4$; Figures 5E,F) than in the presence of CAM-E, and they were also even more elongated, with ϵ exceeding 2 (Figure 5F at the highest value of ρ. Once again, these results supported well the observations made by cryoTEM (see above).

Capsid morphologies are closely related to their elastic properties.

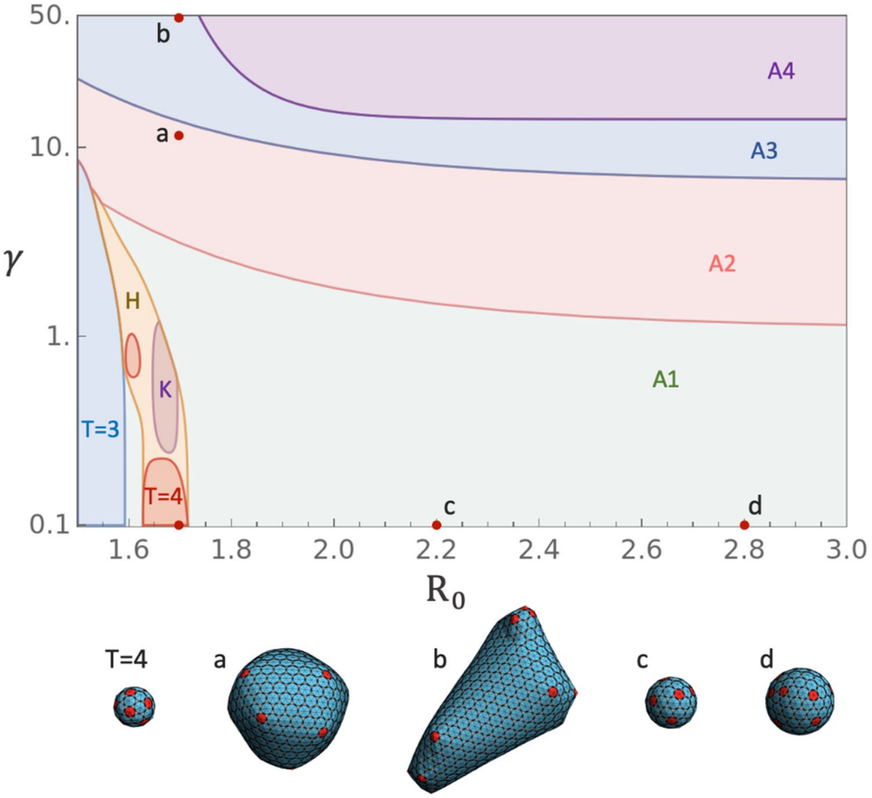

Figure 6. Phase diagram of capsid morphologies calculated by coarse-grained simulations as a function of spontaneous radius of curvature R_0 and γ. At small $R_0 \simeq 1.7$, the complete shells had icosahedral symmetry with triangulation numbers $T = 3$ and $T = 4$, as illustrated on the bottom of the phase diagram. The H shape surrounded the $T = 4$ region and consisted of a rotated $T = 4$ shell with the same number of triangular subunits (80) and the same radius. Inside the H region, we also observed the K shape, which had less triangular subunits (78) and a smaller radius. A1 to A4 denote regions with different ranges of asphericity: 0 to 0.0002 for A1; 0.0002 to 0.01 for A2; 0.01 to 0.1 for A3; and above 0.1 for A4. Below the phase diagram are representative shapes at different points (R_0, γ) in the phase space.

Figure 6 shows coarse-grained simulations of the complete capsid morphologies calculated as a function of the spontaneous radius of curvature R_0 and $\gamma = k_\mathrm{s} l_0^2 / k_\mathrm{b}$, where l_0 stands for the equilibrium bond length of each subunit, whereas k_s and k_b are the stretching and bending moduli between two subunits, respectively. Note that it is possible to map γ in our discrete model to the Föppl-von-Kármán number (FvK) defined as YR^2/κ in the continuum limit, with Y the Young modulus, R the radius of curvature and κ the bending rigidity (Figure S11). The phase diagram clearly shows that the capsids became more elongated (increasing asphericity) as γ was raised (see also a more comprehensive phase diagram on Figure S12). Thus, by fixing R_0 to 1.7, we could reproduce the native $T = 4$ capsids, the larger and slightly elongated capsids obtained with CAM-E (see representative shape **a** on Figure 6) and the strongly deformed, aberrant capsids formed in the presence of CAM-A (see representative shape **b** on Figure 6) by tuning γ from 0.1 to 50. Increasing R_0 at low γ-values yielded spherical shells with larger radii, but not deformed shells as observed in the presence of modulators. Therefore, increasing γ was a prerequisite to form aspherical capsids, and since we did not observe any spherical object larger than the native $T = 4$ capsid, we suggest that modulators influence mostly γ, or equivalently, the FvK number. It should be mentioned that the values of γ in our discrete model were $\gamma \simeq 1$ for $T = 4$ and $\gamma \simeq 5$ for H capsids, approximately corresponding to FvK $\simeq 20$ and FvK $\simeq 100$ in the continuum model, respectively. These values are small compared to those reported in the literature for HBV (FvK $\simeq$ 400) using thin-shell approximation.[35] The discrepancy could be due to the limitations employed in the thin-shell theory. We note that recent all-atom simulations revealed that HBV shells are very flexible[36]; thus, the low value of γ seems to be one of the essential features of HBV Core proteins, which need to be flexible and capable of a conformational switch in order to assemble to two different T numbers.

Discussion

Chronic infections by Hepatitis B virus are a major public health burden and drugs allowing HBV cure are being intensively developed. Among these, CAMs have grown into both promising compounds, a dozen of which are currently in clinical trials, and illuminating probes of the molecular steps of the viral cycle involving HBV Core.[27,37] Indeed, the recently adopted nomenclature of CAM-E and CAM-A[27] reflects the results of years of biochemical, biophysical and virological experiments that suggested that most CAMs paradoxically seem to promote capsid assembly, this nomenclature was also related to the theory that triggering (empty) capsid assembly at the wrong time and space was severely detrimental to the virus.[38] Here we show that capsids of the HBV Core NTD assembled in the presence of a prototypical CAM-E, JNJ-632, are actually structurally distinct from regular $T = 4$ capsids formed under the same conditions. With CAM-E, they depart from icosahedral symmetry and normal stoichiometry, forming larger (> 120 dimeric subunits) and ellipsoidal capsids even at a concentration where assembly is not detectably accelerated (modulator-to-subunit molar ratio $\rho = 0.7$). Thus, there could be modes of actions of CAM beyond accelerating assembly even for CAM-E. All CAMs seem to bind at the same interface between Core dimers as seen by X-ray crystallography.[39,40] This interface is present in 240 copies in regular assembled $T = 4$ capsids. Here our SAXS data show that local organization is not drastically changed by the CAMs in any of the assemblies formed. Thus, the number of potential CAM sites likely remains two per dimer and below $\rho = 2$, not all potential binding sites can be occupied by CAMs. At CAM-E $\rho = 7.0$, alterations of the icosahedral structure are more pronounced, but the objects assembled seem to still be closed capsids and their sizes plateau at ~160 subunits vs. 120 in regular capsids. This is in sharp contrast to the behavior in assembly with

CAM-A. Here we confirm and extend the previous reports that CAM-A induces grossly abnormal assembly (see below).

The kinetic modelling was carried out using ‘model A’, but in certain circumstances, capsid assembly can be described by the classical nucleation theory.[41,42] The latter theory is well suited to situations where the energy barrier of the formation of closed shells is high with respect to the thermal energy. This mainly occurs when the subunit concentration is near the critical aggregation concentration that itself depends on the subunit binding energy through $c_* \sim \exp(\beta \Delta g)$.[43] We assessed the classical nucleation theory and observed that the values of $-\beta \Delta g$ required to reproduce the fractions of subunits in capsids during the early time steps, were nearly the same as those inferred from ‘model A’ (see Supporting Text and Figure S13). However, and quite importantly, in all cases, the resulting height of the energy barrier at the beginning of capsid assembly was lower than the thermal energy (see Supporting Text), and consequently, the classical nucleation theory could not be applied here. In the reported experiments, capsid assembly was triggered by a quench, namely, a rapid change of physicochemical conditions bringing free subunits out of equilibrium, and ‘model A’ was precisely worked out to account for the dynamics of phase transition in quenched systems. Notice that without modulator, we found $-\beta \Delta g = 9.0$, a value close to 8.2 as estimated by Asor and coworkers[18] under the same ionic conditions, but with a temperature of 25°C, which may account for the slight difference.

The striking effects of CAM-A cannot be solely ascribed to an enhanced interaction between subunits mediated by the modulator due to its hydrophobic nature. Although CAM-E at $\rho = 0.7$ and CAM-A at $\rho = 0.34$ induced similar subunit binding energies ($-\beta \Delta g \simeq 10$), the capsid morphologies were significantly different, capsids in the presence of CAM-A being larger and more elongated. Additionally, assembly with CAM-A exhibited a long timescale of several hours during which the objects were still growing and most of them eventually remained uncapped. This is characteristic of a kinetic trap:[43] large, uncapped objects are formed within too short a time, thus depleting free subunits, which are thereby not in sufficient amount to form closed-up structures. Additional subunits must be released from small objects to allow the large ones to keep growing and become closed, a process reminiscent of the Ostwald ripening in emulsions.

At the microscopic level, we propose that the modulators have a twofold mechanism of action. Firstly, they enhance the binding energy between dimeric subunit owing to their hydrophobic nature. The higher subunit binding energy $-\beta\Delta g$ translates into increased stretching energies at the capsid scale, and subsequently, higher FvK numbers. A second effect may arise from their steric hindrance that can disturb the binding between adjacent subunits, since modulators sit at the surface of contact. As a result, the subunits in contact gain more flexibility to deviate around their preferred angle of curvature, thus lowering the bending rigidity of the capsid, which contributes to further increase the FvK number. It is noteworthy that, because of the slight differences between the four quasi-equivalent conformations 'A', 'B', 'C' and 'D' of Core in the $T = 4$ capsid, there are four variations in the aforementioned dimer-dimer interface depending on whether it is A-to-A (pentameric contact), or B-to-C, C-to-D or D-to-B (three hexameric contacts). Different CAMs preferentially bind to different of these interface variations, giving a rationale as to how they may modify capsid curvature.[39]

We show here that JNJ-632, classified as a CAM-E, actually misdirects capsid assembly, albeit in a subtler way than the CAM-A BAY 41-4109. This has just been also reported by Lecoq and coworkers,[44] who showed that CAM-E, including JNJ-632, actually homogenize the four different dimer-dimer contacts, albeit to a lesser extent than CAM-A. Although negative-staining electron microscopy could not distinguish capsids reassembled with and without JNJ-632 at the mesoscopic level, solid-state nuclear magnetic resonance clearly showed that JNJ-632 modified dimer-dimer contacts at the molecular level, almost to the same extent as one CAM-A. Here we also demonstrate by TR-SAXS that JNJ-632 increases the average stoichiometry of assembly above 120 dimers, and by cryoTEM that the larger capsids are detectably ellipsoidal.

Conclusions

The purpose of this study was to elucidate how two classes of modulators disturb the assembly pathway and alter the energetics of HBV capsids. We bring light to several important features of CAMs: *e.g.,* some CAM-E actually can misdirect assembly; CAMs modulate, but do not prevent or grossly modify, the local arrangement of subunits; CAMs increase the subunit binding energy; and CAM binding also alters elastic energy. These molecules promote capsid assembly by increasing the amount of buried hydrophobic surface area at the interface between dimers, thus raising subunit binding energy. However, some CAMs overfill this hydrophobic pocket, causing the formation of aberrant, much longer, and distorted capsids. Here, we demonstrate that it is possible to simulate the formation of such objects by increasing the Föppl-von-Kármán number (FvK). This number is proportional to γ, which in turn is proportional to the ratio of stretching to bending energy. Thus, by lowering the bending energy while leaving the stretching energy unchanged (thereby increasing the FvK value), dimers could more easily deviate from their spontaneous curvature and form large, flat, non-curved structures such as those observed during assembly in the presence of type A modulators. Class E modulators, which do not lead to the formation of this type of structure, would therefore have a less substantial impact on this bending energy. We believe that further quantitative and physical investigations are necessary to understand the mechanisms of action of modulators. Although this research alone cannot be used to design antiviral molecules, it does enable us to carry out fingerprinting. The techniques used here provide both clarification and insights on the mode of action of these molecules. These techniques will therefore enable us to characterize, from a structural and molecular point of view, the mode of action of synthesized CAMs with observable antiviral activity after screening on cellular systems. Since the HBV Core protein can self-assemble *in vitro*, it constitutes a good cargo protein engineering system for the transport and delivery of large and small molecules. The ability to chemically trigger the dissociation of virus-like particles derived from HBV Core proteins has already been demonstrated[45]. Therefore, the assembly of highly stable cargo capsids utilized to deliver therapeutic molecules could be regulated by these assembly modulators.

Materials and Methods

Sample preparation

Expression of Cp149 capsids. Hepatitis B virus capsids were expressed in *E. coli* as previously described.[11,20] Briefly, *E. coli* BL21*Codon Plus cells were transfected with the pRSF-T7-HBc149 opt plasmid and grown at 37°C overnight on an LB-agar plate containing 50 mg.L^{-1} of kanamycin and 34 mg.L^{-1} of chloramphenicol. A single colony was then inoculated into 2×10 mL of LB medium and grown at 37°C overnight. After that, it was diluted into 2×500 mL of LB medium containing the same concentration of antibiotics and incubated at 37°C for 5 h. When OD_{600} reached 0.8, the induction was done by adding 1 mM IPTG and the bacterial cells were grown at 25°C overnight.

Capsid purification. After the expression of Cp149 capsids, bacterial cells were pelleted by centrifugation and resuspended in lysis buffer containing 300 mM NaCl, 2 mM DTT, 50 mM Tris-HCl pH 7.5. 1 g.L^{-1} of lysozyme, 0.5% Triton-X-100 and protease inhibitor cocktail were added to the cells and mixed on ice for 1 h. Then 4 μL of commercial Pierce nuclease were added to the solution and mixed at room temperature for 45 min. The cells were lysed by sonication and centrifuged for 45 min. The supernatant was deposited on a 10% to 60% sucrose gradient and centrifuged with a SW-32 Ti rotor (Beckman, France) at 28,800 rpm for 3 h at 4°C. The fractions containing Cp149 capsids were identified by 16% SDS-PAGE gel and precipitated by slowly adding ammonium sulfate until 40% saturation was reached. After incubation for 1 h at room temperature and 2 h on ice, the solution was centrifuged. The pellet was resuspended in a purification buffer composed of 5% sucrose, 1 mM DTT, 50 mM Tris HCl, pH 7.5 and centrifuged to remove insoluble pellet. The supernatant containing soluble Cp149 capsids was stored at 4°C.

Subunit purification. The purification of Cp149 subunits was adapted from a protocol developed previously.[20] Cp149 capsids were dialyzed at 4°C in the disassembly buffer composed of 1 mM DTT, 50 mM CHES pH 9 and completely dissociated by adding solid urea up to 3 M. Following the dissociation, the subunits were then purified by size exclusion chromatography (SEC) using a Superdex S200 increase GL pre-equilibrated with the disassembly buffer. Fractions where dimers had been identified were dialyzed in the assembly buffer containing 500 mM NaCl, 1 mM DTT, 50 mM Tris-HCl pH 7.5, Assembled capsids were separated from dimers by size exclusion chromatography. Finally, an additional dissociation-purification step was added to remove inactive protein. Subunits could be stored at -80°C for several months without any activity loss.

Time-resolved and static small-angle X-ray scattering

Prior to any measurement, Cp149 subunits were incubated for 1 h with 3 M of solid urea and dialyzed in a buffer containing 1 mM DTT, 50 mM Tris-HCl pH 7.5 for 2 h.

Time-resolved small-angle X-ray scattering experiments were carried out at the SWING beamline of the SOLEIL synchrotron facility. The sample-to-detector distance was set to 2.5 or 5.0 m which provided momentum transfers q ranging from 2.92×10^{-3} to 0.463 $Å^{-1}$ or 1.46×10^{-3} to 0.232 $Å^{-1}$. Assembly was triggered with a stopped-flow mixer (BioLogic SFM-400) by mixing subunits, the assembly buffer composed of ammonium acetate and 50 mM Tris-HCl pH 7.5, and the modulators JNJ-632 and Bay 41-4109 at different modulators-to-subunit ratio, into a 1 mm diameter quartz capillary, at 37°C. Beam exposure time was set to 20 ms.

The samples for static small-angle X-ray scattering measurements were prepared by manually mixing subunits with 150 mM ammonium acetate in the presence of modulators, incubated at 37°C for 1 hour and stored for 2 to 3 days at 4°C. During measurements, the samples were injected into a 1.5 mm diameter quartz capillary cell using an autosampler and the sample-to-detector distance was set to 2.5 m. The temperature was maintained at 37°C with a heat bath.

The two-dimensional images were radially averaged to obtain the one-dimensional scattering profiles using Foxtrot,[46] and intensities were converted into absolute units after subtraction of the contribution of buffer solutions. The forward scattering intensity $I_0 = I(q \to 0)$ and the radius of gyration R_g – or the mean (weight-averaged) radius of gyration $\langle R_\mathrm{g} \rangle_\mathrm{w}$ for mixtures – were determined with AUTORG and PRIMUS from the ATSAS suite[47] with the condition $q\langle R_\mathrm{g} \rangle_\mathrm{w} < 1.3$ defining the Guinier region. I_0 was used to estimate the mean aggregation number $\langle N \rangle_\mathrm{w}$ corresponding to the weight-averaged number of subunits per object:

$$\langle N \rangle_\mathrm{w}(t) = \frac{c^\mathrm{S}}{c} \frac{I_0(t)}{I_0^\mathrm{S}} \quad (4)$$

where I_0^S stands for the forward scattering intensity of subunits measured separately at a molar concentration c^S, while c is the initial molar concentration of subunits in an assembly experiment.

Kinetic modelling

The form factors of subunits and $T = 4$ capsids were computed with CRYSOL[48] from the crystal structures (Protein Data Bank code 1QGT). The subunit concentration before quench c in Eqs. 2 and 3 was taken by normalizing the subunit molar concentration to the reference state $c_{\text{ref}} = 21.7\ \text{mM}$, this value being estimated from the volume occupied by a single subunit. The semi-experimental intensities of Figure 4A were obtained by fitting Eq. 1 to experimental scattering intensities with the fraction of subunits in capsids $\phi_{\text{se}}(t)$ as a fitting parameter for each time point, with bound constraints $0 \leq \phi_{\text{se}}(\text{t}) \leq 1$. The distributions of Γ and $-\beta\Delta g$ in Figures 4B,C were computed with a Hamiltonian Markov chain Monte Carlo algorithm: prior distributions were set to be uniform and the likelihood was chosen as gaussian using the experimental uncertainties of the scattering intensities. The form factors were fixed to their calculated values, and at each Monte Carlo step, $\phi(t)$ was calculated by solving Eqs. 2 and 3. Afterwards, the modelled intensities were computed from Eq. 1 and compared with the experimental intensities through the likelihood. Sampling was carried out using Matlab® with 10,000 uncorrelated Monte Carlo steps.

Throughout the text, the goodness-of-fit parameter is defined as

$$\chi^2 = \frac{1}{(N_{\text{data}} - N_{\text{param}})} \sum_{i,j} \left[\frac{I(q_i, t_j) - I_{\text{m}}(q_i, t_j)}{\sigma_{i,j}} \right]^2 \tag{5}$$

where N_{data} is the number of data points, N_{param} the number of parameters including the data points entering the form factors, I_{m} the modelled scattering intensities, and σ the experimental uncertainties.

The subunit binding energies in the presence of modulators were estimated as follows: firstly, a value of $-\beta\Delta g$, and subsequently ϕ_∞ through Eq. 3, was chosen. The evolution of $\phi_{\text{m}}(t)$ – the index ‘m’ stands for ‘modelled’ – was calculated by solving Eq. 2, Γ being set to 440 s^{-1}. The traces are plotted as dashed lines on Figures 5A,D and are referred to as ‘model A’. With the assumption that the ten last experimental scattering curves were measured near the equilibrium, their average was used to estimate the form factor of the final capsids $P_N(q)$ via Eq. 1 after plugging ϕ_∞ into $\phi(t)$, while $P_1(q)$ was calculated from the crystal structure of subunits using CRYSOL.[48] $P_N(q)$ are plotted on Figures 5B,E. The semi-experimental fractions of subunits in capsids $\phi_{\mathrm{se}}(t)$ were inferred from Eq. 1 applied on experimental scattering intensities through a least-squares minimization with bound constraints ($0 \leq \phi_{\mathrm{se}}(t) \leq 1$) at each time point. These fractions are displayed on Figures 5A,D as discs. $-\beta\Delta g$ was adjusted manually in such a way that $\phi_{\mathrm{m}}(t)$ and $\phi_{\mathrm{se}}(t)$ coincided over the short time scales, *i.e.*, typically for time points verifying $\phi_{\mathrm{m}}(t) \sim \phi_{\mathrm{se}}(t) < 0.5$. The number of subunits in capsid N was determined by extrapolation of $NP_N(q)$ to the low q-values and reinjected into Eqs. 2 and 3 in such a way that ‘model A’ was self-consistent.

The method for the determination of the radius of gyration and of the axial ratio on Figure 5 was proposed by Roig-Solvas and coworkers[49] as an extension of the Guinier approximation. A Taylor expansion of the scattering intensity of an ellipsoid with semi-axes R, R, ϵR – ϵ being the axial ratio –, about $q = 0$, yields

$$I(q) = I_0\left[1 - \tfrac{1}{3}R_{\mathrm{g}}^2 q^2 + \tfrac{1}{21}\left(R_{\mathrm{g}}^4 + \tfrac{1}{5}A_{\mathrm{F}}^2\right) q^4 + \mathcal{O}(q^6)\right] \tag{6}$$

with R_{g} the radius of gyration and A_{F} the anisotropy factor. For a prolate ellipsoid ($\epsilon \geq 1$), the axial ratio is related to the anisotropy factor through $\epsilon = \sqrt{\left(1 + A_{\mathrm{F}}/R_{\mathrm{g}}\right)/\left(1 - A_{\mathrm{F}}/\left(2R_{\mathrm{g}}\right)\right)}$. The axial ratio is 1.0 for the spherical case and increases as the structure becomes elongated.

Accordingly, $P_N(q)$ was fitted with an 8^{th}-order polynomial over the range $1 \leq qR_\text{g} \leq 3$, and both R_g and ϵ were determined by identification of the three first polynomial coefficients.

Cryotransmission electron microscopy and image processing

For the cryoTEM experiments, the assembly was triggered by mixing manually 30 µM of subunits in a buffer containing 1 mM DTT, 20 mM Tris-HCl pH 7.5, in the presence or not of modulators (20 or 200 µM – *i.e.*, $\rho = 0.7$ or 7.0 – for CAM-E, and 100 µM – *i.e.*, $\rho = 3.4$ – for CAM-A) and 200 mM ammonium acetate. The latter salt concentration was found to be essential for the viability of capsids embedded in thin ice after cryofixation. Capsids assembled without modulator or with CAM-E were incubated at 37°C for 1 h and concentrated ten times to allow a faster collection of a large number of single particles during image processing. Capsids assembled with CAM-A were incubated at 37°C for 24 h with no further concentration. 4 μL of solutions were deposited onto glow-discharged Quantifoil holey-carbon grids (R2/2 or R2/1). The grids were blotted with filter paper for 2 s before automated plunging into liquid ethane cooled down by liquid nitrogen using FEI Vitroblot. The grids were stored in liquid nitrogen until use. The frozen samples were transferred into a Gatan 626 cryo-holder and imaging was carried out at –180°C on a JEOL JEM-2010 microscope equipped with a 200-kV field emission gun. The samples were imaged with a magnification of ×50,000 using a minimal dose system. Images were recorded with a Gatan Ultrascan 4K CCD camera at 2 μm of nominal defocus.

CryoTEM particle picking, classification and reconstruction were carried out using standard protocols with the cryoSPARC software.[50] After CTF correction, a first 2D classification was performed using manually picked and extracted particles. Then, from the selected 2D classes, the particles were verified manually and extracted from the micrographs, and a second round of 2D classification was performed. The 3D maps were obtained after several steps of *ab initio* reconstruction and homogeneous or heterogeneous refinement jobs.

Sample	Number of micrographs	Number of particles after 2D classification	Number of particles for 3D maps
No modulator	39	3,478	3,478
CAM-E ($\rho = 0.7$)	42	3,912	3,912
CAM-E ($\rho = 7.0$)	54	3,010	3,010
CAM-A ($\rho = 3.4$)	95	—	—

Table 1. Number of micrographs and single particles used for image processing.

Coarse-grained simulations

The assembly simulations were performed through triangulation growth model,[51–53] where we coarse-grained the subunits as the edges of each triangle, as shown in Figure 7. Since the subunit concentration was low, at each step of the shell growth, we assumed that the subunits had enough time to find an optimal position to maximize its neighbors due to the gain of hydrophobic interaction. Meanwhile, the energy of the shell was minimized at each step, which can be written as a sum of stretching energy and bending energy

$$E_{\mathrm{shell}} = E_{\mathrm{s}} + E_{\mathrm{b}} = \frac{1}{2} k_{\mathrm{s}} \sum_{\mathrm{i}} (l_i - l_0)^2 + k_{\mathrm{b}} \sum_i [1 - \cos(\theta_i - \theta_0)] \tag{7}$$

where k_{s} and k_{b} are the stretching and bending moduli, l_i and θ_i are the length and dihedral angles of bond i, and l_0, θ_0 are the equilibrium bond length and preferred angle. The shell energy is a discretization of continuum energy[54]

$$E = \frac{1}{2}\int dA \frac{Y}{1+\nu}\left(u_{ij}^2 + \frac{\nu}{1-\nu}u_{kk}^2\right) + \int dA \left[\frac{1}{2}\kappa(H - H_0)^2 + \kappa_g K\right] \quad (8)$$

where u_{ij}, Y and ν are the strain tensor, 2D Young's modulus, and Poisson's ratio, respectively. H, H_0, K, κ and κ_g are mean curvature, spontaneous curvature, Gaussian curvature, bending rigidity, and Gaussian rigidity, while dA is the area element. The asphericity is defined as

$$\mathcal{A} = \frac{1}{N_v}\sum_{i=1}^{N_v} \frac{(R_i - \langle R \rangle)^2}{\langle R \rangle^2} \quad (9)$$

with N_v the vertex number, R_i the radial distance of vertex i, and $\langle R \rangle$ the mean radius.

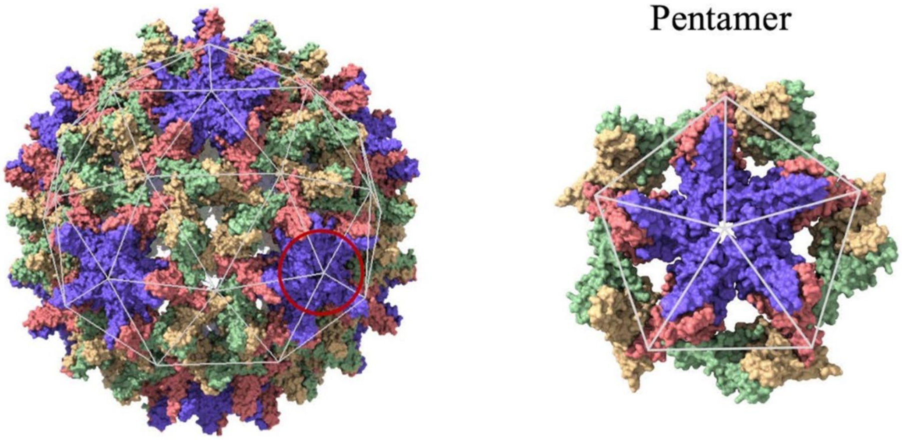

Figure 7. Coarse-grained model. Illustration of a $T = 4$ HBV capsid along with triangular subunits used in the simulations.

Acknowledgments

We thank S. Fieulaine and R. Ruedas for their help during sample preparation and TR-SAXS experiments. K.K., S.B. and G.T. acknowledge financial support from the Agence Nationale de Recherche sur le Sida et les hépatites virales (contracts ANRS ECTZ117006 and ECTZ123887). K.K, S.B. and G.T. also acknowledge the synchrotron SOLEIL for allocating beamtime and A. Thureau for his technical assistance at the SWING beamline. The electron microscopy imaging is supported by "Investissements d'Avenir" LabEx PALM (ANR-10-LABX-0039-PALM). S.L. and

R.Z. acknowledge support from NSF DMR-2131963 and the University of California Multicampus Research Programs and Initiatives (grant No. M21PR3267).

Supporting Information

The Supporting Information is available online free of charge. Supporting text detailing SVD analysis and classical nucleation theory; supporting figures showing SAXS patterns of subunits and assembled capsids in the presence or not of CAM-E, 3D reconstructions of capsids in the presence or not of CAM-E, SVD analyses with and without modulators, mapping on continuum FvK number, full phase diagram of simulated capsid morphologies, and modelling with the classical nucleation theory (PDF).

Author contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

References

1. Beasley, R. P. Hepatitis B virus. The major etiology of hepatocellular carcinoma. *Cancer* **61**, 1942–1956 (1988).
2. Taverniti, V. *et al.* Capsid Assembly Modulators as Antiviral Agents against HBV: Molecular Mechanisms and Clinical Perspectives. *J Clin Med* **11**, 1349 (2022).
3. Zlotnick, A. *et al.* Core protein: A pleiotropic keystone in the HBV lifecycle. *Antiviral Research* **121**, 82–93 (2015).
4. Birnbaum, F. & Nassal, M. Hepatitis B virus nucleocapsid assembly: primary structure requirements in the core protein. *J Virol* **64**, 3319–3330 (1990).
5. Venkatakrishnan, B. & Zlotnick, A. The Structural Biology of Hepatitis B Virus: Form and Function. *Annual Review of Virology* **3**, 429–451 (2016).
6. Nassal, M. The arginine-rich domain of the hepatitis B virus core protein is required for pregenome encapsidation and productive viral positive-strand DNA synthesis but not for virus assembly. *J Virol* **66**, 4107–4116 (1992).
7. Perlmutter, J. D. & Hagan, M. F. Mechanisms of Virus Assembly. *Annu Rev Phys Chem* **66**, 217–239 (2015).
8. Zlotnick, A., Johnson, J. M., Wingfield, P. W., Stahl, S. J. & Endres, D. A Theoretical Model Successfully Identifies Features of Hepatitis B Virus Capsid Assembly. *Biochemistry* **38**, 14644–14652 (1999).
9. Endres, D. & Zlotnick, A. Model-Based Analysis of Assembly Kinetics for Virus Capsids or Other Spherical Polymers. *Biophysical Journal* **83**, 1217–1230 (2002).

10. Oliver, R. C. *et al.* Assembly of Capsids from Hepatitis B Virus Core Protein Progresses through Highly Populated Intermediates in the Presence and Absence of RNA. *ACS Nano* **14**, 10226–10238 (2020).
11. Ceres, P. & Zlotnick, A. Weak Protein−Protein Interactions Are Sufficient To Drive Assembly of Hepatitis B Virus Capsids. *Biochemistry* **41**, 11525–11531 (2002).
12. Patterson, A., Zhao, Z., Waymire, E., Zlotnick, A. & Bothner, B. Dynamics of Hepatitis B Virus Capsid Protein Dimer Regulate Assembly through an Allosteric Network. *ACS Chem. Biol.* **15**, 2273–2280 (2020).
13. Garmann, R. F., Comas-Garcia, M., Gopal, A., Knobler, C. M. & Gelbart, W. M. The Assembly Pathway of an Icosahedral Single-Stranded RNA Virus Depends on the Strength of Inter-Subunit Attractions. *Journal of Molecular Biology* **426**, 1050–1060 (2014).
14. Narayanan, T., Wacklin, H., Konovalov, O. & Lund, R. Recent applications of synchrotron radiation and neutrons in the study of soft matter. *Crystallogr. Rev.* vol. 23 160–226 (2017).
15. Kler, S. *et al.* RNA Encapsidation by SV40-Derived Nanoparticles Follows a Rapid Two-State Mechanism. *J. Am. Chem. Soc.* **134**, 8823–8830 (2012).
16. Tresset, G. *et al.* Norovirus Capsid Proteins Self-Assemble through Biphasic Kinetics via Long-Lived Stave-like Intermediates. *J. Am. Chem. Soc.* **135**, 15373–15381 (2013).
17. Chevreuil, M. *et al.* Nonequilibrium self-assembly dynamics of icosahedral viral capsids packaging genome or polyelectrolyte. *Nat Commun* **9**, 3071 (2018).
18. Asor, R., Schlicksup, C. J., Zhao, Z., Zlotnick, A. & Raviv, U. Rapidly Forming Early Intermediate Structures Dictate the Pathway of Capsid Assembly. *J. Am. Chem. Soc.* **142**, 7868–7882 (2020).

19. Law-Hine, D., Zeghal, M., Bressanelli, S., Constantin, D. & Tresset, G. Identification of a major intermediate along the self-assembly pathway of an icosahedral viral capsid by using an analytical model of a spherical patch. *Soft Matter* **12**, 6728–6736 (2016).
20. Chevreuil, M. *et al.* Nonsymmetrical dynamics of the HBV capsid assembly and disassembly evidenced by their transient species. *J. Phys. Chem. B* **124**, 9987–9995 (2020).
21. Law-Hine, D. *et al.* Reconstruction of the disassembly pathway of an icosahedral viral capsid and shape determination of two successive intermediates. *J. Phys. Chem. Lett.* **6**, 3471–3476 (2015).
22. Burke, A. *et al.* Nanoparticle-Templated Self-Assembly of Viral Capsids Probed by Time-Resolved Absorbance Spectroscopy and X-Ray Scattering. *Phys. Rev. Applied* **10**, 054065 (2018).
23. Zoulim, F. New insight on hepatitis B virus persistence from the study of intrahepatic viral cccDNA. *Journal of Hepatology* **42**, 302–308 (2005).
24. Stray, S. J. *et al.* A heteroaryldihydropyrimidine activates and can misdirect hepatitis B virus capsid assembly. *Proceedings of the National Academy of Sciences* **102**, 8138–8143 (2005).
25. Berke, J. M. *et al.* Capsid Assembly Modulators Have a Dual Mechanism of Action in Primary Human Hepatocytes Infected with Hepatitis B Virus. *Antimicrob Agents Chemother* **61**, (2017).
26. Corcuera, A. *et al.* Novel non-heteroarylpyrimidine (HAP) capsid assembly modifiers have a different mode of action from HAPs in vitro. *Antiviral Research* **158**, 135–142 (2018).
27. Zoulim, F. *et al.* Nomenclature of HBV core protein-targeting antivirals. *Nat. Rev. Gastroenterol. Hepatol.* **19**, 748–750 (2022).

28. Zhou, Z. *et al.* Heteroaryldihydropyrimidine (HAP) and Sulfamoylbenzamide (SBA) Inhibit Hepatitis B Virus Replication by Different Molecular Mechanisms. *Sci Rep* **7**, 42374 (2017).

29. Campagna, M. R. *et al.* Sulfamoylbenzamide Derivatives Inhibit the Assembly of Hepatitis B Virus Nucleocapsids. *J Virol* **87**, 6931–6942 (2013).

30. Cole, A. G. Modulators of HBV capsid assembly as an approach to treating hepatitis B virus infection. *Current Opinion in Pharmacology* **30**, 131–137 (2016).

31. Vandyck, K. *et al.* Synthesis and Evaluation of N-Phenyl-3-sulfamoyl-benzamide Derivatives as Capsid Assembly Modulators Inhibiting Hepatitis B Virus (HBV). *J. Med. Chem.* **61**, 6247–6260 (2018).

32. Weber, O. *et al.* Inhibition of human hepatitis B virus (HBV) by a novel non-nucleosidic compound in a transgenic mouse model. *Antiviral Research* **54**, 69–78 (2002).

33. Stray, S. J. & Zlotnick, A. BAY 41-4109 has multiple effects on Hepatitis B virus capsid assembly. *Journal of Molecular Recognition* **19**, 542–548 (2006).

34. van der Schoot, P. & Zandi, R. Kinetic theory of virus capsid assembly. *Phys. Biol.* **4**, 296–304 (2007).

35. Roos, W. H., Bruinsma, R. & Wuite, G. J. L. Physical virology. *Nature Phys.* **6**, 733–743 (2010).

36. Hadden, J. A. *et al.* All-atom molecular dynamics of the HBV capsid reveals insights into biological function and cryo-EM resolution limits. *eLife* **7**, e32478 (2018).

37. Lahlali, T. *et al.* Novel Potent Capsid Assembly Modulators Regulate Multiple Steps of the Hepatitis B Virus Life Cycle. *Antimicrob Agents Chemother* **62**, e00835-18 (2018).

38. Zlotnick, A. & Mukhopadhyay, S. Virus assembly, allostery and antivirals. *Trends Microbiol.* **19**, 14–23 (2011).

39. Katen, S. P., Tan, Z., Chirapu, S. R., Finn, M. G. & Zlotnick, A. Assembly-Directed Antivirals Differentially Bind Quasiequivalent Pockets to Modify Hepatitis B Virus Capsid Tertiary and Quaternary Structure. *Structure* **21**, 1406–1416 (2013).

40. Klumpp, K. *et al.* High-resolution crystal structure of a hepatitis B virus replication inhibitor bound to the viral core protein. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 15196–15201 (2015).

41. Zandi, R., van der Schoot, P., Reguera, D., Kegel, W. & Reiss, H. Classical nucleation theory of virus capsids. *Biophys. J.* **90**, 1939–1948 (2006).

42. Timmermans, S. B. P. E. *et al.* The Dynamics of Viruslike Capsid Assembly and Disassembly. *J. Am. Chem. Soc.* **144**, 12608–12612 (2022).

43. Hagan, M. F. & Grason, G. M. Equilibrium mechanisms of self-limiting assembly. *Rev. Mod. Phys.* **93**, 025008 (2021).

44. Lecoq, L. *et al.* Molecular elucidation of drug-induced abnormal assemblies of the hepatitis B virus capsid protein by solid-state NMR. *Nat. Commun.* **14**, 471 (2023).

45. Starr, C. A. *et al.* Engineering Metastability into a Virus-like Particle to Enable Triggered Dissociation. *J. Am. Chem. Soc.* **145**, 2322–2331 (2023).

46. Thureau, A., Roblin, P. & Pérez, J. BioSAXS on the SWING beamline at Synchrotron SOLEIL. *J Appl Cryst* **54**, 1698–1710 (2021).

47. Franke, D. *et al.* ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Cryst.* **50**, 1212–1225 (2017).

48. Svergun, D. I., Barberato, C. & Koch, M. H. J. CRYSOL - a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* **28**, 768–773 (1995).

49. Roig-Solvas, B., Brooks, D. & Makowski, L. A direct approach to estimate the anisotropy of protein structures from small-angle X-ray scattering. *J. Appl. Crystallogr.* **52**, 274–283 (2019).
50. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
51. Wagner, J. & Zandi, R. The Robust Assembly of Small Symmetric Nanoshells. *Biophys. J.* **109**, 956–965 (2015).
52. Panahandeh, S., Li, S. & Zandi, R. The equilibrium structure of self-assembled protein nano-cages. *Nanoscale* **10**, 22802–22809 (2018).
53. Panahandeh, S. *et al.* How a Virus Circumvents Energy Barriers to Form Symmetric Shells. *ACS Nano* **14**, 3170–3180 (2020).
54. Seung, H. S. & Nelson, D. R. Defects in flexible membranes with crystalline order. *Phys. Rev. A* **38**, 1005–1018 (1988).