A pocket-factor–triggered conformational switch in the hepatitis B virus capsid

Lauriane Lecoq[a,1], Shishan Wang[a,1], Marie Dujardin[a], Peter Zimmermann[b], Leonard Schuster[b], Marie-Laure Fogeron[a], Mathilde Briday[a], Maarten Schledorn[c], Thomas Wiegand[c], Laura Cole[a], Roland Montserret[a], Stéphane Bressanelli[d], Beat H. Meier[c], Michael Nassal[b,2,3], and Anja Böckmann[a,2,3]

[a]Molecular Microbiology and Structural Biochemistry, Labex Ecofect, UMR 5086 CNRS/Université de Lyon, Lyon 69367, France; [b]Department of Medicine II/Molecular Biology, University Hospital Freiburg, Medical Center, University of Freiburg, Freiburg 79106, Germany; [c]Physical Chemistry, ETH Zurich, Zurich 8093, Switzerland; and [d]Institute for Integrative Biology of the Cell, Commissariat à l'énergie atomique, CNRS, Université Paris-Saclay, Gif sur Yvette Cedex 91198, France

Edited by Robert Tycko, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD, and approved February 4, 2021 (received for review November 6, 2020)

Viral hepatitis is growing into an epidemic illness, and it is urgent to neutralize the main culprit, hepatitis B virus (HBV), a small-enveloped retrotranscribing DNA virus. An intriguing observation in HB virion morphogenesis is that capsids with immature genomes are rarely enveloped and secreted. This prompted, in 1982, the postulate that a regulated conformation switch in the capsid triggers envelopment. Using solid-state NMR, we identified a stable alternative conformation of the capsid. The structural variations focus on the hydrophobic pocket of the core protein, a hot spot in capsid–envelope interactions. This structural switch is triggered by specific, high-affinity binding of a pocket factor. The conformational change induced by the binding is reminiscent of a maturation signal. This leads us to formulate the "synergistic double interaction" hypothesis, which explains the regulation of capsid envelopment and indicates a concept for therapeutic interference with HBV envelopment.

hepatitis B virus | solid-state NMR | hydrophobic pocket | Triton

More than 250 million people are chronically infected with hepatitis B virus (HBV), the major cause for terminal liver disease. Killing nearly one million people every year, the death toll of HBV rivals that of HIV (1); current therapies can rarely cure infection (2). HBV is a small-enveloped DNA virus (Fig. 1*A*, reviewed in ref. 3) with a 3.2 kb relaxed circular DNA (rcDNA) genome that encodes three sequence-related envelope or surface proteins (small [S], middle [M], and large [L]; Fig. 1*B*), the capsid-forming core protein (Cp), the multifunctional polymerase, and an epigenetic regulator of HBV transcription, HBx (reviewed in refs. 4–6, respectively). Cp (Fig. 1*C*) consists of an N-terminal assembly domain and an arginine-rich C-terminal domain (CTD) that, among other functions, is needed for RNA packaging (7, 8). Cp forms stable dimers, and 120 (or 90) copies of them assemble into capsids with T = 4 (or T = 3, minor class) icosahedral symmetry (9), including in heterologous expression systems. In the cytoplasm of hepatic cells, nucleocapsids assemble from Cp dimers and a complex of viral pregenomic RNA (pgRNA) and the polymerase. Inside the capsid, the polymerase reverse transcribes the pgRNA into single-stranded, minus-strand DNA and, subsequently, into partially double-stranded rcDNA (reviewed in ref. 10), assisted by dynamic changes in Cp's phosphorylation status (11, 12) (Fig. 1*D*). These mature capsids are then enveloped through, possibly synergistic, interactions (Fig. 1 *B* and *C*) of the S/M/L cytosolic loop (13, 14) and the preS matrix domain (12–16) with the Cp spike tip (15) and/or base (16), including a hydrophobic pocket at the dimer interface (17). For interaction with Cp, preS is oriented inside the envelope (Fig. 1*B*), while later on, it partly points to the outside (Fig. 1*A*) to interact with the HBV receptor (18).

Envelopment seems to be a regulated process in HBV. While cytoplasmic capsids carry viral genomes in all phases of maturation and occur also at high abundance as empty capsids (19, 20), secreted virions contain mostly rcDNA (21, 22). This observation prompted the "maturation signal" hypothesis (21) (Fig. 1*D*, green light), which posits a conformational signaling mechanism that couples a capsid's internal genome status to the envelopment proficiency of the capsid surface. Options for transmission of this signal included modified interactions between the N-terminal and C-terminal Cp domains within a monomer, or with respect to each other in the capsid, or differences in Cp–nucleic acid interactions owing to different Cp phosphorylation states (10, 23–25) and/or different types of nucleic acid (19) (Fig. 1*D*). In line with such a structure-change scenario, envelopment regulation can be modulated by certain mutations in Cp's assembly domain (22) (Fig. 1*D*), resulting in immature single-stranded DNA (ssDNA) intermediates (F97L) (23), no envelopment (and consequently no secretion) [P5T, P5A, P5S, L60V, L60A, L95A, K96A, etc. (15, 23)], or higher genome maturity (i.e., more DNA in double-stranded form) [genotype G (26, 27), Fig. 1*C*, cyan]. De facto, getting ahold of such maturation-associated structural changes has failed up to now.

Further complexity was added by the later finding that most cytoplasmic capsids carry no viral nucleic acid, yet can still be enveloped to yield "empty virions" (19, 20) (Fig. 1*E*), likely in large excess over true virions. This propelled a "single-strand blocking"

Significance

Viral hepatitis causes more deaths than tuberculosis and HIV-1 infection. Most cases are due to chronic infection with hepatitis B virus (HBV), which afflicts >250 million people. Current therapies are rarely curative, and new approaches are needed. Here, we report the discovery (by nuclear magnetic resonance) of a small molecule binder in the hydrophobic pocket in the HBV capsid. This structural element is, in an unknown manner, central in capsid envelopment. Binding of the pocket factor induces a distinct, stable conformation in the capsid, as expected for a signaling switch. This brings not only a new molecular view on the mechanism underlying capsid envelopment, but it also opens a rationale for its inhibition.

Author contributions: L.L., S.B., B.H.M., M.N., and A.B. designed research; L.L., S.W., M.D., P.Z., L.S., M.-L.F., M.B., M.S., T.W., L.C., and R.M. performed research; L.L., S.W., M.D., P.Z., L.S., M.-L.F., M.B., B.H.M., M.N., and A.B. analyzed data; and L.L., B.H.M., M.N., and A.B. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]L.L. and S.W. contributed equally to this work.

[2]M.N. and A.B. contributed equally to this work.

[3]To whom correspondence may be addressed. Email: michael.nassal@uniklinik-freiburg.de or a.bockmann@ibcp.fr.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2022464118/-/DCSupplemental.

Published April 20, 2021.

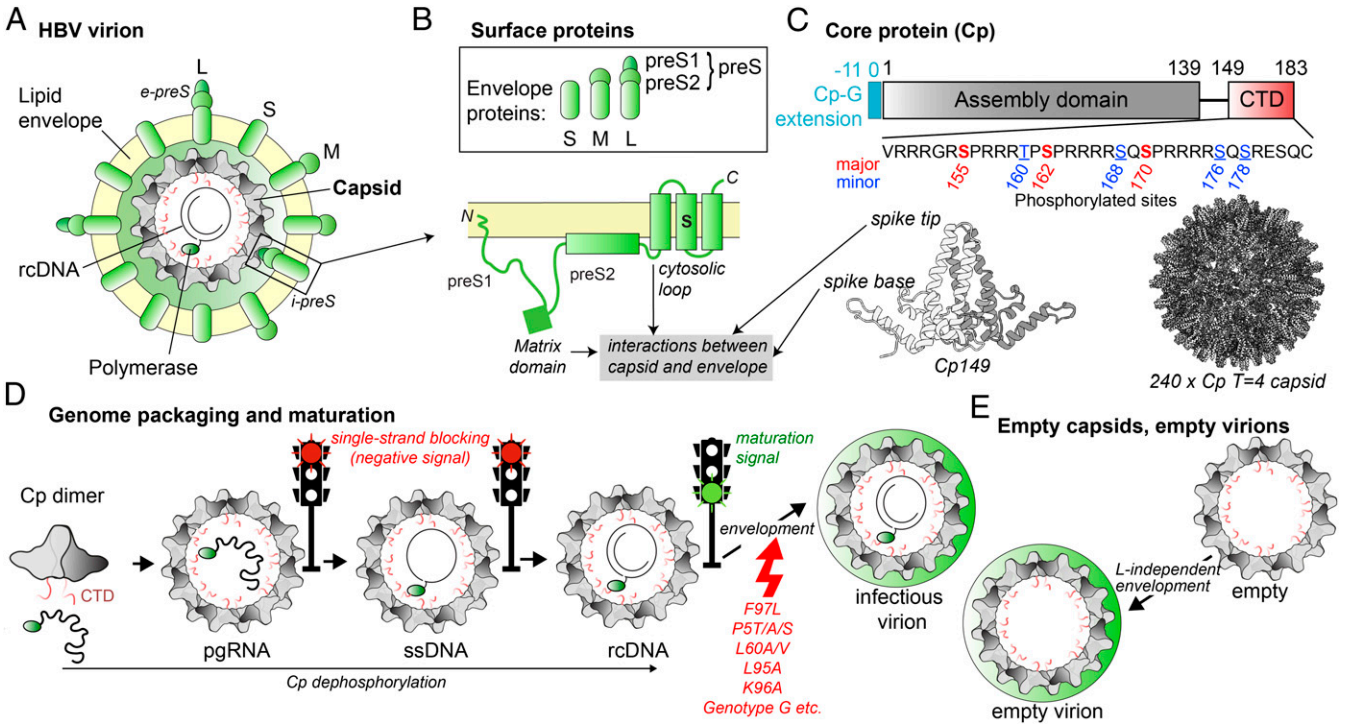

Fig. 1. Structure and maturation of HBV virions. (*A*) Mature HBV is made by a nucleocapsid containing rcDNA, surrounded by an envelope in which the three L, M, and S surface proteins are inserted. (*B*) The three surface proteins share the S domain, and M and L have additional N-terminal preS1 and preS2 extensions. (*C*) The core protein shows an assembly domain and a CTD, which can be phosphorylated. The assembly domain forms a dimer, 120 copies of which assemble into the T = 4 capsid. Genotype G (Cp-G) shows a 12 amino-acid extension at the N terminus of core. (*D*) The core protein assembles, via its CTD, around the pgRNA/polymerase complex to form the immature core. The RNA is reverse transcribed into ssDNA and then into partially double-stranded rcDNA. The CTD is subsequently, at least partially, exposed (71). The mature capsid becomes enveloped, and the virion is secreted. The maturation signal and single strand blocking hypothesis aim at providing possible rationales for envelopment regulation. Mutations can cause altered genome maturity (16, 55). (*E*) Empty capsid envelopment can occur without the need for L. The resulting particles are then secreted as empty virions.

model (19) (red lights in Fig. 1*D*), which proposes a negative signal exposed on pgRNA or ssDNA capsids. New data suggest that envelopment of empty capsids might actually take place through distinct interactions as it occurs even in the absence of L (24) (Fig. 1*E*). Overall, however, the molecular mechanisms underlying envelopment of mature DNA-containing as well empty capsids remain unclear.

In view of the viral particle heterogeneity in vivo, nearly all high-resolution X-ray crystallographic and cryogenic electron microscopy (cryo-EM) structures of HBV Cp are derived from bacterially produced capsid-like particles (CLPs) (9, 25, 26, 28–30). Likely due to structural disorder, the CTD has not been resolved in any study.

A variety of model CLPs mimicking distinct capsid states have been developed, including full-length Cp183 CLPs from *Escherichia coli*, which harbor ~3,500 to 4,000 nucleotides of bacterial RNA per capsid (11, 27), and empty CTD-less Cp149 CLPs (7). The abundance of the smaller T = 3 particles can be greatly increased through deletions in the linker (Cp140) (31, 32). Furthermore, Cp variants with mutations of disulfide-forming cysteines (33), secretion-relevant residues such as F97L, or mimicking partial phosphorylation (e.g., Cp183-EEE; S155E, S162E, and S170E) are easily accessible, and protocols for in vitro CLP disassembly and reassembly have been developed (30). Cp phosphorylated at seven sites (P7-Cp183) can be obtained via coexpression with serine-arginine–rich protein kinase 1 (SRPK1) (11). Cryo-EM reconstructions show only minor differences between Cp183, P7-Cp183, and F97L Cp183 (26).

As an approach toward a high-resolution definition of the HBV capsid conformation, we used solid-state nuclear magnetic resonance (NMR), a technique able to reveal even subtle structural differences in large particles, such as virus capsids and their interactants (reviewed in refs. 34 and 35) in a sedimented form (36–39). NMR, as a spectroscopic method, can record a signal for each ^{1}H, ^{13}C, or ^{15}N nucleus; as the NMR chemical shifts are sensitive to the local electronic environment around each nucleus, they directly report conformational information with utmost sensitivity (40, 41). Once the NMR signals of a specimen are known from sequential assignment, as is the case for the HBV capsid (42, 43), chemical-shift differences (CSDs) occurring in response to different modifications or treatments can be mapped to the atomic level on a representative three-dimensional (3D) structure.

Enabled by the biochemical and biophysical tools described above, here, we identify a stable alternative conformation of the HBV capsid. We show that this conformation is accessed in most of the investigated capsid states, with the interesting exceptions of genotype G and cell-free synthesized (CF-Cp183) capsids. We establish that the origin of this structural switch is the occupancy of Cp's hydrophobic pocket by a molecular binder, identified as Triton X-100, and that Cp mutations interfering with the binding of this pocket factor abolish envelopment. Our findings shed light on the dynamic aspects of the HBV capsid structure known to be central for its different functions. They lead us to advance a new hypothesis in signaling envelopment proficiency and define the hydrophobic pocket as a new druggable target for intervention with the HBV life cycle.

Results

NMR Spectra Reveal Two Different Capsid Conformations. We recorded NMR spectra on sediments of *E. coli* natively self-assembled capsids (Fig. 2*A* [blue], and *SI Appendix*, Fig. S1) and capsids reassembled from Cp dimers and in vitro transcribed pgRNA

(Fig. 2*A* [pink] and *SI Appendix*, Fig. S2). Intriguingly, we observed substantial CSDs between the peaks in the two spectra, indicating distinct conformational differences at specific amino acid positions (Fig. 1*B*); we termed the underlying conformations as forms A and B. The spectra of form B were assigned in a similar manner as described for form A (43).

One possible factor causing the nonidentical capsid conformations was the different RNA content (i.e., either length- and sequence-heterogeneous *E. coli* RNA or the defined pgRNA). However, other options remained. Passage through the partially denaturing disassembly–reassembly conditions could have destroyed an intermediate state established in *E. coli* by chaperones, disulfide-bond promoting enzymes, or the like. Reassembly could have happened in a different icosahedral symmetry, or the conformational change could have been induced by the binding of a molecule which was removed during the disassembly/reassembly steps.

Most, but Not all, Capsid Modifications Allow Access to Both A and B Conformations. To pinpoint the structural origin of the two conformations, we compared capsids from further Cp types (*SI Appendix*, Table S2), namely, from sevenfold phosphorylated P7-Cp183 (*SI Appendix*, Fig. S2) (11), the partial phosphorylation mimic Cp183-EEE (11), the immature secretion variant Cp183-F97L, CTD-less Cp149, T = 3-dominated Cp140 capsids, genotype-G capsids (44), and cell-free synthesized CF-Cp183 (45). P7-Cp183, Cp140, and Cp149 are devoid of nucleic acids (*SI Appendix*, Fig. S3). Negative-stain EM micrographs (Fig. 2*C*) revealed a similar capsid appearance for all forms investigated and the expected smaller size for the Cp140 T3 capsids.

We recorded NMR spectra for all capsids forms, and Fig. 2*D* shows the I59 Cδ1-Cβ signal as a signature cross peak that distinguishes A and B forms. For CF-Cp183, the proton resonance of residue W62 in a ^{1}H-detected proton-nitrogen correlation (hNH) spectrum served as signature (Fig. 2*E*). Pairwise comparisons revealed that P7-Cp183, Cp183-F97L, Cp183-EEE, and Cp149 can all principally exist in form A and form B (Fig. 2*F* and *SI Appendix*, Table S2). However, in most preparations, one of the peaks and, hence, one of the conformations, prevailed; in general, this was form A for the *E. coli* and form B for the reassembled samples (Fig. 2*F*). Only the *E. coli*–isolated Cp140 containing >85% of T = 3 particles presented a mixture of A and B forms, but B predominated in the reassembled preparation (Fig. 2*D*).

Hence, preparation history, rather than an intrinsic property, determines the fraction of Cp in conformation A. The only exceptions were genotype G capsids (Fig. 2*D*, and full aliphatic region in *SI Appendix*, Fig. S4) and CF-Cp183 (Fig. 2*E*, and full hNH spectrum in *SI Appendix*, Fig. S5), which were only detectable as form B. For genotype G capsids, these data suggest that the N-terminal extra sequence on Cp (*SI Appendix*, Fig. S4*B*) impedes access to form A; for CF-Cp183, they imply either a negative impact from some eukaryotic modification(s), as possible in the wheat-germ system used, or the lack of a form A–promoting factor that is present during isolation of capsids from *E. coli*.

The Differences between Forms A and B Localize to the Hydrophobic Pocket. CSDs, as site-specific indicators, allow to reveal structural differences at the level of individual atoms in a protein. Fig. 3*A* shows the CSDs between forms A and B for Cp183 (for other capsid types, see *SI Appendix*, Fig. S6), for which sizes of up to 1.8 ppm indicate significant conformational differences (40). Pronounced CSDs (in red) concentrate at the N terminus and around positions 60 and 100; in addition, V27 and L42 show isolated differences. While distant in primary sequence, these regions colocalize in the 3D fold at the spike base and the hydrophobic pocket. In Fig. 3*B*, residues with large CSDs are shown as red spheres on the Cp149 structure (29). The spatial arrangement of the residues is shown in Fig. 3*C*, with lines connecting residues <4 Å apart. In this network, extensive intramonomer contacts (lines connecting same-color residues) exist between residues in helices 3 and 4, between helices 2 and 3, and between the N terminus and helix 4. Intermonomer contacts exist between the N terminus and helices 2 and 3, the salt bridge–forming residues E64 and K96, and the two juxtaposed C61-C61 residues. Only E8 remains isolated from this network. CSDs are absent from the interdimer interfaces which form the fivefold and quasi-sixfold vertices in the capsid; also, the spike tip is completely unaffected.

Remarkably, these conformational differences cover all residues (except L31) of the hydrophobic pocket, defined to include residues V27, L31, A58, I59, C61, W62, M93, K96, I/F97, L100, and L101 (22), as shown in Fig. 3 *D* and *E*, on the protein surface. Further residues forming a groove around the hydrophobic pocket also display pronounced CSDs, including the E64-K96 salt bridge lining its entrance (Fig. 3*E*) and the surface-exposed N-terminal residues.

Altogether, our data indicate that the conformational differences between forms A and B localize mainly to the hydrophobic pocket but extend to remote regions, notably the very N-terminal amino acids.

Conformation A Is Induced by a Pocket Factor. Considering the accumulation of CSDs at hydrophobic pocket residues (Fig. 3*E*), we hypothesized that binding to the pocket of a molecule present during *E. coli* expression or purification was key to conformation A. This was also supported by the spontaneous, though very slow (on the order of months), conversion of *E. coli*–isolated Cp149 capsids into form B upon storage (Fig. 2*F*, percentage of form B after 1 mo storage) or faster upon gel filtration (Fig. 2*F* and *SI Appendix*, Table S2), a step included for Cp dimer purification before reassembly. To validate this pocket factor hypothesis, we designed two Cp mutants, P5W and L60W, which place a bulky Trp residue at the entrance of the pocket (Fig. 4 *A* and *B*). Indeed, as shown in Fig. 4*C* and *SI Appendix*, Table S2, spectra of Cp183-P5W and Cp183-L60W capsids isolated from *E. coli* conformed exclusively to state B. Mutation-specific CSDs involved numerous residues beyond the mutation site, particularly at the N terminus and around positions 60 (helix α3) and 100 (helix α4), congruent with, though less pronounced than, the differences between forms A and B in Cp183 (*SI Appendix*, Fig. S7 *B* and *C*). In comparison, mutation C61A did not impact remote residues except M93 (*SI Appendix*, Fig. S7*A*).

In order to identify the pocket factor, we extracted a solution containing form A capsids with chloroform/methanol and recorded NMR spectra of the chloroform phase (*SI Appendix*, Fig. S8*C*). A comparison of the resulting signals with chemical shift databases revealed the molecule as 2-[4-(2,4,4-trimethylpentan-2-yl)phenoxy] ethanol (also called t-octylphenoxypolyethoxy-ethanol), a compound commonly used in virus inactivation and better known as Triton X-100, a detergent closely related to nonyl phenoxypolyethoxylethanol (NP-40), and IGEPAL CA-630. Recording a spectrum of commercial Triton X-100 confirmed the assignment (*SI Appendix*, Fig. S8*C*), whereas the chloroform phase of analogously analyzed reassembled capsid did not yield signals (*SI Appendix*, Fig. S8*A*).

As a detergent, Triton X-100 is often added during Cp purification to the *E. coli* cell extract. To verify its responsibility for the conformational switch, we added Triton X-100 to reassembled capsids in form B; this, indeed, induced a conversion to form A (Fig. 4*C*). Assessing the affinity of Triton X-100 to Cp183 by isothermal titration calorimetry (ITC) revealed a dissociation constant (K_D) of 8.3 ± 0.5 μM (*SI Appendix*, Fig. S9). NMR titrations of the Cp149 dimer with Triton X-100 confirm a K_D value around or below the micromolar range, since they display two separate signals for residues showing CSDs between forms A and B, without exchange between them on the NMR-typical millisecond time scale.

Altogether, these data demonstrate that occupancy of the hydrophobic pocket with a binding molecule (here, Triton X-100) is

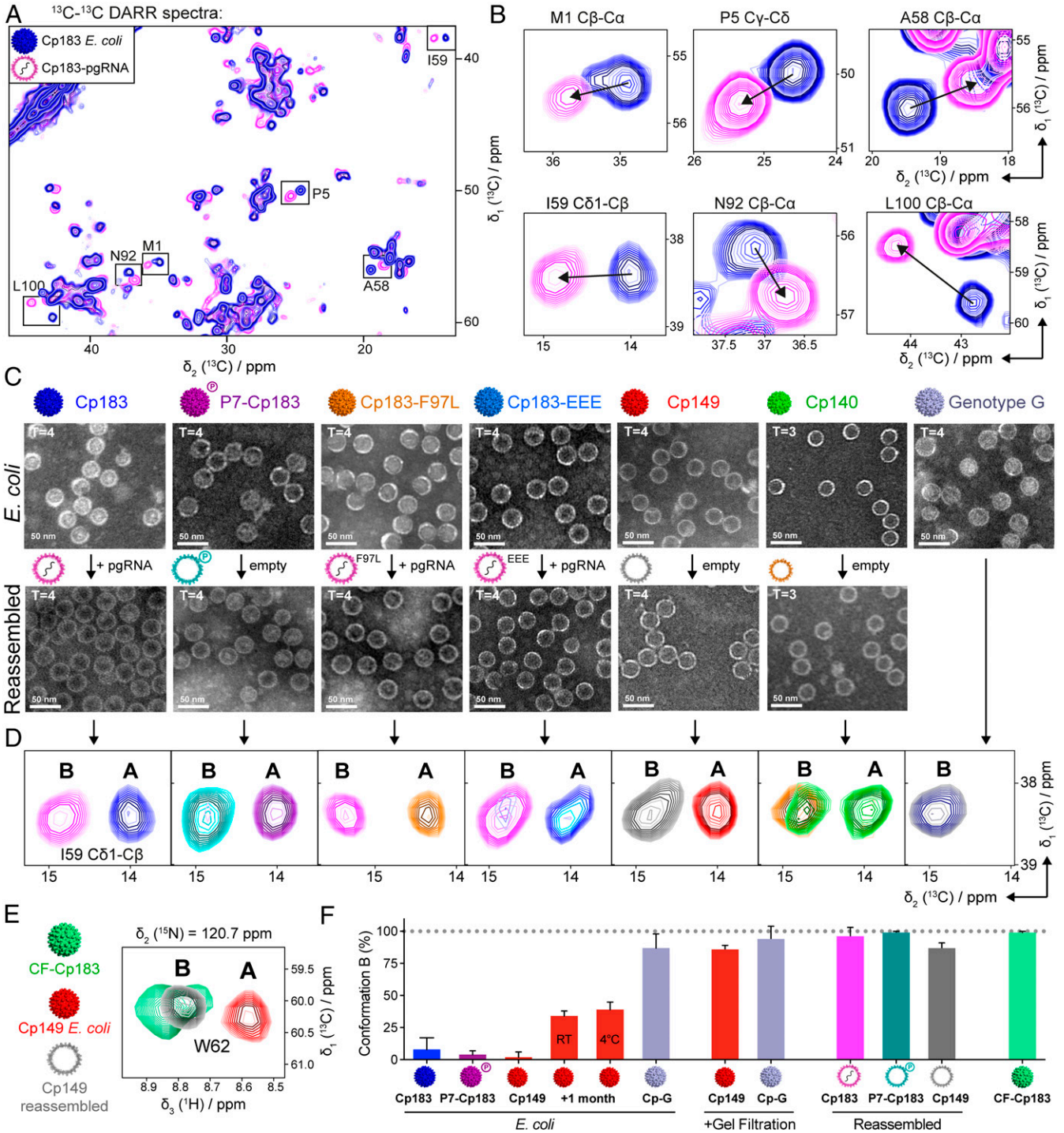

Fig. 2. All Cp show two conformations except for genotype G and cell-free synthesized capsids. (*A*) Overlay of 2D ^{13}C DARR spectra of Cp183 *E. coli* (blue) and Cp183-pgRNA (pink). Examples of CSDs are framed and labeled with the corresponding residue number. (*B*) Zoom on selected resonances from *A*. (*C*) Negative-stain EM of the different Cps: Cp183, P7-Cp183, Cp183-F97L, Cp183-EEE, Cp149, and Cp-GenotypeG (Cp-G) form T = 4 capsids in both *E. coli* and upon reassembly; Cp140 assembles mainly in T = 3 capsids. (*D*) I59 Cδ1-Cβ peaks of 2D DARR spectra for *E. coli* isolated versus in vitro reassembled capsids. For Cp149, Cp183, Cp183-F97L, Cp183-EEE, and P7-Cp183, *E. coli*–isolated capsids conform mainly to form A, and reassembled capsids mainly to form B. Exceptions are *E. coli*–isolated Cp140 capsids, which comprised a mixture of A and B forms and Cp-G capsids that conformed exclusively to B, regardless of the preparation procedure. Color-coding of the NMR peaks is according to the capsid in *C*. (*E*) 3D proton-nitrogen-alpha carbon correlation (hCANH) spectrum extract showing W62 of cell-free synthesized CF-Cp183 (green) versus Cp149 *E. coli* (red) and reassembled (gray) capsids. CF-Cp183 are in form B. (*F*) Percentage of conformation B in the different capsid forms. In addition, two Cp149 *E. coli* capsid preparations left for 1 mo at different temperatures are shown as well as Cp149 and Cp-G *E. coli* capsids after a gel filtration step. The percentage of B was calculated from six representative and isolated peaks in the DARR spectrum (see *Materials and Methods*). Error bars represent SDs.

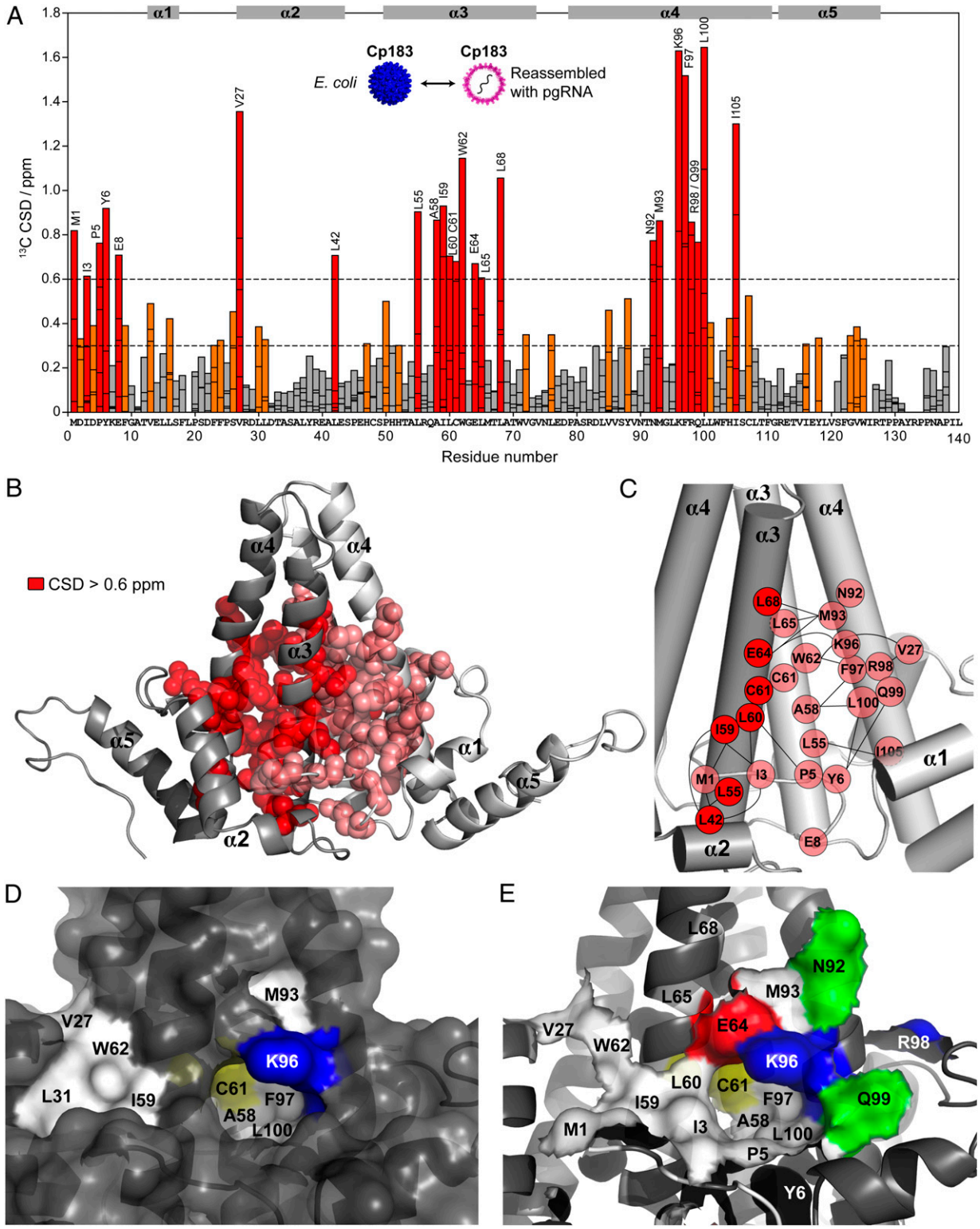

Fig. 3. Conformational differences between forms A and B. (*A*) Histogram showing CSDs between A and B forms ($CSD = |\delta_C[A] - \delta_C[B]|$). A single bar is plotted for each assigned carbon atom (see for example V27, where the five transparently superimposed bars correspond to the CSDs of the five assigned carbon resonances). The protein sequence up to residue 140 is given below the graph, and CSDs >0.6 ppm are shown in red, <0.3 ppm in gray, and in between in orange. (*B*) Residues with CSDs >0.6 ppm are shown as spheres on the X-ray structure of a Cp149 dimer [chains C and D of Protein Data Bank (PDB) accession code 1QGT (29)] and localized at the base of the spike. Color code is as follows: dark gray and red for chain C; light gray and light red for chain D. (*C*) Proximities in the 3D structure (≤4 Å) between residues with CSDs >0.6 ppm are shown by black lines. (*D*) Residues from the hydrophobic pocket according to ref. 53. (*E*) Residues showing large CSDs between A and B forms (in red in *A*) in surface mode on the X-ray structure (color coded according to residue type as follows: white, hydrophobic; red, acidic; blue, basic; green, polar; and yellow, cysteine).

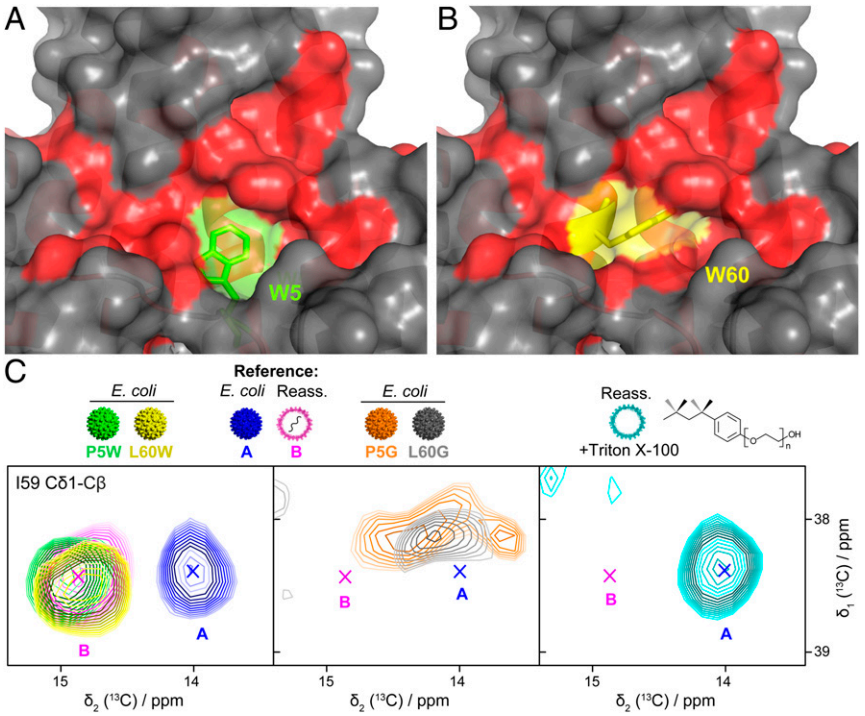

Fig. 4. Mutational obstruction of the hydrophobic pocket blocks access to form A, while adding Triton X-100 to wild-type capsids restores form A. In silico mutated W5 (*A*) and W60 (*B*) as predicted by PyMol (Schrödinger, LLC) highlighted on the Cp149 structure [PDB: 1QGT (29)], picturing how the P5W and L60W replacements can obstruct the hydrophobic pocket. Residues involved in the conformational A/B switch are colored in red. (*C*) I59 Cδ1-Cβ correlation peak extracted from a 2D DARR recorded on Cp183 capsids isolated from *E. coli* in form A (in blue), reassembled with pgRNA in form B (in pink), and compared to mutants P5W (in green), L60W (in yellow), P5G (in orange), and L60G (in gray). P5W and L60W capsids isolated from *E. coli* adopted >95% form B, while P5G and L60G showed a broad intermediate peak. Addition of Triton X-100 at a 8:1 molar ratio to capsids reassembled in form B (*Right*) fully restored the form A (in cyan).

responsible for the appearance of form A and that loss of this molecule induces form B. Triton X-100 as pocket factor binds with high specificity and affinity, as only a limited set of residues showed significant chemical shift changes, and no multiple peaks for a single residue were observed, indicating full saturation of the binding sites when added in eightfold excess. In the simplest model, the dynamic space of the capsid can thus be described by two minima (full versus empty pocket), providing a basis for a binary switch.

Other Cp Modifications Induce at Most Minor Conformational Differences. In order to relate the extent of the pocket factor–induced conformational change between forms A and B to one or another of the experimentally addressable capsid modifications, we pairwise compared the CSDs induced by the presence versus absence of the CTD, CTD phosphorylation, different nucleic acids, T = 3 versus T = 4 symmetry, several site-specific mutations, and the oxidation state of cysteine C61 (Fig. 5 and *SI Appendix*, Fig. S10).

The most pronounced CSDs were seen when comparing Cp183 with Cp183-F97L (Fig. 5*A* and *SI Appendix*, Fig. S10*A*). Remarkably, especially in the reassembled capsids, many of the most strongly affected residues in the F97L mutant coincide with those showing large CSDs between forms A and B in wild-type Cp183.

Interestingly, the impact of the presence or absence of pgRNA is small, as highlighted by the comparison between empty (phosphorylated) and pgRNA-containing Cp183 capsids (Fig. 5*D*). One can note that the presence of the CTD induced a slight attenuation of two glutamates located at the inner capsid lining (Fig. 5*G* and *SI Appendix*, Fig. S11), pointing to their interaction with the CTD. CSDs comparing T = 3 and T = 4 geometries are only minor, and are shown in Fig. 5*E* and *SI Appendix*, Figs. S10*E* and S12. Data shown in *SI Appendix* Figs. S13 and S14 confirm that the capsid's C61 oxidation state is unrelated to forms A and B.

Taken together, these comparisons show that, while point mutations can detectably impact capsid conformation, none of the capsid modifications induced structural differences that were as substantial as those distinguishing forms A and B.

Virological Impact of Selected Cp Mutations Affecting A-form Accessibility. To unravel a potential biological role of the accessibility of form A for a given Cp sequence, we sought to address, in transfected human hepatoma cells, the envelopment capacity of Cp mutants P5W and L60W (i.e., with the largest possible sidechain), which, as recombinant CLPs, were exclusively found in form B (Fig. 4 and *SI Appendix*, Fig. S7). For comparison, we included mutants P5G and L60G carrying the smallest possible sidechain.

pgRNA serves as a substrate for reverse transcription yet also as mRNA for Cp and polymerase proteins. To avoid any impact of Cp open reading frame (ORF) mutations on replication, we employed a trans-complementation system (46, 47) in which always the same vector encoding a Cp-deficient HBV genome (pCH-9/3091sCX-core$^-$; short pCH_core$^-$) is cotransfected with a separate Cp expression vector, such as pTRUF-HBc183 (short pHBc_wt) encoding wild-type Cp or the desired mutant. Formation of replication-competent nucleocapsids was monitored by the viral DNA contents of intracellular capsids and/or nonenveloped ("naked") capsids that are released in varying amounts into the culture supernatant (48); envelopment proficiency was addressed by the formation of virions, that is, enveloped viral DNA-containing capsids based on their distinct mobility from naked capsids in native agarose gel electrophoresis (46). For better resolution, the concentrated particles from the cell culture supernatants were preseparated on a Nycodenz gradient according to their different buoyant densities, as established using particles from the stable wild-type HBV-producing cell line HepG2.117 (46) (*SI Appendix*, Fig. S15 *A* and *B*).

The hepatoma cells transiently transfected with a wild-type HBV vector (pCHwt) or cotransfected with pCH_core$^-$ plus pHBc_wt both produced virions and naked capsids, although the latter particles predominated (*SI Appendix*, Fig. S16). In contrast, P5W and L60W mutant Cps, as well as P5G and L60G, all generated viral DNA-containing capsids but no distinguishable virion signals, as shown in Fig. 6 for the gradient fractions 8 to 10 from the transition region between virions and naked capsids. Hence, under conditions where complementation with wild-type Cp generated reasonable amounts of virions, all four mutant Cps

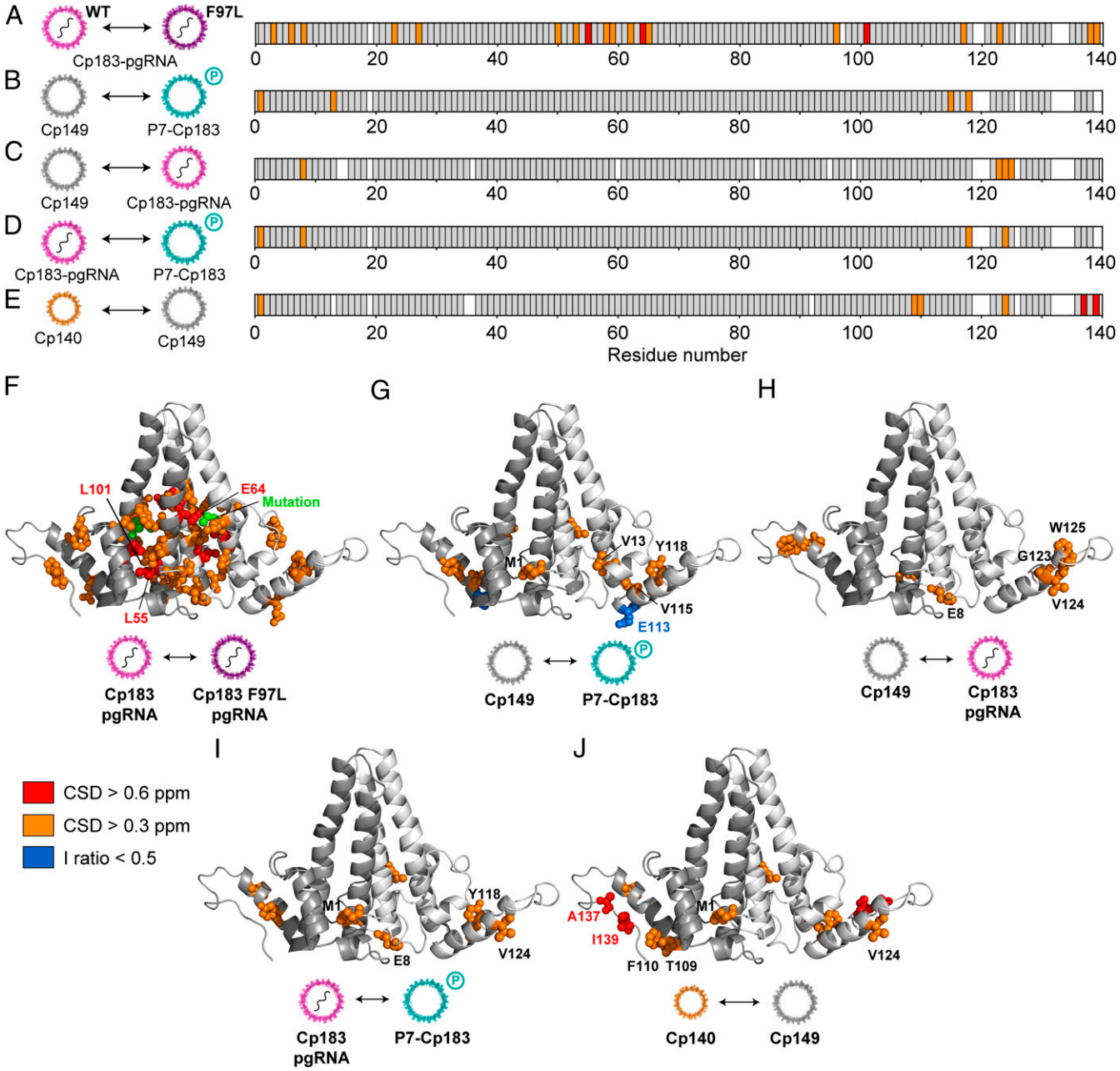

Fig. 5. CTD phosphorylation, RNA content, and capsid geometry have little impact on the capsid structure. (*A–E*) ^{13}C CSDs mapped on the amino-acid sequence in a compact form with the same color coding as in Fig. 2*A* (for full plots, see *SI Appendix*, Fig. S12 *A–E*) and (*F–J*) CSDs plotted on the 3D structures. CSDs are shown for F97L mutation on Cp183 (*A* and *F*); empty capsids with and without CTD (Cp149 versus P7-Cp183) (*B* and *G*); absence/presence of RNA (Cp149 versus Cp183) (*C* and *H*); CTD phosphorylation and nucleic acid content (Cp183 versus P7-Cp183) (*D* and *I*); and T = 3 versus T = 4 capsid geometry (Cp140 versus Cp149) (*E* and *J*); on the 3D structure [PDB 1QGT (52)], residues with medium and large CSDs are shown as orange and red spheres, respectively. E113 which lost more than 50% of intensity in presence of the CTD is shown in blue and the mutated residue F97 in green.

failed to support envelopment, although they did support capsid-internal genome replication.

The similar phenotypes caused by the replacement of P5 and L60 with residues as disparate as Trp and Gly were surprising. We therefore also recorded NMR spectra of both Gly variants, which revealed substantial spectral changes (*SI Appendix*, Fig. S7 *D* and *E*). A clear distinction between A and B forms was difficult, as the signals of the residues in and around the hydrophobic pocket of the Gly mutants were indicative of a slightly different conformation, having both features of A and B forms. Also, the presence of broadened resonance lines observed for many distinct A- and B-form resonances might indicate a dynamic equilibrium between A and B (intermediate, fast chemical-exchange effects) (Fig. 4*C*). The rather strong impact of the Gly mutations could relate to the helix-breaker effect of Gly in L60G and/or the disruption of the bridging function of P5 or L60 in the L60-P5-Y6 hydrophobic triad. So, while the W mutations obstruct the pocket, the G mutations destabilize a defined structure of the pocket, which in a lock and key interaction model, would reduce the structural complementarity and hence the affinity of the pocket for the substrate.

Taken together, both obstruction of the hydrophobic pocket as well as modification of its binding affinity impair capsid envelopment,

although they remain compatible with the structural requirements for capsid-internal replication.

Discussion

We have investigated by solid-state NMR over 10 different assemblies of wild-type Cp-based HBV capsids, covering full-length versus truncated, empty versus RNA-filled, and various mutant Cps. We determined the NMR chemical shifts for all variants and pairwise compared them in order to fingerprint structural alterations induced by the presence of the CTD, nucleic acid-content, icosahedral symmetry, and phosphorylation and oxidation state. Our data show invariably that the conformational alterations associated with these variations are all small (summarized in Fig. 7*A*) compared to the single most important conformational change we detected, independently of the type of construct, between the two distinct forms A and B identified in this work. NMR chemical shift changes clearly localize the conformational differences to Cp's hydrophobic pocket and its delimiting residues, and we elucidated the cause to be a high-affinity molecular binder, identified as Triton X-100. Form B corresponds to the unbound state, while the addition of Triton X-100 fills the hydrophobic pocket and induces form A. Interestingly, this binding is distinct from that of the small molecule capsid assembly modulators (e.g., the heteroaryldihydropyrimidine [HAP] compounds and others), which all target the so-called hydrophobic cavity at the interdimer interface (49). As Cp mutations preventing access to the pocket lock the capsid in form B and impede formation of enveloped virions, we propose that form A relates to a previously undetected envelopment-ready conformation of the HBV capsid.

A first question raised by our data is to which form the known HBV capsid structures correspond. The X-ray crystallographic structures likely represent the unbound state B since they did not show extra density. Furthermore, even if the respective capsids were isolated from *E. coli* via very similar protocols as used here (29, 50, 51), obtaining highly diffracting crystals took several months (50, 52). Hence, as in our long-term stored samples (Fig. 1*F*), the bound detergent was likely lost from the pocket before crystal formation. Similarly, the increase in structural order upon long-term storage of Cp183 and P7-Cp183 capsids seen in recent high-resolution EM structures (11) suggests that they as well correspond to form B (see below for a single exception), in line with loss of form A from an initially mixed A plus B population that, by EM, may be detected as heterogeneity.

Structures with an unknown density in the hydrophobic pocket have been described in two studies. In the first, an extra density was observed in 8.8 Å cryo-EM reconstructions of capsids from serum virions from which the envelope was stripped (53). Such stripping generally employs one of the closely related mild detergents NP-40 or Triton X-100; hence, the observed extra density plausibly derives from such detergent as in form A. It is not surprising that the same study found no extra density in the hydrophobic pocket of the *E. coli*–derived RNA-filled reference capsids because the reported purification procedure did not involve Triton X-100 or a similar detergent (53). In the second study, extra density was observed in high-resolution (<3 Å) cryo-EM reconstructions of *E. coli*–produced Cp183-F97L capsids that, as here, had been purified in the presence of Triton X-100 (26). The density mapped to the vicinity of residues 5, 6, 60, 64, 93, and 96 and thus to residues for which we saw some of the largest CSDs. Given the evidence for Triton X-100 binding to the HBV capsid, the pocket factor observed in the cryo-EM structure is certainly Triton X-100. That no pocket factor was visible in wild-type Cp183 particles may suggest that the enlarged pocket in Cp183-F97L (26) strengthens the interaction. Other options include insufficient overall resolution in the other studies yet also the hitherto unknown importance of storage conditions and duration for occupancy of the pocket. At any rate, the large chemical shift effects illustrate the extraordinary sensitivity of NMR to detect, and resolve with atomic resolution, even minor changes in geometry (41) and its complementarity to the other high-resolution structure determination methods (35). It is thus remarkable that the presence of pgRNA does not induce any important change in chemical shifts. This is of particular interest in the context of the single-strand blocking hypothesis, which, plainly interpreted, predicts an envelopment-impairing alteration in capsid structure, specifically by encapsidated single-stranded nucleic acid (18). The absence of significant CSDs invalidates, in the present in vitro setup, the hypothesis of a conformational signal (Fig. 5*D*; see also *SI Appendix*, Fig. S10*D*).

In sum, our NMR data provide clear-cut evidence that the HBV capsid can bind an exogenous factor in its hydrophobic pocket, which provides the basis for a binary conformational switch. A more general question is the physiological relevance of the hydrophobic pocket and its interaction with Triton X-100 or, more likely, a natural, structurally related pocket factor. The capsid's hydrophobic pocket is conserved throughout HBV isolates [6 out of 10 residues are fully conserved, and the remaining four are highly similar according to the Cp genetic variability as reported in the HBVDB (54)], and the importance of its integrity is highlighted by various hydrophobic pockets and P5 and L60 variants, which are proficient for capsid-internal replication but defective for envelopment (16, 55). As only enveloped HBV can spread, envelopment is as critical for the virus as genome replication. Several scenarios can be envisaged in respect to how the hydrophobic pocket, a defined void in the HBV capsid structure that can be filled from the outside, could be involved in envelopment (Fig. 7).

In one scenario (Fig. 7*B*), the natural pocket factor mimicked by Triton X-100 is of viral origin. Given its known importance for virion formation, the matrix domain at the preS1/preS2 interface (56) and, in particular, its hydrophobic residues are prime candidates (represented by a green square in Fig. 7). Coupled to another maturation-related event, in other words, second DNA strand extension and/or a preS-independent capsid-envelope

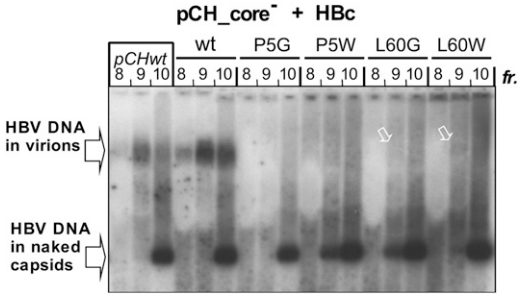

Fig. 6. Cp mutants with impaired access to form A support capsid-internal HBV replication but not capsid envelopment. HepG2 cells were transfected with a wild-type HBV expression vector (pCHwt) or cotransfected with an analogous Cp-deficient HBV vector (pCH_core⁻) plus a separate HBc expression vector encoding wild-type Cp (wt) or the indicated mutants. Particles in the culture supernatants were preseparated in Nycodenz gradients according to their buoyant densities (virions ~1.18 g/mL; naked capsids ~1.22 g/mL) and then via native agarose gel electrophoresis by electrophoretic mobility. Signals represent HBV-specific DNA in virions and in naked capsids, as revealed by hybridization with a radioactively labeled HBV probe. The presence of envelope proteins plus Cp (virions) versus only Cp (naked capsids) at the respective gel positions was shown by immunoblotting, as detailed in *SI Appendix*, Figs. S13 and S14. Note the lack of DNA signals at the virion position for all mutant Cps, perhaps except very faint signals for the L60 mutants (arrows), despite wild-type Cp-like DNA signals at the naked capsid position; hence, the four Cp mutants support capsid-internal replication but are deficient for envelopment. The lighter areas between virion and capsid positions arise from nonviral bulk protein in the polyethylene glycol (PEG) precipitated particle preparations.

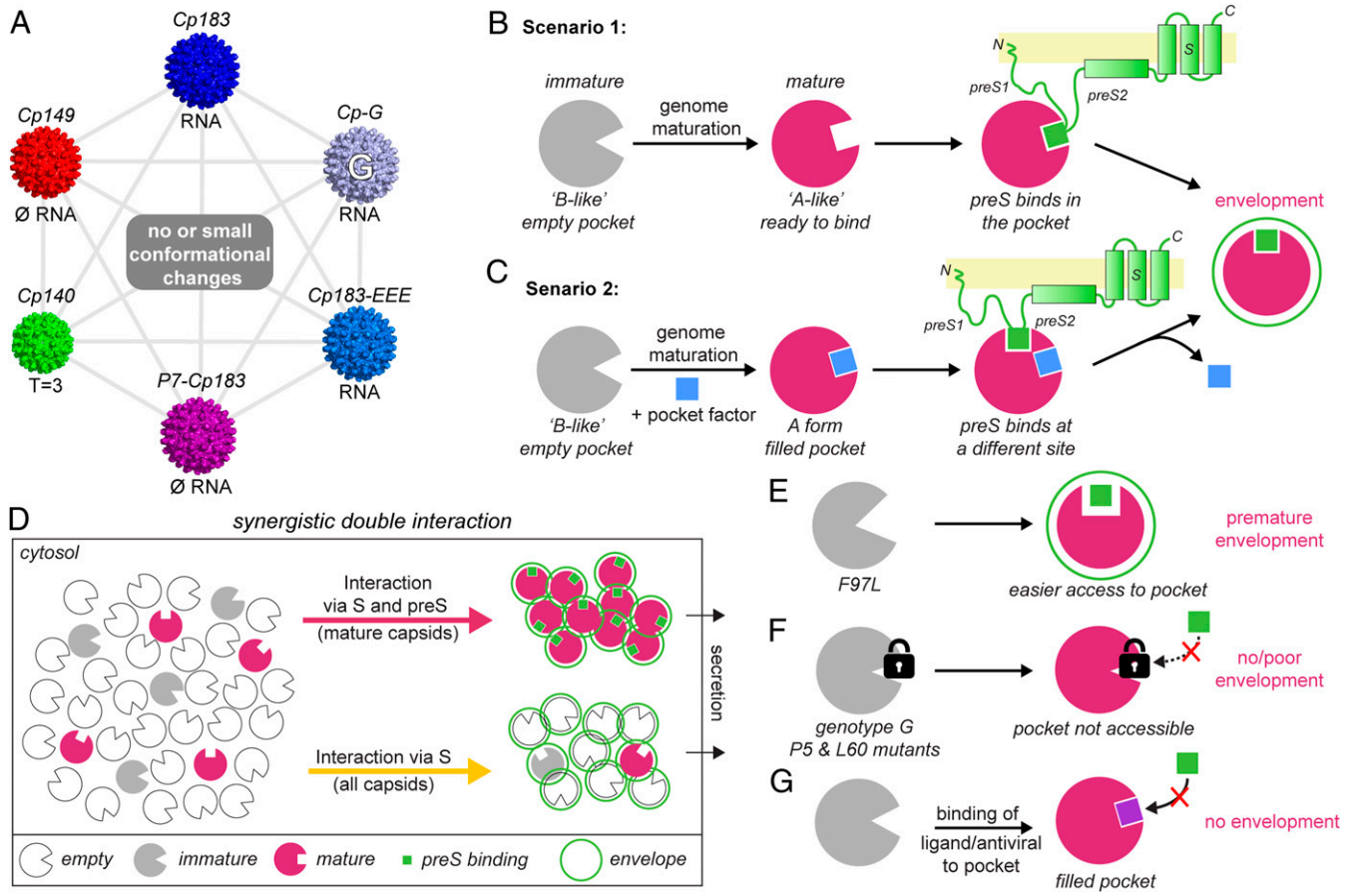

Fig. 7. Possible role of the two different capsid forms in the context of viral particle formation. (*A*) The presence of the CTD, its phosphorylation, and the presence of nucleic acids or the T = 3/T = 4 symmetry do not induce significant conformational change. (*B*) In scenario 1, capsid maturation changes the hydrophobic pocket so that is becomes proficient for interaction, likely with the preS matrix domain, represented as a green square. (*C*) In scenario 2, a pocket factor, causing a conformation similar to form A, binds the capsid. The bound pocket factor enables preS binding at a different site in the capsid. The pocket factor is then either released upon preS binding (as shown) or remains in the enveloped particle. (*D*) Cartoon of the "synergistic double interaction" hypothesis, where two different mechanisms combine to produce the observed phenotype (e.g., mainly empty and mature capsids are enveloped). (*E*) In the F97L mutant, the pocket is structurally modified, thereby favoring the interaction with preS and resulting in premature envelopment. (*F*) Genotype G capsids as well as Cp183-P5 and L60 mutant capsids have no or only poorly accessible hydrophobic pockets, resulting in reduced or impaired envelopment. (*G*) Binding of a ligand to the hydrophobic pocket results in an antiviral effect through inhibition of preS binding, therefore preventing envelopment.

protein interaction (see below), the capsid would adopt a new envelopment-ready conformation that enables a stable interaction between preS and the Cp hydrophobic pocket. The widened pocket in the F97L variant which in form B (empty pocket) generated the most similar CSD pattern (Fig. 5 *A* and *F*) to wild-type Cp capsids in form A (filled pocket) may already be closer to this new conformation (Fig. 7*E*). A faster acquisition of the membranous envelope would lead to an earlier cut off of the deoxynucleotide triphosphate (dNTP) supply into the capsid interior, such that virions contain less mature DNA. A role for the myristoyl residue on Gly2 of preS1 is less likely because the myristoylation-preventing G2A mutant still allows formation of morphologically intact, although noninfectious, virions (57, 58). Also, we did not detect any specific binding of myristic acid to HBV core proteins.

In a second scenario (Fig. 7*C*), the natural pocket factor is a cellular molecule whose binding, again coupled to an upstream event such as genome maturation, induces the capsid to adopt an envelopment-ready conformation. Occupancy of the pocket would enable preS binding yet to a separate interaction site. As also for the first scenario, mutations of pocket-relevant Cp residues could affect accessibility and/or shape complementarity of the pocket and, in this way, impact the efficiency of the capsid–envelope interaction (Fig. 7*F*). Examples of such interaction-modulating pockets are the bile acid binding pocket in murine norovirus capsid protruding domain (59) or the linoleic acid binding pocket in the SARS coronavirus-2 spike protein, which, when occupied, locks the protein in a conformation with reduced receptor interaction in vitro (60).

In a third scenario, the Triton X-100 complementary void in the Cp structure per se is the functionally relevant feature, with the free space lowering the energy barrier for the conformational transitions required during the viral life cycle; Triton X-100 binding would then be a consequence of the existence of the pocket. Such models are proposed for a long-known pocket in picorna virus capsids that is targeted by the antiviral WIN compounds (61) and a distinct, newly identified drug-binding pocket in rhino- and enterovirus capsids (62). Mutations in the HBV capsid pocket geometry could modify such a lubricant function, altering the barrier for a shift toward an envelopment-proficient conformation upon an upstream maturation-induced trigger. However, given the specific envelopment phenotypes of mutations which interfere with pocket accessibility, such as P5W and L60W or the N-terminal extra sequence in genotype G Cp, and conversely of F97L, which facilitates pocket access, this "functional void" model appears less likely.

All three hypotheses imply a genome maturation-related upstream event as a key for the accessibility and/or occupancy of the hydrophobic pocket in Cp, which then regulates the capsid–envelope interaction. A plausible trigger is the increasing conversion of flexible minus-strand DNA into much stiffer dsDNA, which eventually spring-loads the capsid structure (63), as supported by the freeze/thaw sensitivity of double-stranded DNA but not ssDNA containing HBV nucleocapsids (64). The respective forces could push on the capsid interior (53), including on the ends of helices α3 and α4, which together form the hydrophobic pocket and thus operate the conformational switch.

Obviously, such a mechanism does not apply to the envelopment of genome-less capsids, which generates the excess empty virions in serum (12). However, recent mutational data indicate that the requirements governing envelopment of genome-containing versus empty capsids are nonidentical; most strikingly, the S envelope protein alone suffices for empty virion formation (24). Hence, these data indicate the existence of a distinct second capsid–envelope interaction that is independent of preS as well as the genome status. Based on this, and on the lack of any detectable structural alteration by encapsidated single-stranded nucleic acid (Fig. 5 *D* and *I*) that could be interpreted as blocking signal, we propose a "synergistic double interaction" model for HBV nucleocapsid envelopment, as outlined in Fig. 7*D*. Its key features are a regulated capsid conformation that is specific for mature genomes (pink arrow in Fig. 7*D*) and is ready to bind preS versus a second weaker and signal-independent capsid interaction with S, perhaps the cytosolic loop, as previously proposed (13–15, 65) (Fig. 7*D*, yellow arrow). This second pathway would, by default, lead to nonselective envelopment of all capsids in the cell at ratios corresponding to their relative abundance. As highly phosphorylated empty capsids are the most abundant species, a combination of both mechanisms will yield the observed phenotype (i.e., secretion of mainly empty and, for wild-type Cp, fully mature capsids) (Fig. 7*D*). Cp mutations relevant for the preS-dependent pathway could then cause differential genome maturation–specific responses as outlined above. Hence, the presented model puts the original maturation signal hypothesis (21) on a plausible mechanistic basis, and the inclusion of a second interaction accounts for empty capsid envelopment.

Whether it is a natural pocket factor that is mimicked by Triton X-100 or the void space of the pocket which enables the conformational transitions of the HBV capsid we have detected remains to be determined (e.g., by high-resolution definition of the envelope–capsid interaction sites) optimally in virions with defined genome status. Importantly, however, in all scenarios, the binding of a pharmacologically acceptable and optimized homolog of Triton X-100 is expected to interfere with the HBV life cycle either as a competitor of the natural pocket factor or by freezing capsid dynamics in one conformation (Fig. 7*G*), as also discussed for the WIN compounds (61). This should warrant follow-up exploration of the therapeutic potential of our findings.

Conclusions

We have identified, by combining the HBV capsid *E. coli* toolbox with the high sensitivity of NMR chemical shifts, a second well-defined structural state for the HBV capsid, which is caused by a molecular binder to the Cp hydrophobic pocket. This transformation causes a conformational switch, which is far more significant than the effects of truncation, presence of nucleic acids, phosphorylation, or specific icosahedral symmetry, all of which show only minor impacts on the capsid structure. We speculate that the two observed states relate to the capsid's conformation involved in the interaction with the envelope preS, since the structural differences between the states colocalize with residues that are central for envelopment and related to altered secretion phenotypes. As modifications in residues gating the hydrophobic pocket block envelopment, and our experimental data do not support the single-strand blocking hypothesis (19), we formulate the "synergistic double interaction" hypothesis, which puts the maturation signal hypothesis on a plausible mechanistic basis and can account for the selective envelopment phenotype observed. Our work places binding to the hydrophobic pocket as a lead suspect on the list of possible conformational signals and should trigger innovative approaches to interfere with capsid envelopment and, ultimately, the formation of infectious HB virions.

Materials and Methods

Expression and Purification Cp140, Cp140-C61A, Cp149, Cp149-P5G, Cp149-L60G, Cp183, Cp183-F97L, Cp183-EEE, Cp183-P5W, Cp183-L60W, P7-Cp183, and Core Protein-Genotype G HBV. Proteins were expressed in *E. coli* and purified as described previously (11, 43), with minor modifications depending on the constructs, as detailed below. Briefly, plasmids of pRSF_T7-HBc140opt, pRSF_T7-HBc140C61Aopt, pET-28a2-HBc149opt, pRSF_T7-HBc183opt, pRSF_T7-HBc183F97Lopt, pRSF_T7-HBc183P5Wopt, pRSF_T7-HBc183L60WLopt, pRSF_T7-HBc149P5Gopt, pRSF_T7-HBc149L60GLopt, pRSF_Tet-SRPK1dNS1_T7-HBc183opt, and pET-28b-HBc-G were transformed into *E.coli* BL21* CodonPlus (DE3) cells and grown at 37 °C in M9 minimal medium culture containing ^{13}C-glucose and $^{15}NH_4Cl$ (for more details on labeled culture, see ref. 43). When the optical density at 600 nm (OD_{600}) reached between 0.7 and 2.0, expression of Cp140, Cp140-C61A, and Cp149 was induced with 1 mM isopropyl ß-D-1-thiogalactopyranoside (IPTG) overnight at 25 °C. For full-length proteins, Cp183 wild-type and mutants, P7-Cp183, and Cp from genotype G HBV (Cp-G), expression was induced when OD_{600} reached 2.0 at a temperature of 20 °C, except for P7-Cp183, for which the temperature after induction was set at 16 °C. For all constructs, cells were collected by centrifugation at 6,000 × *g* for 20 min and resuspended in 15 mL (for 1 L culture) TN300 buffer (50 mM Tris, 300 mM NaCl, 2.5 mM ethylenediaminetetraacetic acid (EDTA), 5 mM DTT, pH 7.5). The cell suspensions were incubated on ice with 1 mg/mL chicken lysozyme, 1× protease inhibitor mixture solution, and 0.5% Triton X-100 for 45 min. A total of 6 μL benzonase nuclease was added to digest nucleic acids for 30 min at room temperature. Cells were broken by sonication using a minimum of 10 cycles of 10 s of sonication and 50 s of cooling down on ice. Cell lysates were centrifuged at 8,000 × *g* for 1 h to remove cell debris. The supernatant was loaded onto a 10 to 60% sucrose gradient buffered with 50 mM Tris pH 7.5, 300 mM NaCl, and 5 mM DTT and centrifuged in SW-32Ti Beckman Coulter swinging bucket rotor at 140,000 × *g* for 3 h at 4 °C. Capsids in gradient fractions were identified by 15% sodium dodecyl sulfate (SDS)–polyacrylamide gel and precipitated by 20% (for Cp140 and Cp149) or 40% (for Cp183, P7-Cp183, and Cp-GenotypeG) saturated $(NH_4)_2SO_4$. After incubation on ice for 1 h and centrifugation at 20,000 × *g* for 1 h, pellets were resuspended in 10 mL purification buffer (50 mM Tris pH 7.5, 5% sucrose, 5 mM DTT, and 1 mM EDTA). The protein solution was centrifuged again for 15 min to remove insoluble pellet. The supernatant containing soluble capsids was dialyzed overnight against purification buffer at 4 °C prior to rotor filling. For two *E. coli* capsids samples (Cp149 and Cp-GenotypeG), an additional gel filtration step was performed where the protein supernatants were loaded onto a HiPrep 16/60 Sephacryl S-200 HR column (120 mL), equilibrated in purification buffer. Fractions containing the protein were pooled for rotor filling. For all *E. coli* nucleic acid–containing capsids, the nucleic acids are isotope labeled with the same isotopes as the protein. The preparation of cell-free capsids in the wheat-germ cell-free system was done as previously described in ref. 45.

In Vitro Transcription of pgRNA. See *SI Appendix, Extended Methods*. In vitro transcribed pgRNA is unlabeled.

Capsid Disassembly and Reassembly In Vitro. See *SI Appendix, Extended Methods*.

Separation of Cp140 T = 3/T = 4 Capsids. T = 3 capsids of Cp140 were separated from T = 4 capsid by sucrose density gradient centrifugation, as described previously (31). Sucrose gradients were prepared from 5, 10, 15, 20, 25, and 30% (weight per volume) stock solutions in TN300 buffer. Each layer of sucrose solution was 6 mL. The capsid sample (1 mL) and was loaded onto the top of the gradient and separated at 140,000 × *g* for 3 h at 4 °C in SW-32Ti Beckman Coulter swinging bucket rotor. Sucrose gradient fractions containing Cp140 capsid were analyzed by electron microscopy, and fractions containing mainly T = 3 capsids were pooled and dialyzed against purification buffer at 4 °C.

Electron Microscopy. For each capsid preparation, 5 μL capsids concentrated at 1 mg/mL were deposited and incubated for 2 min on a carbon-coated grid. After removing solution excess with absorbent paper, the negative staining was performed by layering the grid on top of 50 μL 2% phosphotungstic acid (weight per volume) at pH 7. After 2 min incubation, the solution excess was removed with absorbent paper. The grids were examined with a JEM-1400 transmission electron microscope operating at 100 kV.

NMR Rotor Filling. After dialysis in purification buffer, soluble capsids isolated from *E. coli* and in vitro reassembled capsids were concentrated using Amicon Ultra centrifugal filter units (Merck, 50 kDa cutoff) to about 20 mg/mL in 1 mL and sedimented into 3.2 mm zirconium rotors by centrifugation (200,000 × *g*, 14 h, and 4 °C) using a homemade filling tool (36). For ^{1}H-detected spectra, <1 mg capsids was sedimented into 0.7 mm rotors by centrifugation (200,000 × *g*, 16 h, and 4 °C). Rotors were immediately closed with watertight caps after the addition of 1 μl saturated sodium trimethylsilylpropanesulfonate (DSS) solution for chemical shift referencing. For 0.7 mm rotors, 15 to 20 μL saturated DSS was added to the protein solution before sedimentation.

Solid-State NMR Spectroscopy. All ^{13}C-detected NMR spectra were recorded using a 3.2 mm triple-resonance (^{1}H, ^{13}C, and ^{15}N) probe head at a static magnetic field of 18.8 T, corresponding to 800 MHz proton resonance frequency (Bruker Avance II). Assignment of the A state was described previously and was performed on Cp149 capsids isolated from *E. coli* (43). Similarly, assignment of the conformation B was obtained at a magic-angle spinning (MAS) frequency of 17.5 kHz using a combination of two-dimensional (2D) and 3D correlation experiments, including 2D Dipolar Assisted Rotation Resonance (DARR), nitrogen-alpha carbon (NCA), nitrogen-carboxyl (NCO), 3D nitrogen-alpha carbon-beta carbon (NCACB), nitrogen-alpha carbon-side chain carbon (NCACX), nitrogen-carboxyl-side chain carbon (NCOCX), alpha carbon-nitrogen-carboxyl (CANCO), and all carbon (CCC) correlation experiments, which were performed on the sample of P7-Cp183 capsids reassembled in vitro. On all the other *E. coli* and reassembled samples, only 2D DARR were recorded. On some samples, additional NCA and NCO were also recorded. A 3D NCACX was obtained on the sample of *E. coli* Cp149 capsids stored for 1 mo at room temperature in order to unambiguously identify the resonances of oxidized C61, and 3D CANCO were recorded on Cp183-F97L and Cp-GenotypeG capsids isolated from *E. coli* as the mutations induced chemical shift changes. Experimental details are given in *SI Appendix*, Tables S3 and S4. All spectra were referenced to DSS and recorded at a sample temperature of 4 °C according to the resonance frequency of the supernatant water (36).

For ^{1}H-detection, 2D fingerprint hNH spectra and 3D proton-alpha carbon-nitrogen (hCANH) spectra (66) were recorded on ^{13}C-^{15}N Cp149 capsids isolated from *E. coli* and in vitro reassembled capsids prepared as described above and on ^{13}C-^{15}N cell-free synthesized CF-Cp183 capsids prepared as described in ref. 45. The RNA packaged in the cell-free capsids is not isotope labeled. Experimental details can be found in refs. 42 and 45. Briefly, spectra were acquired on a wide-bore 850 MHz Bruker Avance III spectrometer with a 0.7 mm triple-resonance MAS probe (Bruker Biospin), operated at 100 kHz MAS with a sample temperature of ~22 °C, and referenced to DSS. All spectra were processed using TopSpin 4.0.3 (Bruker Biospin) and analyzed with the CcpNmr Analysis package, version 2.4.2 (67, 68).

^{31}P CP-MAS were recorded at 17 kHz on a 500 MHz Bruker spectrometer on the same rotors as the ones used for the ^{13}C DARR spectra. Experimental details are given in *SI Appendix*, Table S1.

NMR Data Analysis and PyMoL Mutagenesis Model Building. The percentage of each conformation was calculated based on the intensity of six isolated peaks corresponding to P5 Cγ-Cδ, P5 Cγ-Cα, and I59 Cδ1-Cβ resonances for which the distinction between the two conformations A and B was clearly visible (leading to two peaks per resonance). If the signal was nonexistent, a symbolic peak was added at the virtual position of the missing conformation. For P5W and L60W mutants, L100 Cβ-Cα and A58 Cβ-Cα resonances were used instead of the two P5 resonances. The peak heights were fitted using the Gaussian fit function of CcpNmr software (67, 68), and the percentage of conformation B was calculated according to

$$\textit{Conformation B}\,(\%) = \frac{1}{n}\sum_{i=1}^{n}\frac{I_B^i}{I_A^i + I_B^i} \times 100,$$

where n denotes the number of cross peaks analyzed (three here), I_A^i is the intensity of the i'th peak for conformation A, and I_B^i is for conformation B for each of the three resonances. The statistical error was calculated as the SD of conformation B percentage between the three resonances.

CSDs between the conformation A and B were calculated for each carbon atoms according to the following equation: $\Delta\delta_C = |\delta_C[A] - \delta_C[B]|$. All assigned carbons, including sidechains, were used for the CSD graphs.

For the prediction of W5 and W60 sidechains position in P5W and L60W mutants, the starting structure was according to the Protein Data Bank access code 1QGT, where P5 and L60 residues were replaced by tryptophan residues using the mutagenesis tool included in PyMoL (Schrödinger, LLC). The rotamers with the least steric clashes were chosen.

Extraction of Triton X-100 from the Protein and Solution NMR Spectroscopy. Chloroform/methanol precipitation was performed as described in *SI Appendix*, Fig. S8 caption. Extracted chloroform phases were transferred into 5 mm solution NMR tubes and ^{1}H-^{13}C heteronuclear single quantum coherence (HSQC) spectra, using sensitivity improvement (69), were recorded on the three samples on a 600 MHz spectrometer (Bruker Avance II), equipped with a triple-resonance cryo-probe. NMR acquisition parameters are detailed in *SI Appendix*, Table S5. ^{1}H-^{13}C NMR chemical shifts of Triton X-100 in chloroform were assigned by comparison with assignments from the literature (70) and are shown in *SI Appendix*, Table S6.

ITC Experiments. Experimental details are given in *SI Appendix*, Fig. S9 caption.

Cell Culture Work. See *SI Appendix, Extended Methods*.

Data Availability. The ^{13}C and ^{15}N backbone and sidechain chemical shifts of the core protein of HBV capsids in conformation B have been deposited in the BioMagResBank (http://www.bmrb.wisc.edu/) under accession code 28122 (72).

ACKNOWLEDGMENTS. We acknowledge Carla Osiowy and Tim Booth (National Microbiology Laboratory, Winnipeg, Canada) for the kind gift of the plasmid of Cp genotype G and Xavier Hanoulle (Université de Lille) for helpful discussions. This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie Grant Agreement 748516. This work was also supported by the French Agence Nationale de Recherches sur le Sida et les hépatites virales (ANRS, ECTZ71388 and ECTZ100488), the French Agence Nationale de Recherche (ANR-14-CE09-0024B), the CNRS (CNRS-Momentum 2018), the Laboratoire d'Excellence (LABEX) Dynamiques Éco-Évolutives des Maladies Infectieuses (ECOFECT) (ANR-11-LABX-0048) within the Université de Lyon program Investissements d'Avenir (ANR-11-IDEX-0007), the Swiss National Science Foundation (Grant 200020_159707), and a grant from the Chinese Scholarship Council to S.W. Financial support from the Infrastructure de Recherche décentralisée RMN Très Hauts Champs (IR-RMN-THC) Fr3050 CNRS for conducting the research is gratefully acknowledged. We thank the Centre d'Imagerie Quantitative Lyon-Est for support at the EM platform.

1. I. Graber-Stiehl, The silent epidemic killing more people than HIV, malaria or TB. *Nature* **564**, 24–26 (2018).
2. M. Nassal, HBV cccDNA: Viral persistence reservoir and key obstacle for a cure of chronic hepatitis B. *Gut* **64**, 1972–1984 (2015).
3. C. Seeger, W. S. Mason, Molecular biology of hepatitis B virus infection. *Virology* **479-480**, 672–686 (2015).
4. S. Seitz, J. Habjanič, A. K. Schütz, R. Bartenschlager, The hepatitis B virus envelope proteins: Molecular gymnastics throughout the viral life cycle. *Annu. Rev. Virol.* **7**, 263–288 (2020).
5. A. Zlotnick *et al.*, Core protein: A pleiotropic keystone in the HBV lifecycle. *Antiviral Res.* **121**, 82–93 (2015).
6. C. M. Livingston, D. Ramakrishnan, M. Strubin, S. P. Fletcher, R. K. Beran, Identifying and characterizing interplay between hepatitis B virus X protein and smc5/6. *Viruses* **9**, 69 (2017).
7. F. Birnbaum, M. Nassal, Hepatitis B virus nucleocapsid assembly: Primary structure requirements in the core protein. *J. Virol.* **64**, 3319–3330 (1990).
8. M. Nassal, 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).
9. R. A. Crowther *et al.*, Three-dimensional structure of hepatitis B virus core particles determined by electron cryomicroscopy. *Cell* **77**, 943–950 (1994).
10. M. Nassal, Hepatitis B viruses: Reverse transcription a different way. *Virus Res.* **134**, 235–249 (2008).
11. J. Heger-Stevic, P. Zimmermann, L. Lecoq, B. Böttcher, M. Nassal, Hepatitis B virus core protein phosphorylation: Identification of the SRPK1 target sites and impact of their occupancy on RNA binding and capsid structure. *PLoS Pathog.* **14**, e1007488 (2018).
12. J. Hu, K. Liu, Complete and incomplete hepatitis B virus particles: Formation, function, and application. *Viruses* **9**, 56 (2017).
13. F. Poisson, A. Severac, C. Hourioux, A. Goudeau, P. Roingeard, Both pre-S1 and S domains of hepatitis B virus envelope proteins interact with the core particle. *Virology* **228**, 115–120 (1997).

14. H. Löffler-Mary, J. Dumortier, C. Klentsch-Zimmer, R. Prange, Hepatitis B virus assembly is sensitive to changes in the cytosolic S loop of the envelope proteins. *Virology* **270**, 358–367 (2000).
15. B. Böttcher *et al*., Peptides that block hepatitis B virus assembly: Analysis by cryomicroscopy, mutagenesis and transfection. *EMBO J.* **17**, 6839–6845 (1998).
16. D. Ponsel, V. Bruss, Mapping of amino acid side chains on the surface of hepatitis B virus capsids required for envelopment and virion formation. *J. Virol.* **77**, 416–422 (2003).
17. S. Le Pogam, C. Shih, Influence of a putative intermolecular interaction between core and the pre-S1 domain of the large envelope protein on hepatitis B virus secretion. *J. Virol.* **76**, 6510–6517 (2002).
18. S. Seitz *et al*., A slow maturation process renders hepatitis B virus infectious. *Cell Host Microbe* **20**, 25–35 (2016).
19. X. Ning *et al*., Secretion of genome-free hepatitis B virus–Single strand blocking model for virion morphogenesis of para-retrovirus. *PLoS Pathog.* **7**, e1002255 (2011).
20. L. Luckenbaugh, K. M. Kitrinos, W. E. Delaney IV, J. Hu, Genome-free hepatitis B virion levels in patient sera as a potential marker to monitor response to antiviral therapy. *J. Viral Hepat.* **22**, 561–570 (2015).
21. J. Summers, W. S. Mason, Replication of the genome of a hepatitis B–Like virus by reverse transcription of an RNA intermediate. *Cell* **29**, 403–415 (1982).
22. B. Ning, C. Shih, Nucleolar localization of human hepatitis B virus capsid protein. *J. Virol.* **78**, 13653–13668 (2004).
23. T. T. Yuan, G. K. Sahu, W. E. Whitehead, R. Greenberg, C. Shih, The mechanism of an immature secretion phenotype of a highly frequent naturally occurring missense mutation at codon 97 of human hepatitis B virus core antigen. *J. Virol.* **73**, 5731–5740 (1999).
24. X. Ning *et al*., Common and distinct capsid and surface protein requirements for secretion of complete and genome-free hepatitis B virions. *J. Virol.* **92**, e00272-18 (2018).
25. J. F. Conway *et al*., Visualization of a 4-helix bundle in the hepatitis B virus capsid by cryo-electron microscopy. *Nature* **386**, 91–94 (1997).
26. B. Böttcher, M. Nassal, Structure of mutant hepatitis B core protein capsids with premature secretion phenotype. *J. Mol. Biol.* **430**, 4941–4954 (2018).
27. J. Z. Porterfield *et al*., Full-length hepatitis B virus core protein packages viral and heterologous RNA with similarly high levels of cooperativity. *J. Virol.* **84**, 7174–7184 (2010).
28. B. Böttcher, S. A. Wynne, R. A. Crowther, Determination of the fold of the core protein of hepatitis B virus by electron cryomicroscopy. *Nature* **386**, 88–91 (1997).
29. S. A. Wynne, R. A. Crowther, A. G. W. Leslie, The crystal structure of the human hepatitis B virus capsid. *Mol. Cell* **3**, 771–780 (1999).
30. X. Yu, L. Jin, J. Jih, C. Shih, Z. H. Zhou, 3.5Å cryoEM structure of hepatitis B virus core assembled from full-length core protein. *PLoS One* **8**, e69729 (2013).
31. A. Zlotnick *et al*., Dimorphism of hepatitis B virus capsids is strongly influenced by the C-terminus of the capsid protein. *Biochemistry* **35**, 7412–7421 (1996).
32. N. R. Watts *et al*., The morphogenic linker peptide of HBV capsid protein forms a mobile array on the interior surface. *EMBO J.* **21**, 876–884 (2002).
33. M. Nassal, A. Rieger, O. Steinau, Topological analysis of the hepatitis B virus core particle by cysteine-cysteine cross-linking. *J. Mol. Biol.* **225**, 1013–1025 (1992).
34. C. M. Quinn *et al*., Magic angle spinning NMR of viruses. *Prog. Nucl. Magn. Reson. Spectrosc.* **86-87**, 21–40 (2015).
35. L. Lecoq, M.-L. Fogeron, B. H. Meier, M. Nassal, A. Böckmann, Solid-state NMR for studying the structure and dynamics of viral assemblies. *Viruses* **12**, 1069 (2020).
36. A. Böckmann *et al*., Characterization of different water pools in solid-state NMR protein samples. *J. Biomol. NMR* **45**, 319–327 (2009).
37. I. Bertini *et al*., Solid-state NMR of proteins sedimented by ultracentrifugation. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 10396–10399 (2011).
38. C. Gardiennet *et al*., A sedimented sample of a 59 kDa dodecameric helicase yields high-resolution solid-state NMR spectra. *Angew. Chem. Int. Ed. Engl.* **51**, 7855–7858 (2012).
39. T. Wiegand *et al*., Sedimentation yields long-term stable protein samples as shown by solid-state NMR. *Front. Mol. Biosci.* **7**, 17 (2020).
40. M. P. Williamson, Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **73**, 1–16 (2013).
41. L. Lecoq *et al*., Localizing conformational hinges by NMR: Where do hepatitis B virus core proteins adapt for capsid assembly? *ChemPhysChem* **19**, 1336–1340 (2018).
42. L. Lecoq *et al*., 100 kHz MAS proton-detected NMR spectroscopy of hepatitis B virus capsids. *Front. Mol. Biosci.* **6**, 58 (2019).
43. L. Lecoq *et al*., Solid-state [^{13}C-^{15}N] NMR resonance assignment of hepatitis B virus core protein. *Biomol. NMR Assign.* **12**, 205–214 (2018).
44. J. J. H. Cotelesage *et al*., Hepatitis B Virus Genotype G forms core-like particles with unique structural properties. *J. Viral Hepat.* **18**, 443–448 (2011).
45. S. Wang *et al*., Combining cell-free protein synthesis and NMR into a tool to study capsid assembly modulation. *Front. Mol. Biosci.* **6**, 67 (2019).
46. D. Sun, M. Nassal, Stable HepG2- and Huh7-based human hepatoma cell lines for efficient regulated expression of infectious hepatitis B virus. *J. Hepatol.* **45**, 636–645 (2006).
47. B. Li *et al*., Suppression of hepatitis B virus antigen production and replication by wild-type HBV dependently replicating HBV shRNA vectors in vitro and in vivo. *Antiviral Res.* **134**, 117–129 (2016).
48. T. Döring, R. Prange, Rab33B and its autophagic Atg5/12/16L1 effector assist in hepatitis B virus naked capsid formation and release. *Cell. Microbiol.* **17**, 747–764 (2015).
49. U. Viswanathan *et al*., Targeting the multifunctional HBV core protein as a potential cure for chronic hepatitis B. *Antiviral Res.* **182**, 104917 (2020).
50. C. R. Bourne, M. G. Finn, A. Zlotnick, Global structural changes in hepatitis B virus capsids induced by the assembly effector HAP1. *J. Virol.* **80**, 11055–11061 (2006).
51. W. S. Tan, I. W. McNae, K. L. Ho, M. D. Walkinshaw, Crystallization and X-ray analysis of the T = 4 particle of hepatitis B capsid protein with an N-terminal extension. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **63**, 642–647 (2007).
52. S. A. Wynne, A. G. W. Leslie, P. J. G. Butler, R. A. Crowther, Crystallization of hepatitis B virus core protein shells: Determination of cryoprotectant conditions and preliminary X-ray characterization. *Acta Crystallogr. D Biol. Crystallogr.* **55**, 557–560 (1999).
53. A. M. Roseman, J. A. Berriman, S. A. Wynne, P. J. G. Butler, R. A. Crowther, A structural model for maturation of the hepatitis B virus core. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15821–15826 (2005).
54. J. Hayer *et al*., HBVdb: A knowledge database for hepatitis B virus. *Nucleic Acids Res.* **41**, D566–D570 (2013).
55. S. Le Pogam, T. T. Yuan, G. K. Sahu, S. Chatterjee, C. Shih, Low-level secretion of human hepatitis B virus virions caused by two independent, naturally occurring mutations (P5T and L60V) in the capsid protein. *J. Virol.* **74**, 9099–9105 (2000).
56. V. Bruss, Envelopment of the hepatitis B virus nucleocapsid. *Virus Res.* **106**, 199–209 (2004).
57. P. Gripon, J. Le Seyec, S. Rumin, C. Guguen-Guillouzo, Myristylation of the hepatitis B virus large surface protein is essential for viral infectivity. *Virology* **213**, 292–299 (1995).
58. V. Bruss, J. Hagelstein, E. Gerhardt, P. R. Galle, Myristylation of the large surface protein is required for hepatitis B virus in vitro infectivity. *Virology* **218**, 396–399 (1996).
59. C. A. Nelson *et al*., Structural basis for murine norovirus engagement of bile acids and the CD300lf receptor. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E9201–E9210 (2018).
60. C. Toelzer *et al*., Free fatty acid binding pocket in the locked structure of SARS-CoV-2 spike protein. *Science* **370**, 725–730 (2020).
61. M. B. Sherman, H. Q. Smith, T. J. Smith, The dynamic life of virus capsids. *Viruses* **12**, 618–621 (2020).
62. R. Abdelnabi *et al*., A novel druggable interprotomer pocket in the capsid of rhino- and enteroviruses. *PLoS Biol.* **17**, e3000281 (2019).
63. M. S. Dhason, J. C. Y. Wang, M. F. Hagan, A. Zlotnick, Differential assembly of Hepatitis B Virus core protein on single- and double-stranded nucleic acid suggest the dsDNA-filled core is spring-loaded. *Virology* **430**, 20–29 (2012).
64. C. Königer *et al*., Involvement of the host DNA-repair enzyme TDP2 in formation of the covalently closed circular DNA persistence reservoir of hepatitis B viruses. *Proc. Natl. Acad. Sci. U.S.A.* **111**, E4244–E4253 (2014).
65. W. S. Tan, M. R. Dyson, K. Murray, Two distinct segments of the hepatitis B virus surface antigen contribute synergistically to its association with the viral core particles. *J. Mol. Biol.* **286**, 797–808 (1999).
66. S. Penzel *et al*., Protein resonance assignment at MAS frequencies approaching 100 kHz: A quantitative comparison of J-coupling and dipolar-coupling-based transfer methods. *J. Biomol. NMR* **63**, 165–186 (2015).
67. W. F. Vranken *et al*., The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins* **59**, 687–696 (2005).
68. T. J. Stevens *et al*., A software framework for analysing solid-state MAS NMR data. *J. Biomol. NMR* **51**, 437–447 (2011).
69. A. G. Palmer III, W. J. Fairbrother, J. Cavanagh, P. E. Wright, M. Rance, Improved resolution in three-dimensional constant-time triple resonance NMR spectroscopy of proteins. *J. Biomol. NMR* **2**, 103–108 (1992).
70. V. G. Rao, S. Mandal, S. Ghosh, C. Banerjee, N. Sarkar, Aggregation behavior of Triton X-100 with a mixture of two room-temperature ionic liquids: Can we identify the mutual penetration of ionic liquids in ionic liquid containing micellar aggregates? *J. Phys. Chem. B* **116**, 13868–13877 (2012).
71. C. T. Yeh, Y. F. Liaw, J. H. Ou, The arginine-rich domain of hepatitis B virus precore and core proteins contains a signal for nuclear transport. *J. Virol.* **64**, 6141–6147 (1990).
72. L. Lecoq *et al*., data_28122. *Biological Magnetic Resonance Data Bank*. https://bmrb.io/data_library/summary/index.php?bmrbId=28122. Deposited 15 May 2020.