Author/s:
Taji, ZH;Bielytskyi, P;Shein, M;Sani, MA;Seitz, S;Schütz, AK

Title:
Transient RNA Interactions Leave a Covalent Imprint on a Viral Capsid Protein

Date:
2022-05-18

Citation:
Taji, Z. H., Bielytskyi, P., Shein, M., Sani, M. A., Seitz, S. & Schütz, A. K. (2022). Transient RNA Interactions Leave a Covalent Imprint on a Viral Capsid Protein. Journal of the American Chemical Society, 144 (19), pp.8536-8550. https://doi.org/10.1021/JACS.1C12439.

Persistent Link:
https://hdl.handle.net/11343/331903

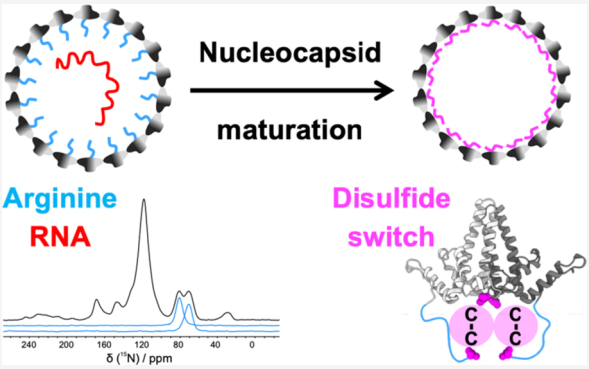
JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
pubs.acs.org/JACS
Article

Transient RNA Interactions Leave a Covalent Imprint on a Viral Capsid Protein

Zahra Harati Taji,[#] Pavlo Bielytskyi,[#] Mikhail Shein, Marc-Antoine Sani,* Stefan Seitz,* and Anne K. Schütz*

Cite This: J. Am. Chem. Soc. 2022, 144, 8536–8550
Read Online
ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The hepatitis B virus (HBV) is the leading cause of persistent liver infections. Its DNA-based genome is synthesized through reverse transcription of an RNA template inside the assembled capsid shell. In addition to the structured assembly domain, the capsid protein harbors a C-terminal extension that mediates both the enclosure of RNA during capsid assembly and the nuclear entry of the capsid during infection. The arginine-rich motifs within this extension, though common to many viruses, have largely escaped atomic-scale investigation. Here, we leverage solution and solid-state nuclear magnetic resonance spectroscopy at ambient and cryogenic temperatures, under dynamic nuclear polarization signal enhancement, to investigate the organization of the genome within the capsid. Transient interactions with phosphate groups of the RNA backbone confine the arginine-rich motifs to the interior capsid space. While no secondary structure is induced in the C-terminal extension, interactions with RNA counteract the formation of a disulfide bond, which covalently tethers this peptide arm onto the inner capsid surface. Electrostatic and covalent contributions thus compete in the spatial regulation of capsid architecture. This disulfide switch represents a coupling mechanism between the structured assembly domain of the capsid and the enclosed nucleic acids. In particular, it enables the redox-dependent regulation of the exposure of the C-terminal extension on the capsid surface, which is required for nuclear uptake of the capsid. Phylogenetic analysis of capsid proteins from hepadnaviruses points toward a function of this switch in the persistence of HBV infections.



INTRODUCTION

The capsid, or protein shell, of viruses packages and encloses the genome during virion formation and protects it until its release. Many viruses encode capsid proteins which assemble into capsid shells with positively charged amino acid residues lining the interior. These arginine or lysine residues often cluster in the termini as basic segments protruding into the interior capsid space. Many single-stranded (ss) RNA viruses harbor flexible peptide arms for this purpose,[1−3] which may in turn contain one or several arginine-rich motifs (ARMs). Pioneering theoretical studies have postulated that nonspecific electrostatic interactions fix the ratio of negative charges of the nucleotides to positive charges of the peptide and determine the localization of the genome inside the nucleocapsid.[4]

Atomic-level investigations of peptide arms and associated nucleic acids within intact nucleocapsids are experimentally challenging. In cryo-electron microscopy (cryo-EM), the resolution is limited by the substantial conformational heterogeneity and the difference in molecular symmetry between the capsid shell and content.[1] Molecular weights of many megadaltons preclude using solution-state nuclear magnetic resonance (NMR), except for very long disordered extensions, where motional freedom independent of the capsid is recovered in the termini. Alternatively, segments truncated from the capsid scaffold serve as proxies.[5−7]

A prototypical nucleocapsid with flexible peptide arms is found in human hepatitis B virus (HBV). The central component of its virion is an icosahedral capsid shell, which carries a partially double-stranded (ds) DNA and is enclosed by an outer lipid membrane with embedded envelope glycoproteins (Figure 1a). The capsid is built from 120 homodimers of the hepatitis B core (C) protein, which occupy quasi-equivalent environments on a symmetrical lattice to yield a five megadalton assembly. The capsid shell is punctuated by pores enabling the passage of small molecules, notably nucleotides in vivo. The C protein in genotype D of human HBV consists of 183 residues (Cp183) with an N-terminal

Received: November 25, 2021
Published: May 5, 2022

https://doi.org/10.1021/jacs.1c12439
J. Am. Chem. Soc. 2022, 144, 8536−8550

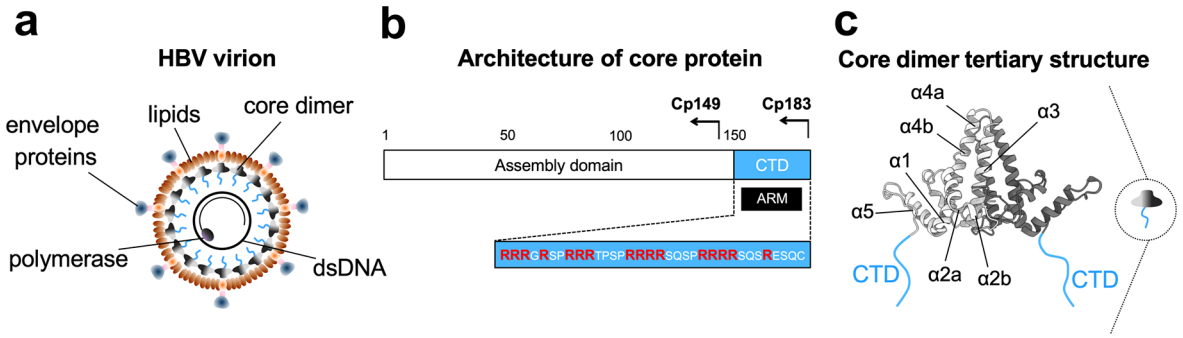


Figure 1. Build-up of the HBV virion. (a) Schematic of the HBV virion. The particle comprises a capsid shell built of core (C) protein homodimers, accommodating circular, partially dsDNA together with a polymerase. It is embedded into a layer of envelope proteins and lipids. (b) C protein consists of 183 amino acid residues and is divided into two parts: The 149 N-terminal residues form the assembly domain, while the last 34 residues make up the CTD and feature nucleic-acid binding ARMs. (c) Assembly domain of the C protein comprises five α-helices connected by loops,[17] whereas the structure of the CTD remains elusive.

assembly domain of predominantly α-helical fold.[8] At the C-terminus, it harbors a nucleic-acid binding domain, called C-terminal domain (CTD), which is around 30 amino acids in length with four arginine-rich repeats interspersed with glycine, proline, serine, glutamine, and threonine residues (Figure 1b,c). This CTD resembles the peptide arms found in many ssRNA viruses.[2,3]

Indeed, the dsDNA genome of the virus matures inside the nucleocapsid from a ss pregenomic RNA template (pgRNA) via reverse transcription. The synthesis of the dsDNA genome and subsequent envelopment of the mature nucleocapsid by the viral membrane are precisely regulated processes. The finding that dsDNA-containing nucleocapsids and empty capsids are more readily enveloped than ssRNA-containing ones sparked the 'maturation signal' hypothesis in which the state of the viral genome is coupled to a binary structural switch within the capsid.[9−12] Mature nucleocapsids can enter the nucleus, where they deposit the viral dsDNA. This process is mediated by a nuclear localization signal, which is located within the CTD and recognized by importin complexes, enabling passage through the nuclear pore complex.[13] Because the CTD has this dual functionality—regulating nucleic acid packaging inside the capsid and presenting signals to the host cell by transient exposure on the capsid exterior—stringent regulation of its localization during the HBV lifecycle is essential.

Although the ∼3200-base pair ssRNA template contains specific packaging motifs which mediate recognition by viral proteins during capsid assembly,[14,15] the interactions with the peptide arms are largely independent of the RNA sequence.[4,16] This leads to the question how numerous unspecific contacts between a given nucleic acid sequence and the CTD combine to enable such stringent regulation of its dual function. Structural insight into the CTD has been limited so far: it is deleted in many studies (capsids assembled from C-terminally truncated C proteins, Cp149[17,18]) or not explicitly modeled into cryo-EM densities, together with nucleic acids, which are detected as a smear in the above-nanometer range.[8,18,19] Many properties of the CTD remain unknown, namely: (i) what determines its propensity for structure − if any − *versus* disorder, (ii) whether disorder is dynamic or static, (iii) what regulates its enclosure inside the capsid *versus* exposure on the surface, and (iv) whether or how it signals the maturation state of the enclosed viral genome to the outer capsid shell.

Here, we address these questions by solution and magic-angle spinning (MAS) solid-state (ss) NMR methodologies geared toward atomic-level investigations of large, heterogeneous viral assemblies. These have proven invaluable in the characterization of viral capsid and envelope proteins, notably HIV,[20−22] influenza A,[23] SARS-CoV-2,[24] or bacteriophages.[20,25,26]

We were able to probe the conformation and mobility of the CTD by dynamic spectral editing and localize interaction sites with nucleic acids via polarization transfer. We discovered a conformational switch in the HBV capsid realized by a disulfide bond that can lock the CTD arm onto the capsid's inner surface, subject to the redox environment and the state of viral genome synthesis. Phylogenetic analysis of C proteins from the family of *Hepadnaviridae* family identified steps in the viral life cycle which could potentially exploit this disulfide-switching mechanism.

RESULTS

Transient Interactions between the RNA Phosphate Backbone and ARMs in the CTD. To characterize the organization of RNA inside the capsids, we expressed the full-length C protein (Cp183, genotype D) in *Escherichia coli* and purified the self-assembled capsids from bacterial cell lysates. Although bacterial RNA lacks the specific packaging signal of the HBV pgRNA template, it is efficiently packaged by the C protein dimers during particle assembly.[16] In this way, capsids and nucleic acids were both uniformly isotopically labeled with ^{15}N and ^{13}C. We analyzed these capsids using ssNMR experiments under MAS utilizing magnetization transfers that rely on dipolar couplings and spin diffusion (PDSD,[27] DARR,[28] NCa,[29] and NhhC[30]) to visualize the rigid capsid assembly while filtering out small mobile molecules and flexible protein or RNA segments, which escape the applied dipolar recoupling methods because of their rapid motion. The obtained spectral resolution was similar to previous studies of Cp149 capsids[31,32] (assembly domain only) while the reported polymorphism ascribed to the use of detergent during purification[33] was circumvented (Materials and Methods Section). With this experimental setup, we expect correlations which identify through-space contacts between nuclear spins in the range of 7−10 Å.

At ambient temperature in the hydrated protein sediment, signals of both the capsid amino acids and ribose (Figure 2a) were detected, confirming the presence of RNA inside the

https://doi.org/10.1021/jacs.1c12439
J. Am. Chem. Soc. 2022, 144, 8536−8550

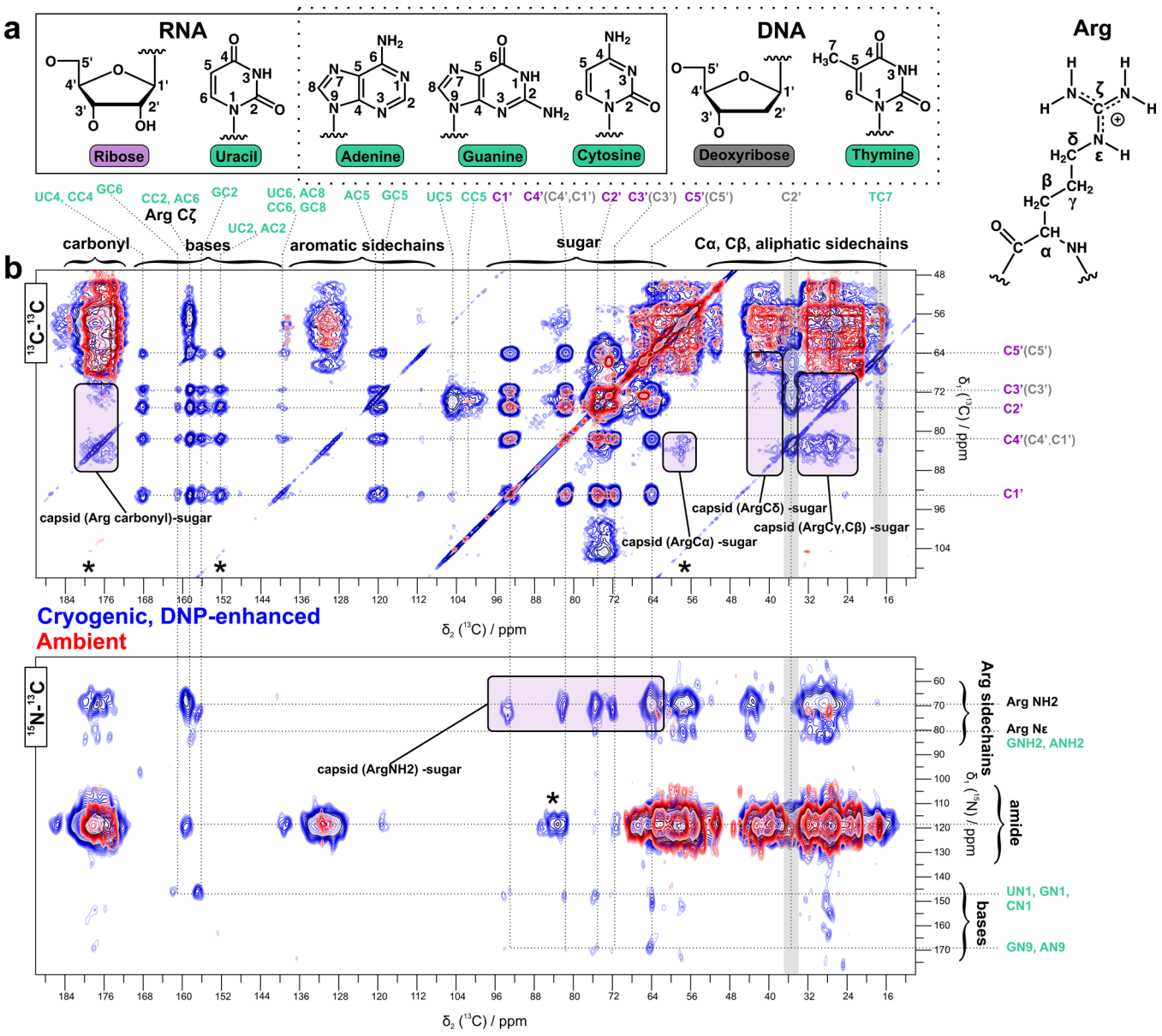


Figure 2. Nucleic acids remain mobile inside the capsid. (a) RNA and DNA bases, sugar moieties, and arginine residues with respective nomenclatures. (b) Top: Zoom of DARR spectra (500 ms mixing); bottom: NhhC spectra (500 μs mixing) acquired on self-assembled Cp183 capsids containing nucleic acids, both uniformly labeled with ^{13}C and ^{15}N, at cryogenic temperature (108 K), B_0 = 18.8 T and MAS frequency ~9.4 kHz with DNP signal enhancement (blue) and at ambient temperature (14 °C) and MAS frequency 15 kHz (red). Correlation peaks unambiguously assigned to capsid-RNA contacts are highlighted in purple; the traces through C2' resonance of deoxyribose and C7 of thymine are highlighted in gray. Note that both capsid-sugar and sugar-base contacts appear solely under cryogenic conditions. Asterisks mark spinning sidebands in all figures. Further details of DNP experiments in Figures S1,2 and Table S2.

capsids (Figure 2b, red contours). However, the signals of RNA bases (Figure 2a) were not observed at the same sensitivity level, pointing to their higher flexibility as compared to the sugar backbone. Similarly, no signals that can be attributed to the CTD were detected. The ARMs would give rise to very prominent arginine spin systems compared to spectra acquired on Cp149 capsids.

We reasoned that the missing components may escape detection because of residual mobility within the capsid scaffold and moved to cryogenic temperature, about 100 K, where major motions of proteins and nucleic acids are frozen.[20,34] Furthermore, such conditions allow for dynamic nuclear polarization (DNP)-based NMR signal enhancement, which capitalizes on the larger Boltzmann polarization of the electron over the nuclear spin reservoir to achieve hyperpolarization of the nuclear spins. The electron-nucleus polarization transfer requires the presence of a radical polarizing agent, here AMUPol,[35] which we dispersed directly in the protein sediment.[36,37] From a comparison of build-up rates for capsid *versus* enclosed nucleic acids (Figure S1), we deduce that the radical penetrated into the capsid interior, most likely through large pores with a diameter up to ~20 Å in the icosahedral lattice.[8,17] Cryogenic temperatures alone lead to significant line broadening in the NMR spectra,[38,39] and additional line broadening may be caused by the radical via paramagnetic relaxation. The presence of the radical inside the sample did not cause signal bleaching, nor influence the

https://doi.org/10.1021/jacs.1c12439
J. Am. Chem. Soc. 2022, 144, 8536−8550

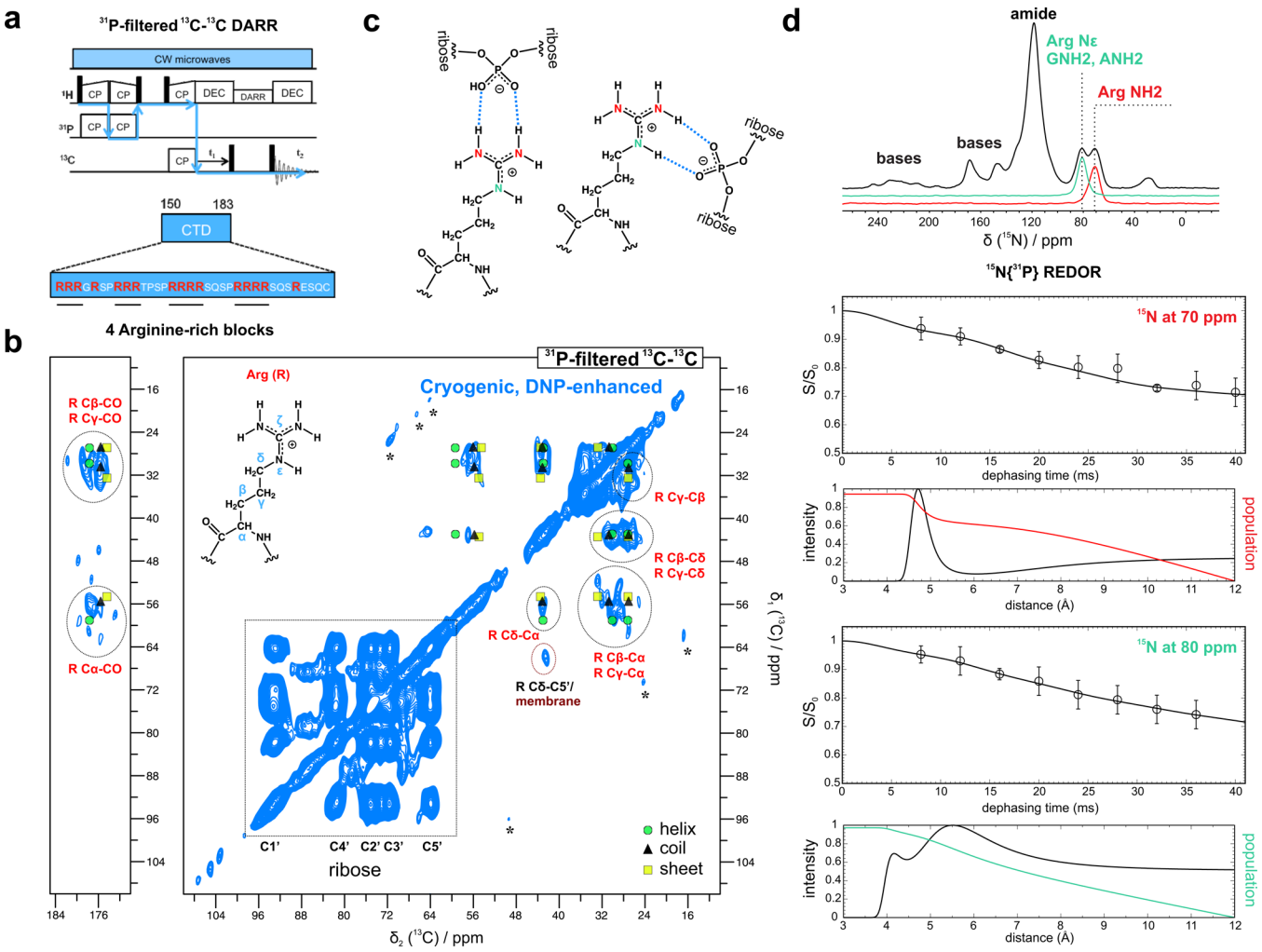


Figure 3. RNA phosphate backbone and ARMs in CTD interact. (a) Top: Pulse sequence to highlight residues in proximity to the phosphate backbone of RNA; black rectangles represent 90° pulses; blue arrows trace magnetization transfers. The sequence filters ^{1}H magnetization via ^{31}P spins by means of two consecutive ^{1}H → ^{31}P and ^{31}P → ^{1}H cross polarization (CP) steps, propagates via proton spin diffusion, and converts to ^{13}C magnetization via ^{1}H → ^{13}C CP, followed by DARR mixing and 2D acquisition, all under DNP signal enhancement. Bottom: 16 arginine residues are distributed throughout the CTD in four blocks. (b) Experiment from panel (a) acquired on capsids assembled from Cp183 filled with RNA (same as Figure 2). Experimental parameters are shown in Table S2 and the full spectrum is shown in Figure S3. Cross peaks in the range ∼60−100 ppm belong to RNA ribose; cross peaks below 60 ppm are compatible with an arginine spin system. Average chemical shift values of arginine residues were derived for α-helices, random coils, and β-sheets.[48] The peak circled in brown can be attributed either to an intermolecular correlation between arginine side chains and ribose or to an intramolecular correlation within the phosphatidyl headgroup of lipid membranes from *E. coli*,[49] traces of which might be present in the sample. A control experiment, however, demonstrated that the NMR contribution from membranes alone does not overlap with spectral regions of interest (Figure S6g). (c) Phosphate-arginine interaction according to the 'arginine fork' motif.[42,50] (d) Top: Comparison of the broad-band excitation ^{1}H-^{15}N CP spectrum of RNA-filled capsids (black) with slices from ^{15}N{^{31}P} J-decoupled REDOR[43] with selective excitation at $\delta(^{15}\mathrm{N})$ 70 ppm (red) and 80 ppm (green). Bottom: BS-REDOR[44] reconstructions yield the respective distance distributions and populations for each signal (full curves Figure S5).

protein structure judging from the spectral fingerprints (Figure S2).

As anticipated, additional signals appeared at cryogenic temperatures. The chemical shifts for all RNA base spin systems can be identified (Figure 2b, blue contours). Furthermore, we could pinpoint unambiguous intermolecular contacts between amino acids of the capsid and RNA, which are mediated by the ARMs in the CTD. Arginine side chains recognize the phosphate-sugar backbone, which is more rigid compared to the flexible bases, for which no contacts to the capsid were detected. In addition, we identified the spectral signature of DNA, which is distinguished by characteristic ^{13}C NMR chemical shifts from C2' of deoxyribose and the C7 atom of the thymine base[40] (Figure 2a,b). Traces of bacterial dsDNA can be incorporated into the capsids during their self-assembly in bacterial cells. Under cryogenic conditions, we did not detect the equivalent of the intermolecular contacts for DNA with arginine residues in the CTD as observed for RNA, indicating that DNA is not able to establish electrostatic interactions to a similar extent. This is corroborated by the finding that the DNA signals do not appear in the ssNMR spectra acquired at ambient temperature, suggesting that both sugar and base moieties of DNA are less immobilized. Thus, our findings are consistent with the theoretically predicted[4] unspecific nature of electrostatic interactions between peptide arms of nucleocapsids and ssRNA.

To confirm our interpretation that phosphate-arginine contacts are the major driving force of the interaction, we

https://doi.org/10.1021/jacs.1c12439
J. Am. Chem. Soc. 2022, 144, 8536−8550

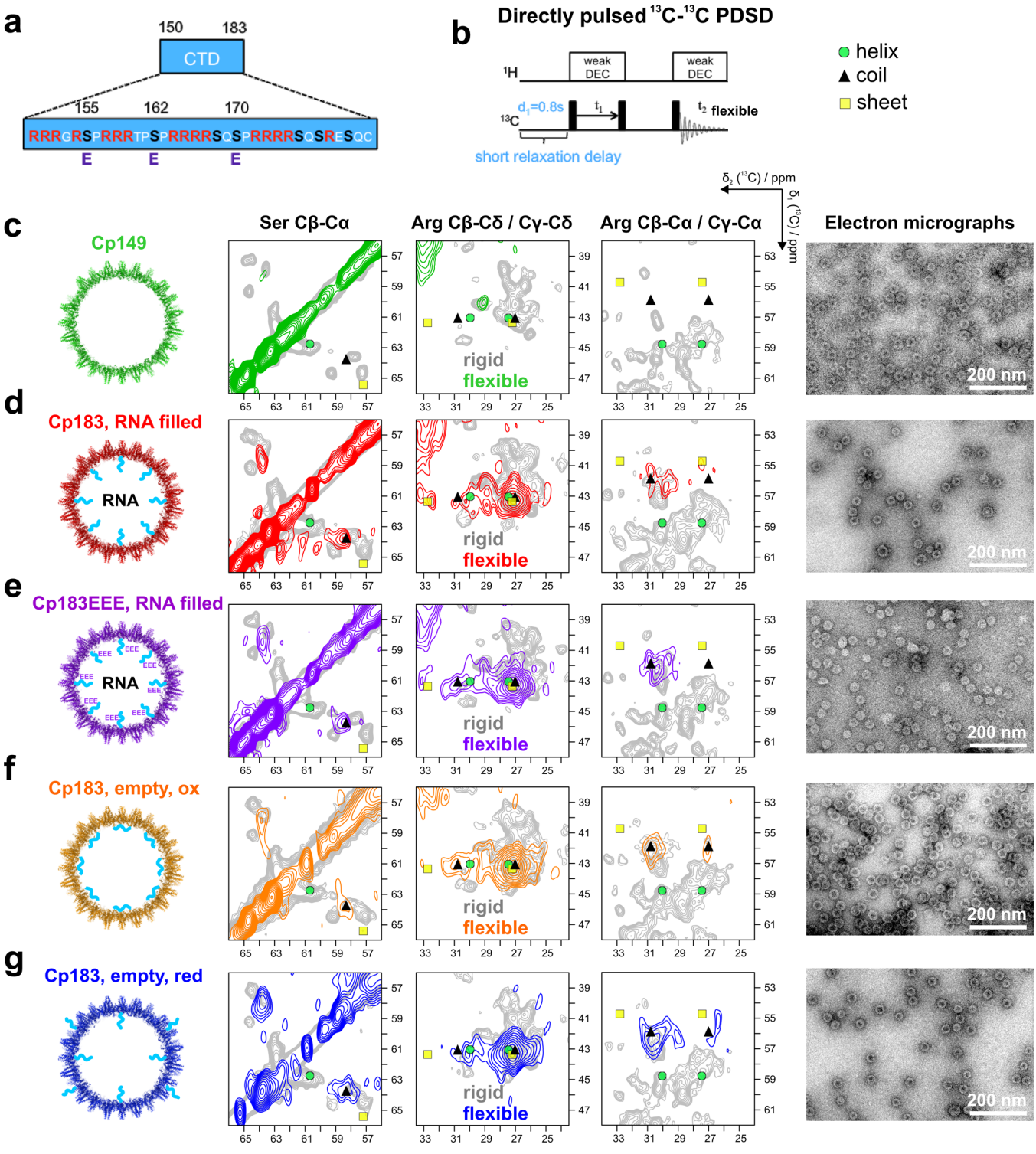


Figure 4. Dynamic regime of the CTD in the context of different capsid types. (a) Arginine distribution and location of serine phosphorylation sites in the CTD. (b) dP PDSD pulse sequence, which relies on differential relaxation properties of spins in dynamic and rigid protein regions. By shortening the relaxation delay and applying weak heteronuclear decoupling, only residues in flexible capsid regions are detected. (c–g) Colored contours: Zooms of Cα-Cβ and side chain regions from dP-PDSD spectra of different capsid types acquired at ambient temperature, highlighting arginine and serine residues in flexible regions. The respective CP-based DARR spectra featuring residues from rigid domains are shown for reference (gray contours). (c) Capsids assembled from Cp149, lacking the CTD; (d) capsids assembled from Cp183 filled with RNA; (e) capsids assembled from Cp183EEE, in which Ser155, Ser162, and Ser170 are mutated to glutamate to mimic phosphorylation, filled with RNA; (f) capsids assembled from Cp183, emptied of RNA, and reassembled in the absence of a reducing agent; (g) capsids assembled from Cp183, emptied of RNA, and reassembled in a reducing environment. Average chemical shift values for α-helices, random coils, and β-sheets of arginine and serine spin systems were derived from statistics obtained at the BMRB.[48] On the right-hand side, negative stain electron micrographs of the respective capsid types are shown. Experimental parameters are shown in Table S3 and full spectra are shown in Figure S6.

https://doi.org/10.1021/jacs.1c12439
J. Am. Chem. Soc. 2022, 144, 8536−8550

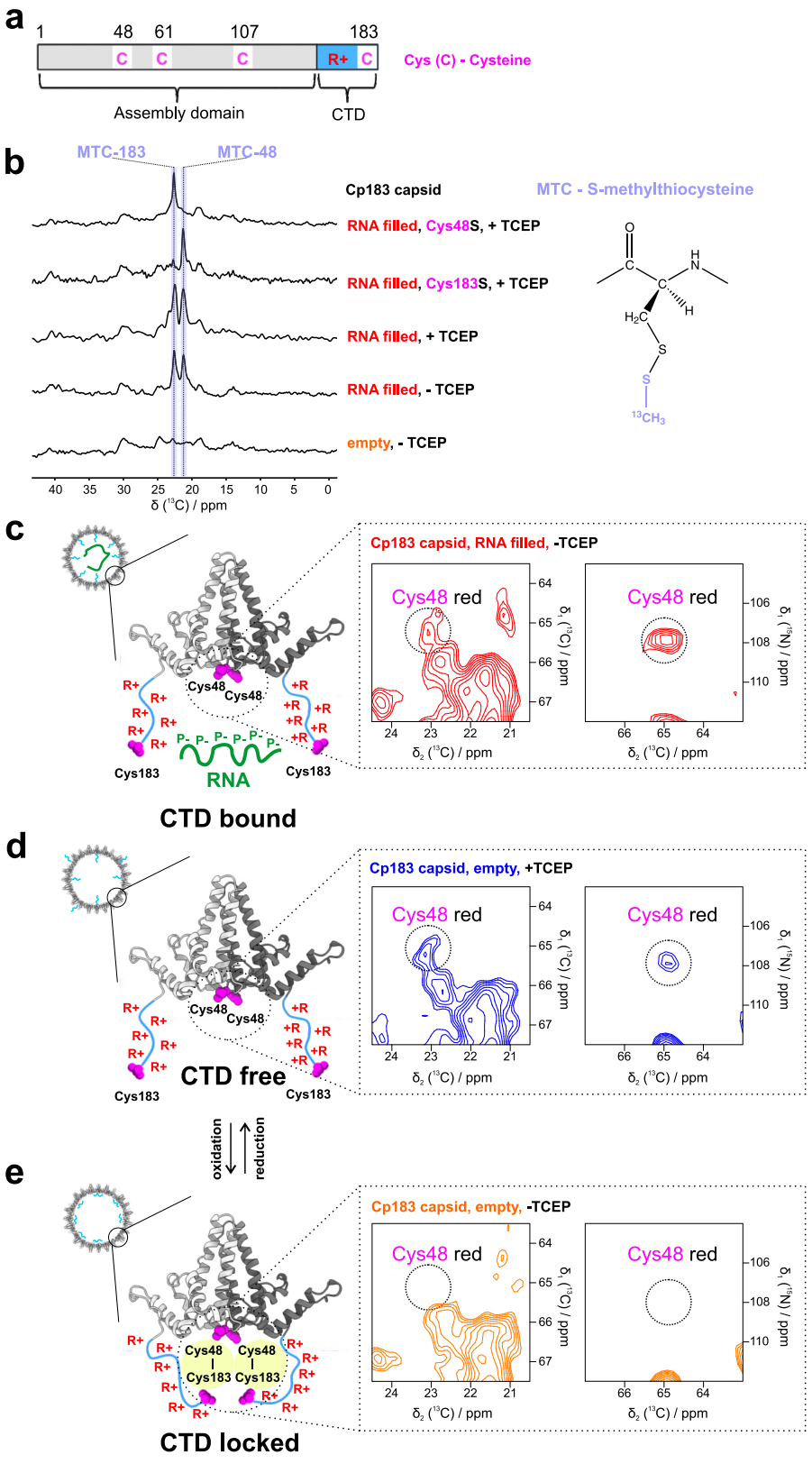


Figure 5. Disulfide forms in empty capsids. (a) Location of cysteine residues in C protein. (b) dp ^{13}C spectra with proton decoupling of capsids labeled with MMTS[60] establish that Cys48 and Cys183 are accessible to oxidation except in RNA-free capsids under oxidative conditions. The corresponding MTC-48 and MTC-183 signals were assigned by mutagenesis (compare upper two spectra). (c–e) Oxidation state of Cys48 as a function of the redox environment and RNA presence inferred from Cα-Cβ and N-Cα correlations. (c) Capsid assembled from Cp183 filled with RNA in the absence of reducing agent (TCEP). (d) Capsid assembled from Cp183, emptied of RNA in a reducing environment. (e) Capsid assembled from Cp183, emptied of RNA in the absence of reducing agent. Cys48 forms a disulfide with Cys183 and thus becomes oxidized. Figure S7 contains further details.

https://doi.org/10.1021/jacs.1c12439
J. Am. Chem. Soc. 2022, 144, 8536–8550

sought to visualize residues which contact the phosphate backbone of the RNA. We implemented a ^{31}P spin filter[40] followed by a 2D ^{13}C-^{13}C correlation[41] (Figure 3a). In addition to the spin system of ribose, which is just adjacent to the phosphate, we saw signals which clearly derive from amino acids and can be unambiguously assigned to an arginine spin system (Figure 3b). Thus, this experiment demonstrates through-space contacts between arginine residues in the ARMs and RNA phosphates.

In protein-RNA complexes, so-called arginine forks (Figure 3c) for RNA recognition have been described, where the guanidinium group contacts either two phosphate moieties or a phosphate and a base.[42] For HBV capsids, our data point toward the first scenario as we observe no arginine-base contacts (Figure 2b). To quantify the distances between the phosphate backbone of RNA and the arginine side chains, we conducted rotational-echo double-resonance (REDOR) experiments. This MAS ssNMR technique enables the measurement of internuclear distances by means of the dipolar coupling. Specifically, we recorded, under cryogenic conditions, ^{15}N{^{31}P}J-decoupled REDOR[25,43] with selective excitation at ^{15}N frequencies corresponding to either the NH_2 amino groups or the Nε atom of the arginine guanidine group. The dephasing curve of the NH_2 groups displays a plateau at ∼0.7, indicating that only a third of them are in proximity to a phosphate group (Figure 3d). The distance distributions of ^{31}P-^{15}N spin pairs obtained via Boltzmann statistics (BS-REDOR[44]) reconstruction indicate a distinct ensemble centered at ∼4.8 Å that corresponds to a third of the dephased population. For comparison, REDOR investigations of a peptide interacting with phospholipids[45] found the Cζ atom of arginine at distances of 4 and 5.1 Å to two ^{31}P spins, which the authors interpreted as a bidentate complex. Similarly, the 4.8 ± 0.4 Å distance ensemble in our data could also encompass bidentate arginine-phosphate fork motifs (Figure 3c). The dephasing curve obtained with selective ^{15}N excitation at ∼80 ppm displays a similar pattern but yielded a broad ensemble centered at ∼5.5 Å that also represents a quarter to a third of the population (Figure 3d). This ensemble likely encompasses contributions of amino groups of adenine and guanine bases, which are subject to dephasing in the range of 6−8 Å from intramolecular contacts to the phosphate backbone.

Although two thirds (i.e., 16/24) of arginine residues in C protein belong to the CTD, no more than half engage the phosphate backbone of the RNA at any given time. This is not due to limited availability of RNA. On the contrary, UV absorbance measurements[46] indicate that an excess of 1.5 negatively charged nucleotides per positive charge in the CTD arm is incorporated into our capsids (Figure S4). This phenomenon is known as 'overcharging'[47] and common to ssRNA viruses.

We conclude that the interaction between arginine side chains and the RNA phosphate backbone is the major electrostatic driving force that confines the CTD to the interior space of RNA-filled nucleocapsids.

CTD Remains Dynamically Disordered Irrespective of the Capsid Content. So far, we have not addressed the conformation of the CTD, whether it is statically or dynamically disordered, and how it is modulated by the presence of RNA and by phosphorylation of serine residues[51] (sites highlighted in Figure 4a). Therefore, we produced different capsid types (Figure 4, left): first, self-assembled RNA-filled capsids built from Cp183, and from a phosphorylation-mimicking mutant (Cp183EEE);[51] second, capsids self-assembled from Cp149; third, capsids emptied of nucleic acids by in vitro reassembly of Cp183 in the presence or absence of a reducing agent.

We applied dynamic spectral editing techniques[52] to separate rigid segments, notably the capsid assembly domain, from flexible segments which retain local mobility even in the context of the capsid. The spectral separation of flexible from rigid is accomplished via complementary pulse sequence elements: transfers based on the isotropic J-coupling or direct excitation via a 90° pulse, combined with short relaxation delays, make experiments selective for dynamic segments of proteins and nucleic acids; meanwhile, transfers based on anisotropic nuclear spin interactions, notably the dipolar coupling, become ineffective in the presence of substantial molecular motions. In the spectra filtered via CP for rigid elements (Figure 4, gray contours; full spectra in Figure S6), we did not notice any impact of nucleic acid presence on the capsid assembly domain, in agreement with previous reports.[33] For none of the investigated capsid types did we observe the clear appearance of signals originating from the CTD in CP-based spectra. This finding argues against the rigidification of the CTD under the tested conditions.

Next, we analyzed the dynamic parts of the nucleocapsids in directly pulsed (dp) ^{13}C PDSD spectra with short relaxation delays (Figure 4b). In all capsid types except for those lacking the CTD, we observed similar signals from dynamic residues that can thus be assigned to the ARMs and interspersed serine residues (Figure 4c−g). The signals coincide with the respective random coil chemical shift range of arginine and serine derived from biological magnetic resonance data bank (BMRB) statistics.[48,53] Changes in CTD dynamics between different capsid types are either minor or below the sensitivity threshold, as was tentatively estimated from the ratios of signal intensities between cross and diagonal peaks (Figure S6). We implemented further experiments to probe for dynamically disordered capsid segments, notably INEPT and INEPT-TOBSY,[54] but did not detect any signals that could be assigned to the CTD; Table S1 contains an overview of NMR experiments used to probe CTD properties.

In summary, irrespective of capsid content, the CTD is mobile at ambient temperature and rapidly samples multiple conformations such that we observe the averaged, or random coil, chemical shift. The interactions of the CTD with enclosed RNA are transient. When molecular motions are frozen under cryogenic conditions, dipolar-based polarization transfer techniques reveal electrostatic interactions between guanidinium groups of arginine and the phosphate-sugar backbone of encapsidated RNA.

Intradimeric Disulfide Can Lock Free CTDs onto the Inner Capsid Surface. Given the mobility of the CTDs and their unspecific interactions with RNA, we cannot easily rationalize how the information about the nucleic acid content inside nucleocapsids could be transmitted via the CTD to the capsid exterior and how the location of the CTD could be regulated. Thus, we speculated that covalent effects could play a role. Indeed, the C protein features four cysteine residues, three in the assembly domain and one at the C-terminus (Figure 5a). Biochemical studies have not yet resolved the functional relevance of these cysteines and reported conflicting types of disulfide linkages.[55−59]

https://doi.org/10.1021/jacs.1c12439
J. Am. Chem. Soc. 2022, 144, 8536−8550

Cys48 located at the base of the C protein dimer and the terminal cysteine Cys183 are accessible to chemical functionalization and hence in principle to disulfide formation. We applied $^{13}CH_3$-S labeling of cysteine side chains using *S*-methyl-^{13}C-methanethiosulfonate (MMTS).[60] The procedure results in the formation of *S*-methylthiocysteine (MTC) with one ^{13}C spin label for each reduced cysteine accessible to MMTS. In contrast, an oxidized cysteine will acquire no spin label, and thus its NMR signal will be identical to the natural abundance background. Addition of the reducing agent tris(2-carboxyethyl)phosphine (TCEP) to the buffer prior to labeling converts all potentially accessible cysteines to a state ready for linkage to MMTS.

In Figure 5b, NMR spectra of $^{13}CH_3$-S labeled RNA-filled capsids contain two main peaks, which were assigned via mutagenesis to MTC-48 and MTC-183. Both signals are present regardless of the addition of TCEP. In contrast, no MTC signals are detected in capsids emptied of RNA when TCEP was absent, suggesting that both Cys48 and Cys183 are engaged in disulfide bonding. Two remaining cysteines (Cys61 and Cys107) were inaccessible to MMTS, presumably because of their location within the interior of the C protein dimer.

Having identified cysteine residues available for oxidation, we tested whether a disulfide would spontaneously form in the aforementioned capsid types. The oxidation state of cysteines can be assessed from their $^{13}C\alpha$ and $^{13}C\beta$ chemical shifts.[61] In capsids containing RNA, Cys48 remained reduced and did not become involved in a disulfide even in the absence of a reducing agent (Figures 5c and S7a), consistent with $^{13}C\alpha/^{13}C\beta$ shifts of 65/23 ppm, respectively. We hypothesized that the interactions with RNA might sequester the CTD in the interior capsid space, thus preventing disulfide formation between Cys48 and Cys183. Indeed, we found that in capsids devoid of RNA, Cys48 can readily switch between two states depending on the redox environment: that is, presence or absence of the reducing agent (Figures 5d,e and S7c,d). While the presence of RNA determined whether Cys48 was able to be oxidized, phosphorylation-mimicking mutations had no effect (Figure S7b,g). In capsids lacking the CTD (Cp149), Cys48 did not become oxidized, ruling out the possibility of a disulfide bond with its neighboring Cys48 across the intradimer interface (Figure S7e,f), pointing toward a bond formed instead with Cys183.

To prove that a disulfide bond is indeed formed between Cys48 and Cys183, we also monitored the CTD. As the latter mostly escapes dipolar transfers, we performed NMR experiments in solution, on in vitro reassembled RNA-free capsids and isolated C protein dimers (Figure 6). Building on studies of Cp149 dimers by solution-state NMR,[62] we isolated Cp183 dimers by disassembling Cp183 capsids and removing the RNA[16] (Figure S8). The amide correlation spectrum (non-TROSY ^{1}H-^{15}N HSQC) of protonated C protein dimers singles out small and flexible protein segments. This spectrum features the typical spectral pattern of an intrinsically disordered region with narrow dispersion and contains signals of approximately 30 amide groups, confirming that the CTD is singled out in this setup. Importantly, we found that the terminal cysteine can be subjected to multiple oxidation–reduction cycles within both the empty capsid (Figure 6a,c) and the isolated dimer (Figure 6b,c). As the latter is sufficient, the disulfide appears to be intradimeric. The conversion of reduced to oxidized dimers was completed within a few hours under exposure to oxygen from the air (Figure S9). In contrast,

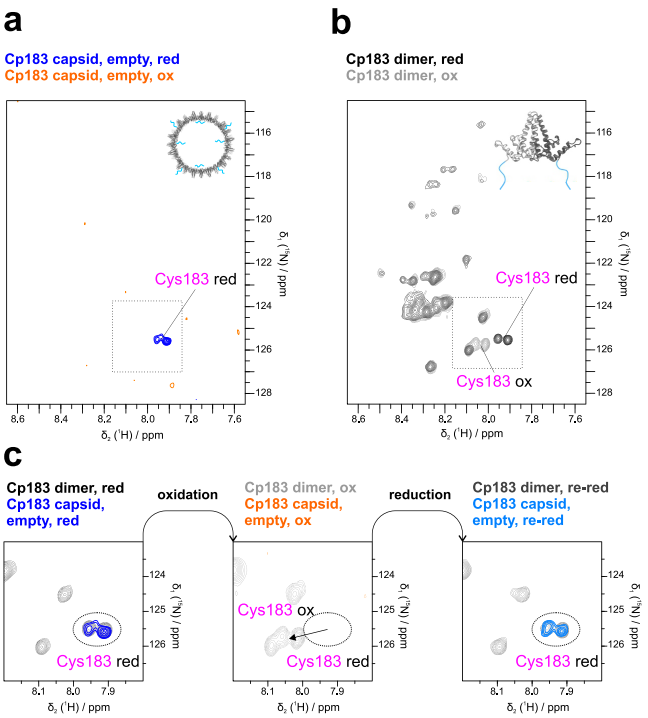


Figure 6. Reversible redox switching of the CTD. Solution spectra (^{1}H-^{15}N HSQC) of in vitro assembled capsids and disassembled C protein dimers at 20 °C, demonstrating reversible oxidation of Cys183. This residue was assigned by mutagenesis and yields the only detectable signal of the capsid. (a) Spectrum of capsids under oxidizing conditions contains no signal from Cys183, demonstrated by contours lowered down to the noise level. (b) In C protein dimers, the CTD displays the narrow chemical shift dispersion expected for an intrinsically disordered region. The signal from Cys183 shifts with oxidation state. (c) Expansions demonstrating oxidation–reduction cycles of Cys183. Under reducing conditions, the signal from Cys183 is visible for both dimers and capsids, implying a free state of the CTD. An oxidative environment induces a disulfide between Cys183 and Cys48, resulting in the disappearance of the former signal from the capsid spectrum and a shift in the dimer spectrum. The consecutive addition of reducing agent results in the reappearance of the Cys183 signal. Note a peak doubling of the terminal Cys183 in empty capsids and dimers alike, irrespective of the oxidation state. This hints at the existence of two distinct CTD populations that do not exchange on the timescale of the NMR experiment.[63]

in dimers bearing the Cys48S mutation, the oxidation of Cys183 occurred on a time scale of days. The reported disulfide linkage between homologous Cys183 residues within the dimer[57] is thus possible but subordinate.

Thus, the CTD can be locked covalently to the inner capsid surface at the base of C protein dimers. The presence of enclosed RNA counteracts this process efficiently, and hence, counter-intuitively, transient electrostatic interactions with nucleic acids can leave a covalent imprint on the C protein.

DNA-Filled Capsids Are Subject to the Disulfide Switch. In the course of maturation, HBV nucleocapsids harbor first ssRNA, then ssDNA, and finally partially dsDNA. The capsids investigated so far enclose predominantly ssRNA of bacterial origin. We sought to establish whether the presence of dsDNA would likewise impede the formation of the Cys48-Cys183 disulfide linkage. In contrast to ss nucleic acids, the C protein has a low affinity for dsDNA. Linear dsDNA fragments are poor templates for capsid assembly from Cp183 dimers in vitro.[64] We therefore attempted the

https://doi.org/10.1021/jacs.1c12439
J. Am. Chem. Soc. 2022, 144, 8536−8550

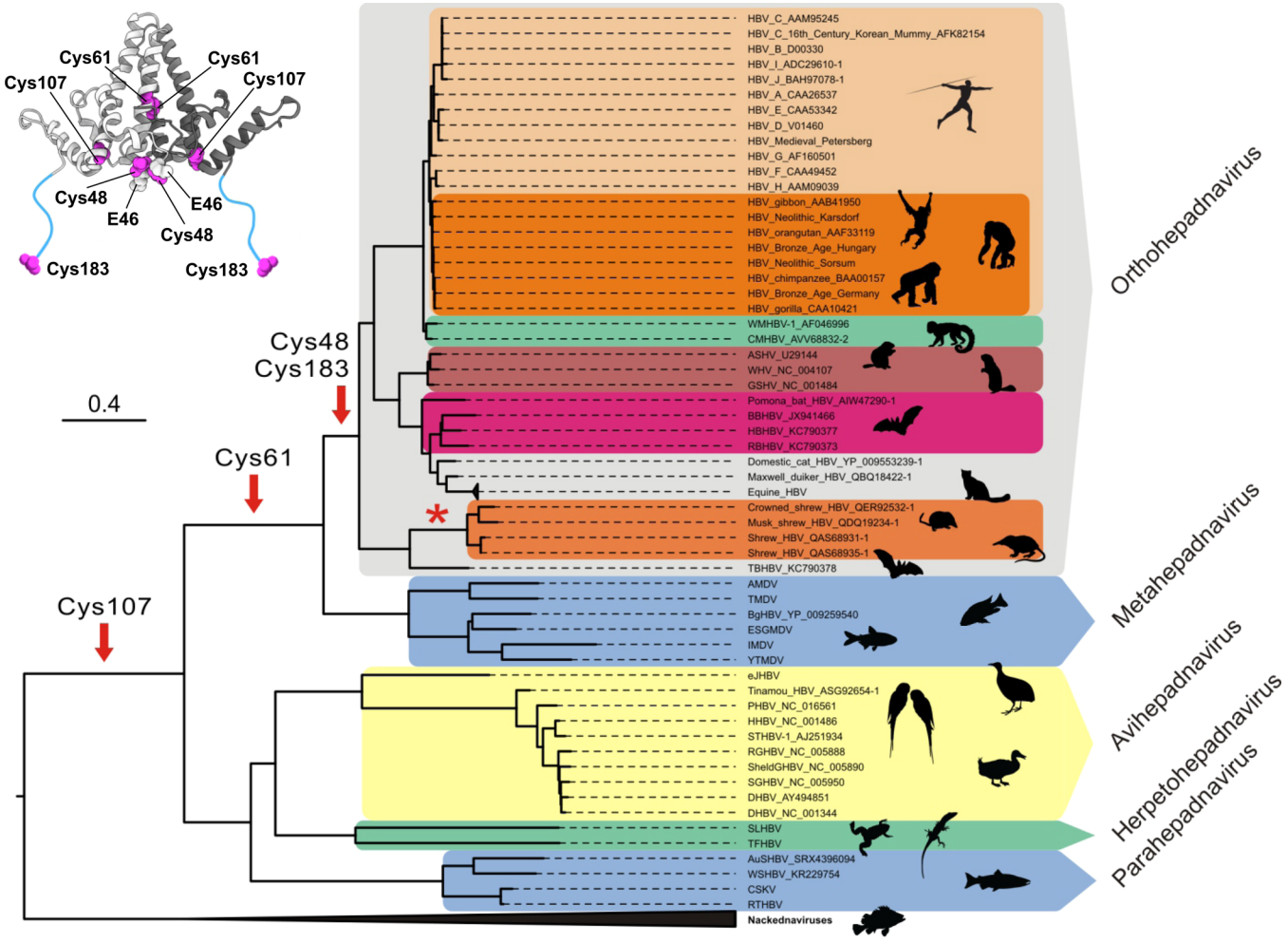


Figure 7. Phylogenetic analysis of hepadnaviral C protein. Cys48 and Cys183, along with Glu46 (Figure S11), appeared simultaneously and are unique to mammalian HBVs (genus *Orthohepadnavirus*) but were subsequently lost in the viral lineage infecting shrews (red asterisk). The scale bar indicates the number of amino acid substitutions per site. The location of four conserved cysteines, as well as Glu46, in the C protein dimer of human HBV is shown on the left. *Homo sapiens*, Cichlid, Mexican tetra, Domestic duck, Frog, Spiny lizard, Coho salmon, and Tiger rockfish reproduced with permission from Ref 65. Copyright 2017 Elsevier (https://creativecommons.org/licenses/by/4.0/).

encapsidation of a more compacted template and opted for a supercoiled circular plasmid dsDNA close to the size of the mature HBV genome (details in Figure S10a–c and Materials and Methods Section). While low yields precluded the full-scale NMR analysis conducted on empty and RNA-filled capsids, the dsDNA-filled capsids parallel empty ones in the inaccessibility of Cys48 and Cys183 sulfhydryl groups for MMTS labeling (Figure S10d). We deduce that the Cys48-Cys183 disulfide formed in our dsDNA-containing capsids.

Disulfide Switch in the C Protein Is Unique to Mammalian Hepadnaviruses. We sought to establish whether this disulfide switch is a relevant functional feature that evolved in HBV or merely an artifact from in vitro studies of capsids. The *Hepadnaviridae* constitute an ancient family of viruses that coevolved with their hosts over several hundreds of millions of years and afflict not only humans but members of all vertebrate classes.[65] We scrutinized the sequence evolution of the C protein, specifically when the four conserved cysteines first appeared (Figures 7, S11). Phylogenetic tree reconstruction revealed that Cys107 is the most ancient and conserved among all HBVs, demarcating them from *Nackednaviridae*, a sister family of nonenveloped fish viruses.[65] Cys61 appeared in the common ancestry of mammalian HBVs (genus *Orthohepadnavirus*) and the most closely related group of fish HBVs assigned to the genus *Metahepadnavirus*. Cys48 and Cys183 emerged last and are unique to the mammalian HBVs. Indeed, they strictly appear as a pair, simultaneously with Glu46 at the base of the four-helix bundle.

Interestingly, there is one exception: both cysteines and the glutamate were subsequently lost in the HBV lineage discovered in shrews.[66,67] The authors describing the discovery of shrew HBVs suggested that these viruses might possess decreased potential for establishing persistent infection. As we will discuss, this observation provides a valuable clue as to which steps in the viral life cycle may rely on the disulfide switch.

DISCUSSION

In this study, we set out to elucidate how positively charged peptide arms of viral capsids enclose nucleic acids—a building principle that is followed by diverse ssRNA viruses.[4] This principle is exemplified by immature nucleocapsids assembled from the C protein of HBV enclosing pregenomic ssRNA, which serves as a template for synthesis of the viral DNA via reverse transcription. We accomplished the characterization of the conformation and mobility of the ARMs in the CTD

https://doi.org/10.1021/jacs.1c12439
J. Am. Chem. Soc. 2022, 144, 8536−8550

peptide arms in the context of empty and RNA-filled capsids using MAS ssNMR. Additionally, we investigated isolated C protein dimers using solution-state NMR.

At ambient temperature, the CTD arms remained dynamically disordered irrespective of the presence of RNA. Under cryogenic conditions, we captured and quantified with REDOR-NMR the transient interactions between the ARMs and RNA, which are driven by multiple unspecific electrostatic contacts between guanidinium groups of the arginine residues and the phosphate moieties of the RNA backbone.[68,69] These became detectable only under cryogenic conditions where nearly all functionally relevant protein motions are frozen; namely, large-scale, local backbone, and side chain, as well as solvent motions,[34,39] such that dynamic disorder manifests as static disorder.[52]

Few experimental observations of how nucleic acids are packaged via flexible capsid ARMs have reached the atomic level so far. Molecular dynamics simulations suggest that basic side chains and phosphate groups are not strictly paired during electrostatically driven coassembly of capsomeres and RNA, which was ascribed to kinetic effects.[47,70] These observations are consistent with our finding that only half of the arginine residues in the ARMs are engaged with the phosphate backbone of RNA at any given time. For the assembled nucleocapsids of ssRNA viruses, overcharging was predicted[47,70] and found experimentally,[71] whereby not every negative charge of the RNA is compensated by a neutralizing amino acid. In contrast, a charge balance hypothesis has been postulated specifically for HBV capsids.[72] While the phenomenon of capsid assembly around nucleic acids is common to many ssRNA viruses, HBV is peculiar in that its mature nucleocapsid hosts dsDNA. Typical overcharging ratios lie between 1.5 and 2 nucleotides per positive charge of the capsid,[47] with our RNA-filled capsids in the lower range.

In addition to these unspecific interactions, we identified a binary conformational switch in the HBV capsid (Figure 8), namely, a disulfide that can tether the otherwise mobile CTD ('CTD free' state) onto the capsid's inner surface ('CTD locked' state). This locking mechanism could regulate the accessibility of the CTD in a redox-dependent manner, acting as a sensor that responds to changes in location, between the nucleus, cytosol, and extracellular space.

The switch is dependent on the state of nucleic acids enclosed in the capsid − it is active only if disulfide bond formation can outcompete nucleic acids that sequester the CTD to the interior capsid space. In fact, we demonstrate that encapsidated ssRNA counteracts disulfide bond formation effectively ('CTD bound' state in Figure 8). It has been hypothesized that during nucleocapsid maturation in vivo, dsDNA emerging via reverse transcription might sequester the CTDs less efficiently,[64] thus enabling the switching mechanism. Indeed, ssRNA and dsDNA behave differently because of different stiffness, or persistence length, of ss (~1 nm) versus ds (~50 nm) nucleic-acid polymers. Core protein dimers bind ssDNA templates with high affinity to produce intact capsids while assembly on dsDNA yields irregular assemblies.[64] It was deduced that dsDNA-containing nucleocapsids were spring-loaded and that the rigidification of the viral genome could release the CTDs. In support of this notion, our NMR analyses identified unambiguous contacts of the ARMs with RNA but not with DNA (Figure 2b). Likewise, in capsids devoid of nucleic acid, the CTDs are subject to the redox-dependent conformational switch ('CTD free' vs 'locked' state in Figure 8). In short, despite their transient nature, RNA interactions can leave a covalent imprint on the capsid protein. With such a disulfide locking mechanism, we report an additional possible layer of regulation that is enabled in the absence of ssRNA and could thus apply to mature dsDNA-filled nucleocapsids and genome-free capsids alike (Figure 8).

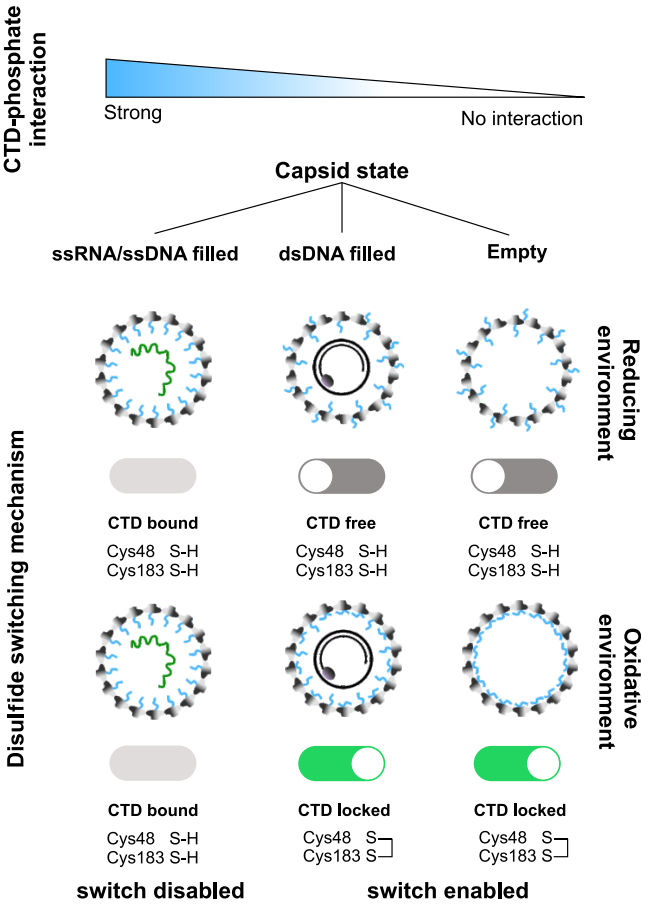


Figure 8. Model of the disulfide switching mechanism in the HBV life cycle. Interaction with ssRNA or ssDNA traps the CTD in the interior of the capsid and prevents disulfide bond formation of the terminal cysteine, making the redox-dependent switch inactive. In contrast, the mobile nature of the CTD in mature dsDNA-filled and empty capsids enables the disulfide switching mechanism, locking the CTD onto the interior of the capsid shell in oxidizing environments (e.g., extracellular space and endoplasmic reticulum).

In addition to recruiting and confining the viral genome during capsid assembly, the CTD also engages the host cell machinery to mediate cell or nuclear entry[13,73] − processes that require at least its transient exposure on the capsid surface.[74] It has been reported that the CTD location is mediated by the type of enclosed nucleic acid: trapped inside in the presence of ssRNA and mobile in mature dsDNA-filled capsids,[11] consistent with the 'CTD bound' and 'CTD free' states identified in our analyses (Figure 8). In most cryo-EM models of the capsid, the bulk of CTDs projects inward.[11,19] In capsids emptied of RNA under reducing conditions, electron density attributed to the CTD is partially exposed on the capsid exterior through the quasi-sixfold pores[11] and accessible to proteolysis[63] and importin binding.[75] Notably, removal of the terminal cysteine[76] and enclosed dsDNA[77] have been documented to increase CTD exposure. Genome-free and mature dsDNA-filled capsids share several characteristics that

https://doi.org/10.1021/jacs.1c12439
J. Am. Chem. Soc. 2022, 144, 8536−8550

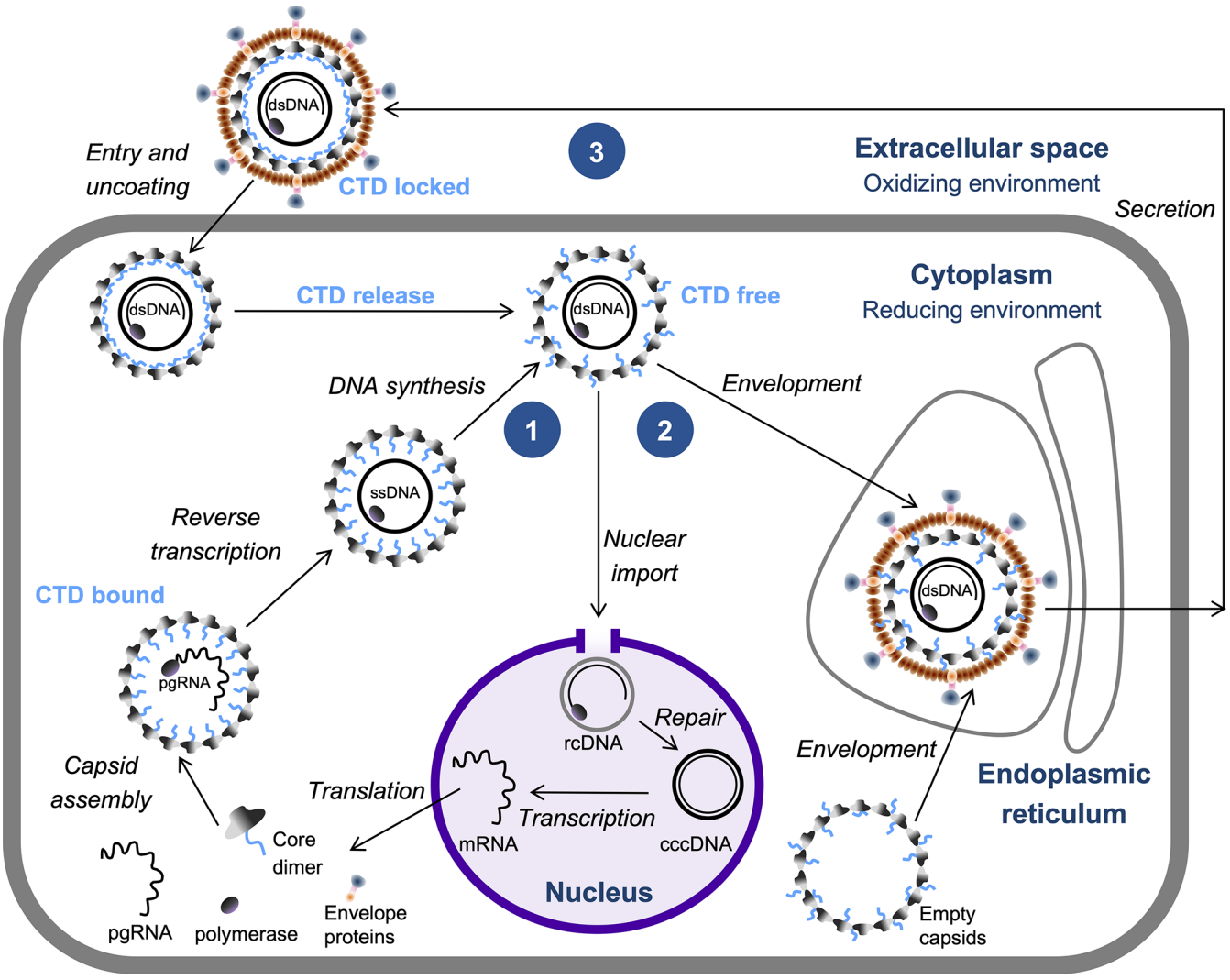


Figure 9. Schematic HBV life cycle, showing entry, uncoating, nuclear import, transcription, translation as well as nucleocapsid assembly, and virion secretion. Three key steps where the disulfide switching mechanism is hypothesized to play role are highlighted by numbers: (1) nucleocapsid maturation; (2) regulation of the ratio of nucleocapsid export *versus* re-import into the nucleus; (3) virion maturation. The hypothetical state of the CTD during the viral lifecycle is highlighted in light blue.

set them apart from their ssRNA-filled counterparts: both partially expose the nuclear localization signal in the CTD on their exterior to bind importin – under reducing conditions in vitro[63,75] and in the reducing environment of the cytosol.[77] Both species become enveloped and exported with high efficiency. Indeed, genome-free virus-like particles are abundant in the blood circulation of chronic HBV carriers.[78] These reports are fully consistent with our NMR-based analyses and are condensed into the model of Figure 8, with the implication that genome-free capsids can serve as a proxy for dsDNA-containing ones. We could extend this picture by the identification of an additional state – 'CTD locked' – in which a disulfide fixes the CTD to the inner capsid surface.

Our phylogenetic analysis confirms that the cysteine residues involved in the disulfide bond, Cys48 and Cys183, emerged as a pair in the common ancestry of mammalian HBVs. Their subsequent simultaneous loss in shrew HBV helps decipher the functional relevance of the switching mechanism encoded by these two residues. Interestingly, the newly discovered shrew HBVs lack genetic elements involved in viral persistence,[67] providing a valuable clue that the disulfide could play a role in establishment and maintenance of chronic infection. Possible steps in the viral life cycle relying on this switch are highlighted in Figure 9:

1. An unexplained phenomenon in HBV biology is the signaling of **nucleocapsid maturation**, which ensures that only replication-competent, dsDNA-containing capsids are packed into virions and released from the host cell. Numerous efforts[33,79] reviewed in ref 12. were aimed at identifying the molecular equivalent of this putative 'maturation signal'. So far, no single structural transition in the capsid has been linked to the capsid maturation state. The disulfide switch could potentially be involved in the specific recognition of mature nucleocapsids ready for envelopment. Reduced efficiency of envelopment has been noted for cysteine-less capsids.[56]
2. Mature nucleocapsids containing dsDNA are enveloped and exported as infectious virions, set to infect more cells, or become re-imported into the nucleus to replenish the pool of episomally persistent viral DNA. This necessitates a recognition sequence, the nuclear

https://doi.org/10.1021/jacs.1c12439
J. Am. Chem. Soc. 2022, 144, 8536–8550

localization signal within the CTD,[13] to be transiently exposed through the capsid pores. The disulfide switch could help regulate the ratio of **nucleocapsid export *versus* re-import into the nucleus**. In support of this notion, surface exposure of the CTD is increased in the absence of Cys183.[76]

3. HBV virions, that is, enveloped nucleocapsids, leave the cell in a noninfectious state in which the major viral entry determinant is hidden in the particles' interior where it interacts with the nucleocapsid.[80] After reversal of this interaction, the entry determinant spontaneously translocates across the lipid bilayer of the viral envelope to become exposed on the particle exterior, where it facilitates receptor-dependent attachment to the next host cell.[80] The slow transition into the 'CTD locked' state proceeding under the more oxidizing conditions of the extracellular space, namely, the bloodstream, might initiate the process of **virion maturation** by triggering the abolishment of the interaction between the entry determinant and the C protein. The disulfide switch operates on a time scale of a few hours, which coincides with the 4−5 h required for the maturation of virions in vitro[80] and with the half-life of virions in serum.[81,82]

CONCLUSIONS

The HBV is impressively minimalistic: around 3200 base pairs encode only four distinct proteins in multiple isoforms. These four proteins sustain the viral life cycle, each assuming multiple conformations and forming transient interactions with each other and the host cell machinery. Two principles prominently underlie the plasticity of this virus: disordered protein segments and disulfide switches. Disorder is amply exploited by viruses[83] and disulfides confer switch-like properties with minimal investment—two strategically placed cysteine residues are sufficient. Here, we discover a disulfide switching mechanism involving a disordered protein segment of the capsid protein and show that it reports on the status of the viral genome.

This instance of ultrastructural dynamics in HBV nucleocapsids could only be fully elucidated by a combination of NMR approaches, from the frozen protein sediment to solution. As disordered, dynamic, and transiently interacting biomolecules are challenging to grasp by many techniques, NMR investigations that span wide molecular size scales and yet capture key atomic-level details of maturation processes bear great potential in the structural biology of viral nucleocapsids.

MATERIALS AND METHODS

Sample preparation, acquisition, and analysis of NMR data and procedures for phylogenetic analysis are described in the Materials and Methods Section in the Supporting Information.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.1c12439.

Supporting figures with full NMR spectra; control experiments; UV measurements; size-exclusion chromatograms; DNA encapsidation; C protein sequence alignment; a table with overview of NMR experiments performed; and tables with experimental parameters for MAS ssNMR experiments (PDF)

AUTHOR INFORMATION

Corresponding Authors

Marc-Antoine Sani − *School of Chemistry, Bio21 Institute, University of Melbourne, Melbourne, Victoria 3010, Australia*; orcid.org/0000-0003-3284-2176; Email: msani@unimelb.edu.au

Stefan Seitz − *Department of Infectious Diseases, Molecular Virology, University of Heidelberg, Heidelberg 69120, Germany; Division of Virus-Associated Carcinogenesis (F170), German Cancer Research Center (DKFZ), Heidelberg 69120, Germany*; Email: stefan.seitz@med.uni-heidelberg.de

Anne K. Schütz − *Bavarian NMR Center, Department of Chemistry, Technical University of Munich, Garching 85748, Germany; Institute of Structural Biology, Helmholtz Zentrum München, Neuherberg 85764, Germany*; orcid.org/0000-0001-6398-5757; Email: anne.schuetz@tum.de

Authors

Zahra Harati Taji − *Bavarian NMR Center, Department of Chemistry, Technical University of Munich, Garching 85748, Germany; Institute of Structural Biology, Helmholtz Zentrum München, Neuherberg 85764, Germany*

Pavlo Bielytskyi − *Bavarian NMR Center, Department of Chemistry, Technical University of Munich, Garching 85748, Germany; Institute of Structural Biology, Helmholtz Zentrum München, Neuherberg 85764, Germany*

Mikhail Shein − *Bavarian NMR Center, Department of Chemistry, Technical University of Munich, Garching 85748, Germany; Institute of Structural Biology, Helmholtz Zentrum München, Neuherberg 85764, Germany*

Complete contact information is available at: https://pubs.acs.org/10.1021/jacs.1c12439

Author Contributions

#Z.H.T. and P.B. contributed equally to this work.

Funding

This work was funded by the German Research Foundation (DFG) through the Emmy Noether program (project number 394455587 to A.K.S.), TRR179 (project number 272983813 and project 23 to A.K.S and S.S.), and SFB1035 (project number 201302640 and project B15 to A.K.S.). The work of S.S. was supported by a grant from the Helmholtz Association's Initiative and Networking Fund (Virological and immunological determinants of COVID-19 pathogenesis − lessons to get prepared for future pandemics, KA1-Co-02 'COVIPA'). M.-A.S. acknowledges the support from the Australian Research Council (ARC) via LIEF funding (LE160100120). Access to NMR spectrometers was provided by Bavarian NMR Centre of Technical University of Munich, Helmholtz Centre Munich and the Bio21 Institute at the University of Melbourne.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Dr. Riddhiman Sarkar and Matthias Brandl for support during DNP experiments and Dr. Gerd Gemmecker and Dr. Sam Asami for support with solution-state NMR. We

thank Dr. Christoph Kaiser and Dr. Carsten Peters for help with electron microscopy measurements.

REFERENCES

(1) Bruinsma, R. F.; Wuite, G. J. L.; Roos, W. H. Physics of viral dynamics. *Nat. Rev. Phys.* **2021**, *3*, 76−91.
(2) Requião, R. D.; Carneiro, R. L.; Moreira, M. H.; Ribeiro-Alves, M.; Rossetto, S.; Palhano, F. L.; Domitrovic, T. Viruses with different genome types adopt a similar strategy to pack nucleic acids based on positively charged protein domains. *Sci. Rep.* **2020**, *10*, 5470.
(3) Comas-Garcia, M. Packaging of Genomic RNA in Positive-Sense Single-Stranded RNA Viruses: A Complex Story. *Viruses* **2019**, *11*, 253.
(4) Belyi, V. A.; Muthukumar, M. Electrostatic origin of the genome packing in viruses. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 17174−17178.
(5) Jensen, M. R.; Communie, G.; Ribeiro, E. A., Jr.; Martinez, N.; Desfosses, A.; Salmon, L.; Mollica, L.; Gabel, F.; Jamin, M.; Longhi, S.; Ruigrok, R. W.; Blackledge, M. Intrinsic disorder in measles virus nucleocapsids. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 9839−9844.
(6) Savastano, A.; Ibáñez de Opakua, A.; Rankovic, M.; Zweckstetter, M. Nucleocapsid protein of SARS-CoV-2 phase separates into RNA-rich polymerase-containing condensates. *Nat. Commun.* **2020**, *11*, 6041.
(7) Busi, B.; Yarava, J. R.; Bertarello, A.; Freymond, F.; Adamski, W.; Maurin, D.; Hiller, M.; Oschkinat, H.; Blackledge, M.; Emsley, L. Similarities and Differences among Protein Dynamics Studied by Variable Temperature Nuclear Magnetic Resonance Relaxation. *J. Phys. Chem B* **2021**, *125*, 2212−2221.
(8) Crowther, R. A.; Kiselev, N. A.; Böttcher, B.; Berriman, J. A.; Borisova, G. P.; Ose, V.; Pumpens, P. Three-dimensional structure of hepatitis B virus core particles determined by electron cryomicroscopy. *Cell* **1994**, *77*, 943−950.
(9) Gerelsaikhan, T.; Tavis, J. E.; Bruss, V. Hepatitis B virus nucleocapsid envelopment does not occur without genomic DNA synthesis. *J. Virol.* **1996**, *70*, 4269−4274.
(10) Ning, X.; Nguyen, D.; Mentzer, L.; Adams, C.; Lee, H.; Ashley, R.; Hafenstein, S.; Hu, J. Secretion of genome-free hepatitis B virus−single strand blocking model for virion morphogenesis of para-retrovirus. *PLoS Pathog.* **2011**, *7*, No. e1002255.
(11) Wang, J. C.-Y.; Dhason, M. S.; Zlotnick, A. Structural organization of pregenomic RNA and the carboxy-terminal domain of the capsid protein of hepatitis B virus. *PLoS Pathog.* **2012**, *8*, No. e1002919.
(12) Seitz, S.; Habjanič, J.; Schütz, A. K.; Bartenschlager, R. The Hepatitis B Virus Envelope Proteins: Molecular Gymnastics Throughout the Viral Life Cycle. *Annu. Rev. Virol.* **2020**, *7*, 263−288.
(13) Kann, M.; Sodeik, B.; Vlachou, A.; Gerlich, W. H.; Helenius, A. Phosphorylation-dependent binding of hepatitis B virus core particles to the nuclear pore complex. *J. Cell Biol.* **1999**, *145*, 45−55.
(14) Patel, N.; White, S. J.; Thompson, R. F.; Bingham, R.; Weiß, E. U.; Maskell, D. P.; Zlotnick, A.; Dykeman, E.; Tuma, R.; Twarock, R.; Ranson, N. A.; Stockley, P. G. HBV RNA pre-genome encodes specific motifs that mediate interactions with the viral core protein that promote nucleocapsid assembly. *Nat. Microbiol.* **2017**, *2*, 17098.
(15) Bartenschlager, R.; Schaller, H. Hepadnaviral assembly is initiated by polymerase binding to the encapsidation signal in the viral RNA genome. *Embo J.* **1992**, *11*, 3413−3420.
(16) Porterfield, J. Z.; Dhason, M. S.; Loeb, D. D.; Nassal, M.; Stray, S. J.; Zlotnick, A. Full-length hepatitis B virus core protein packages viral and heterologous RNA with similarly high levels of cooperativity. *J. Virol.* **2010**, *84*, 7174−7184.
(17) Wynne, S. A.; Crowther, R. A.; Leslie, A. G. W. The crystal structure of the human hepatitis B virus capsid. *Mol. Cell* **1999**, *3*, 771−780.
(18) Conway, J. F.; Cheng, N.; Zlotnick, A.; Wingfield, P. T.; Stahl, S. J.; Steven, A. C. Visualization of a 4-helix bundle in the hepatitis B virus capsid by cryo-electron microscopy. *Nature* **1997**, *386*, 91−94.
(19) Yu, X.; Jin, L.; Jih, J.; Shih, C.; Hong Zhou, Z. 3.5A cryoEM structure of hepatitis B virus core assembled from full-length core protein. *PLoS One* **2013**, *8*, No. e69729.
(20) Gupta, R.; Zhang, H.; Lu, M.; Hou, G.; Caporini, M.; Rosay, M.; Maas, W.; Struppe, J.; Ahn, J.; Byeon, I. L.; Oschkinat, H.; Jaudzems, K.; Barbet-Massin, E.; Emsley, L.; Pintacuda, G.; Lesage, A.; Gronenborn, A. M.; Polenova, T. Dynamic Nuclear Polarization Magic-Angle Spinning Nuclear Magnetic Resonance Combined with Molecular Dynamics Simulations Permits Detection of Order and Disorder in Viral Assemblies. *J. Phys. Chem. B* **2019**, *123*, 5048−5058.
(21) Lu, M.; Russell, R. W.; Bryer, A. J.; Quinn, C. M.; Hou, G.; Zhang, H.; Schwieters, C. D.; Perilla, J. R.; Gronenborn, A. M.; Polenova, T. Atomic-resolution structure of HIV-1 capsid tubes by magic-angle spinning NMR. *Nat. Struct. Mol. Biol.* **2020**, *27*, 863−869.
(22) Kwon, B.; Lee, M.; Waring, A. J.; Hong, M. Oligomeric Structure and Three-Dimensional Fold of the HIV gp41 Membrane-Proximal External Region and Transmembrane Domain in Phospholipid Bilayers. *J. Am. Chem. Soc.* **2018**, *140*, 8246−8259.
(23) Movellan, K. T.; Wegstroth, M.; Overkamp, K.; Leonov, A.; Becker, S.; Andreas, L. B. Imidazole-Imidazole Hydrogen Bonding in the pH-Sensing Histidine Side Chains of Influenza A M2. *J. Am. Chem. Soc.* **2020**, *142*, 2704−2708.
(24) Mandala, V. S.; McKay, M. J.; Shcherbakov, A. A.; Dregni, A. J.; Kolocouris, A.; Hong, M. Structure and drug binding of the SARS-CoV-2 envelope protein transmembrane domain in lipid bilayers. *Nat. Struct. Mol. Biol.* **2020**, *27*, 1202−1208.
(25) Yu, T. Y.; Schaefer, J. REDOR NMR characterization of DNA packaging in bacteriophage T4. *J. Mol. Biol.* **2008**, *382*, 1031−1042.
(26) Sergeyev, I. V.; Itin, B.; Rogawski, R.; Day, L. A.; McDermott, A. E. Efficient assignment and NMR analysis of an intact virus using sequential side-chain correlations and DNP sensitization. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 5171−5176.
(27) Bloembergen, N. On the interaction of nuclear spins in a crystalline lattice. *Physica* **1949**, *15*, 386−426.
(28) Takegoshi, K.; Nakamura, S.; Terao, T. ^{13}C-^{1}H dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chem. Phys. Lett.* **2001**, *344*, 631−637.
(29) Baldus, M.; Petkova, A. T.; Herzfeld, J.; Griffin, R. G. Cross polarization in the tilted frame: assignment and spectral simplification in heteronuclear spin systems. *Mol. Phys.* **1998**, *95*, 1197−1207.
(30) Lange, A.; Luca, S.; Baldus, M. Structural constraints from proton-mediated rare-spin correlation spectroscopy in rotating solids. *J. Am. Chem. Soc.* **2002**, *124*, 9704−9705.
(31) Lecoq, L.; Wang, S.; Wiegand, T.; Bressanelli, S.; Nassal, M.; Meier, B. H.; Böckmann, A. Solid-state [^{13}C-^{15}N] NMR resonance assignment of hepatitis B virus core protein. *Biomol. NMR Assignments* **2018**, *12*, 205−214.
(32) Lecoq, L.; Wang, S.; Wiegand, T.; Bressanelli, S.; Nassal, M.; Meier, B. H.; Böckmann, A. Localizing Conformational Hinges by NMR: Where Do Hepatitis B Virus Core Proteins Adapt for Capsid Assembly? *ChemPhysChem* **2018**, *19*, 1336−1340.
(33) Lecoq, L.; Wang, S.; Dujardin, M.; Zimmermann, P.; Schuster, L.; Fogeron, M. L.; Briday, M.; Schledorn, M.; Wiegand, T.; Cole, L.; Montserret, R.; Bressanelli, S.; Meier, B. H.; Nassal, M.; Böckmann, A. A pocket-factor-triggered conformational switch in the hepatitis B virus capsid. *Proc. Natl. Acad. Sci. U. S. A.* **2021**, *118*, No. e2022464118.
(34) Lewandowski, J. R.; Halse, M. E.; Blackledge, M.; Emsley, L. Protein dynamics. Direct observation of hierarchical protein dynamics. *Science* **2015**, *348*, 578−581.
(35) Sauvée, C.; Rosay, M.; Casano, G.; Aussenac, F.; Weber, R. T.; Ouari, O.; Tordo, P. Highly efficient, water-soluble polarizing agents for dynamic nuclear polarization at high frequency. *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 10858−10861.
(36) Ravera, E.; Corzilius, B.; Michaelis, V. K.; Rosa, C.; Griffin, R. G.; Luchinat, C.; Bertini, I. Dynamic nuclear polarization of sedimented solutes. *J. Am. Chem. Soc.* **2013**, *135*, 1641−1644.

https://doi.org/10.1021/jacs.1c12439
J. Am. Chem. Soc. 2022, 144, 8536−8550

(37) Ravera, E.; Corzilius, B.; Michaelis, V. K.; Luchinat, C.; Griffin, R. G.; Bertini, I. DNP-enhanced MAS NMR of bovine serum albumin sediments and solutions. *J. Phys. Chem. B* **2014**, *118*, 2957−2965.
(38) Bauer, T.; Dotta, C.; Balacescu, L.; Gath, J.; Hunkeler, A.; Böckmann, A.; Meier, B. H. Line-Broadening in Low-Temperature Solid-State NMR Spectra of Fibrils. *J. Biomol. NMR* **2017**, *67*, 51−61.
(39) Linden, A. H.; Franks, W. T.; Akbey, U.; Lange, S.; van Rossum, B. J.; Oschkinat, H. Cryogenic temperature effects and resolution upon slow cooling of protein preparations in solid state NMR. *J. Biomol. NMR* **2011**, *51*, 283−292.
(40) Morag, O.; Abramov, G.; Goldbourt, A. Complete chemical shift assignment of the ssDNA in the filamentous bacteriophage fd reports on its conformation and on its interface with the capsid shell. *J. Am. Chem. Soc.* **2014**, *136*, 2292−2301.
(41) Wiegand, T.; Liao, W. C.; Ong, T. C.; Däpp, A.; Cadalbert, R.; Copéret, C.; Böckmann, A.; Meier, B. H. Protein-nucleotide contacts in motor proteins detected by DNP-enhanced solid-state NMR. *J. Biomol. NMR* **2017**, *69*, 157−164.
(42) Chavali, S. S.; Cavender, C. E.; Mathews, D. H.; Wedekind, J. E. Arginine Forks Are a Widespread Motif to Recognize Phosphate Backbones and Guanine Nucleobases in the RNA Major Groove. *J. Am. Chem. Soc.* **2020**, *142*, 19835−19839.
(43) Jaroniec, C. P.; Tounge, B. A.; Rienstra, C. M.; Herzfeld, J.; Griffin, R. G. Measurement of $^{13}C-^{15}N$ Distances in Uniformly ^{13}C Labeled Biomolecules: J-Decoupled REDOR. *J. Am. Chem. Soc.* **1999**, *121*, 10237−10238.
(44) Gehman, J. D.; Separovic, F.; Lu, K.; Mehta, A. K. Boltzmann statistics rotational-echo double-resonance analysis. *J. Phys. Chem. B* **2007**, *111*, 7802−7811.
(45) Tang, M.; Waring, A. J.; Hong, M. Phosphate-Mediated Arginine Insertion into Lipid Membranes and Pore Formation by a Cationic Membrane Peptide from Solid-State NMR. *J. Am. Chem. Soc.* **2007**, *129*, 11438−11446.
(46) Porterfield, J. Z.; Zlotnick, A. A simple and general method for determining the protein and nucleic acid content of viruses by UV absorbance. *Virology* **2010**, *407*, 281−288.
(47) Perlmutter, J. D.; Hagan, M. F. Mechanisms of virus assembly. *Annu. Rev. Phys. Chem.* **2015**, *66*, 217−239.
(48) Ulrich, E. L.; Akutsu, H.; Doreleijers, J. F.; Harano, Y.; Ioannidis, Y. E.; Lin, J.; Livny, M.; Mading, S.; Maziuk, D.; Miller, Z.; Nakatani, E.; Schulte, C. F.; Tolmie, D. E.; Kent Wenger, R.; Yao, H.; Markley, J. L. BioMagResBank. *Nucleic Acids Res.* **2008**, *36*, D402−D408.
(49) Borcik, C. G.; Versteeg, D. B.; Wylie, B. J. An Inward-Rectifier Potassium Channel Coordinates the Properties of Biologically Derived Membranes. *Biophys. J.* **2019**, *116*, 1701−1718.
(50) Calnan, B. J.; Biancalana, S.; Hudson, D.; Frankel, A. D. Analysis of arginine-rich peptides from the HIV Tat protein reveals unusual features of RNA-protein recognition. *Genes Dev.* **1991**, *5*, 201−210.
(51) de Rocquigny, H.; Rat, V.; Pastor, F.; Darlix, J. L.; Hourioux, C.; Roingeard, P. Phosphorylation of the Arginine-Rich C-Terminal Domains of the Hepatitis B Virus (HBV) Core Protein as a Fine Regulator of the Interaction between HBc and Nucleic Acid. *Viruses* **2020**, *12*, 738.
(52) Siemer, A. B. Advances in studying protein disorder with solid-state NMR. *Solid State Nucl. Magn. Reson.* **2020**, *106*, No. 101643.
(53) Fritzsching, K. J.; Yang, Y.; Schmidt-Rohr, K.; Hong, M. Practical use of chemical shift databases for protein solid-state NMR: 2D chemical shift maps and amino-acid assignment with secondary-structure information. *J. Biomol. NMR* **2013**, *56*, 155−167.
(54) Hardy, E. H.; Verel, R.; Meier, B. H. Fast MAS total through-bond correlation spectroscopy. *J. Magn. Reson.* **2001**, *148*, 459−464.
(55) Gallina, A.; Bonelli, F.; Zentilin, L.; Rindi, G.; Muttini, M.; Milanesi, G. A recombinant hepatitis B core antigen polypeptide with the protamine-like domain deleted self-assembles into capsid particles but fails to bind nucleic acids. *J. Virol.* **1989**, *63*, 4645−4652.
(56) Nassal, M. Conserved cysteines of the hepatitis B virus core protein are not required for assembly of replication-competent core particles nor for their envelopment. *Virology* **1992**, *190*, 499−505.
(57) Nassal, M.; Rieger, A.; Steinau, O. Topological analysis of the hepatitis B virus core particle by cysteine-cysteine cross-linking. *J. Mol. Biol.* **1992**, *225*, 1013−1025.
(58) Zhou, S.; Standring, D. N. Cys residues of the hepatitis B virus capsid protein are not essential for the assembly of viral core particles but can influence their stability. *J. Virol.* **1992**, *66*, 5393−5398.
(59) Böttcher, B.; Nassal, M. Structure of Mutant Hepatitis B Core Protein Capsids with Premature Secretion Phenotype. *J. Mol. Biol.* **2018**, *430*, 4941−4954.
(60) Religa, T. L.; Ruschak, A. M.; Rosenzweig, R.; Kay, L. E. Site-directed methyl group labeling as an NMR probe of structure and dynamics in supramolecular protein systems: applications to the proteasome and to the ClpP protease. *J. Am. Chem. Soc.* **2011**, *133*, 9063−9068.
(61) Sharma, D.; Rajarathnam, K. 13C NMR chemical shifts can predict disulfide bond formation. *J. Biomol. NMR* **2000**, *18*, 165−171.
(62) Freund, S. M.; Johnson, C. M.; Jaulent, A. M.; Ferguson, N. Moving towards high-resolution descriptions of the molecular interactions and structural rearrangements of the human hepatitis B core protein. *J. Mol. Biol.* **2008**, *384*, 1301−1313.
(63) Selzer, L.; Kant, R.; Wang, J. C.-Y.; Bothner, B.; Zlotnick, A. Hepatitis B Virus Core Protein Phosphorylation Sites Affect Capsid Stability and Transient Exposure of the C-terminal Domain. *J. Biol. Chem.* **2015**, *290*, 28584−28593.
(64) Dhason, M. S.; Wang, J. C.-Y.; Hagan, M. F.; Zlotnick, A. Differential assembly of Hepatitis B Virus core protein on single- and double-stranded nucleic acid suggest the dsDNA-filled core is spring-loaded. *Virology* **2012**, *430*, 20−29.
(65) Lauber, C.; Seitz, S.; Mattei, S.; Suh, A.; Beck, J.; Herstein, J.; Börold, J.; Salzburger, W.; Kaderali, L.; Briggs, J. A. G.; Bartenschlager, R. Deciphering the Origin and Evolution of Hepatitis B Viruses by Means of a Family of Non-enveloped Fish Viruses. *Cell Host Microbe* **2017**, *22*, 387−399.e6.
(66) Nie, F. Y.; Tian, J. H.; Lin, X. D.; Yu, B.; Xing, J. G.; Cao, J. H.; Holmes, E. C.; Ma, R. Z.; Zhang, Y. Z. Discovery of a highly divergent hepadnavirus in shrews from China. *Virology* **2019**, *531*, 162−170.
(67) Rasche, A.; Lehmann, F.; König, A.; Goldmann, N.; Corman, V. M.; Moreira-Soto, A.; Geipel, A.; van Riel, D.; Vakulenko, Y. A.; Sander, A. L.; Niekamp, H.; Kepper, R.; Schlegel, M.; Akoua-Koffi, C.; Souza, B.; Sahr, F.; Olayemi, A.; Schulze, V.; Petraityte-Burneikiene, R.; Kazaks, A.; Lowjaga, K.; Geyer, J.; Kuiken, T.; Drosten, C.; Lukashev, A. N.; Fichet-Calvet, E.; Ulrich, R. G.; Glebe, D.; Drexler, J. F. Highly diversified shrew hepatitis B viruses corroborate ancient origins and divergent infection patterns of mammalian hepadnaviruses. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 17007−17012.
(68) Woods, A. S.; Ferré, S. Amazing stability of the arginine-phosphate electrostatic interaction. *J. Proteome Res.* **2005**, *4*, 1397−1402.
(69) Ellis, J. J.; Broom, M.; Jones, S. Protein-RNA interactions: structural analysis and functional classes. *Proteins* **2007**, *66*, 903−911.
(70) Perlmutter, J. D.; Qiao, C.; Hagan, M. F. Viral genome structures are optimal for capsid assembly. *Elife* **2013**, *2*, No. e00632.
(71) Garmann, R. F.; Comas-Garcia, M.; Koay, M. S.; Cornelissen, J. J.; Knobler, C. M.; Gelbart, W. M. Role of electrostatics in the assembly pathway of a single-stranded RNA virus. *J. Virol.* **2014**, *88*, 10472−10479.
(72) Su, P. Y.; Yang, C. J.; Chu, T. H.; Chang, C. H.; Chiang, C.; Tang, F. M.; Lee, C. Y.; Shih, C. HBV maintains electrostatic homeostasis by modulating negative charges from phosphoserine and encapsidated nucleic acids. *Sci. Rep.* **2016**, *6*, 38959.
(73) Gallucci, L.; Kann, M. Nuclear Import of Hepatitis B Virus Capsids and Genome. *Viruses* **2017**, *9*, 21.
(74) Hilmer, J. K.; Zlotnick, A.; Bothner, B. Conformational equilibria and rates of localized motion within hepatitis B virus capsids. *J. Mol. Biol.* **2008**, *375*, 581−594.

https://doi.org/10.1021/jacs.1c12439
J. Am. Chem. Soc. 2022, 144, 8536−8550

(75) Chen, C.; Wang, J. C.-Y.; Pierson, E. E.; Keifer, D. Z.; Delaleau, M.; Gallucci, L.; Cazenave, C.; Kann, M.; Jarrold, M. F.; Zlotnick, A. Importin β Can Bind Hepatitis B Virus Core Protein and Empty Core-Like Particles and Induce Structural Changes. *PLoS Pathog.* **2016**, *12*, No. e1005802.
(76) Seifer, M.; Standring, D. N. A protease-sensitive hinge linking the two domains of the hepatitis B virus core protein is exposed on the viral capsid surface. *J. Virol.* **1994**, *68*, 5548−5555.
(77) Rabe, B.; Vlachou, A.; Panté, N.; Helenius, A.; Kann, M. Nuclear import of hepatitis B virus capsids and release of the viral genome. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 9849−9854.
(78) Alberti, A.; Diana, S.; Scullard, G. H.; Eddleston, A. L. W. F.; Williams, R. Full and empty Dane particles in chronic hepatitis B virus infection: Relation to hepatitis B e antigen and presence of liver damage. *Gastroenterology* **1978**, *75*, 869−874.
(79) Roseman, A. M.; Berriman, J. A.; Wynne, S. A.; Butler, P. J.; Crowther, R. A. A structural model for maturation of the hepatitis B virus core. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 15821−15826.
(80) Seitz, S.; Iancu, C.; Volz, T.; Mier, W.; Dandri, M.; Urban, S.; Bartenschlager, R. A Slow Maturation Process Renders Hepatitis B Virus Infectious. *Cell Host Microbe* **2016**, *20*, 25−35.
(81) Murray, J. M.; Purcell, R. H.; Wieland, S. F. The half-life of hepatitis B virions. *Hepatology* **2006**, *44*, 1117−1121.
(82) Murray, J. M.; Wieland, S. F.; Purcell, R. H.; Chisari, F. V. Dynamics of hepatitis B virus clearance in chimpanzees. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 17780−17785.
(83) Xue, B.; Blocquel, D.; Habchi, J.; Uversky, A. V.; Kurgan, L.; Uversky, V. N.; Longhi, S. Structural disorder in viral proteins. *Chem. Rev.* **2014**, *114*, 6880−6911.

https://doi.org/10.1021/jacs.1c12439
J. Am. Chem. Soc. 2022, 144, 8536−8550