STRUCTURAL BIOLOGY

Inherent symmetry and flexibility in hepatitis B virus subviral particles

Quan Wang[1,2,3]*†, Tao Wang[4]†, Lin Cao[5]†, An Mu[6,7]†, Sheng Fu[6]†, Peipei Wang[1]†, Yan Gao[1], Wenxin Ji[3], Zhenyu Liu[4], Zhanqiang Du[5], Luke W. Guddat[8], Wenchi Zhang[6], Shuang Li[6], Xuemei Li[6], Zhiyong Lou[4], Xiangxi Wang[6], Zhongyu Hu[9]*, Zihe Rao[1,4,5,6]*

Chronic hepatitis B virus (HBV) infection poses a major global health challenge with massive morbidity and mortality. Despite a preventive vaccine, current treatments provide limited virus clearance, necessitating lifelong commitment. The HBV surface antigen (HBsAg) is crucial for diagnosis and prognosis, yet its high-resolution structure and assembly on the virus envelope remain elusive. Utilizing extensive datasets and advanced cryo–electron microscopy analysis, we present structural insights into HBsAg at a near-atomic resolution of 3.7 angstroms. HBsAg homodimers assemble into subviral particles with D_2- and D_4-like quasisymmetry, elucidating the dense-packing rules and structural adaptability of HBsAg. These findings provide insights into how HBsAg assembles into higher-order filaments and interacts with the capsid to form virions.

The World Health Organization (WHO) estimates that approximately 300 million individuals worldwide live with chronic hepatitis B, resulting in over 800,000 escalating deaths each year (*1*). Although there is a preventive vaccine, hepatitis B still stands as the primary cause of chronic hepatitis, cirrhosis, and hepatocellular carcinoma globally, presenting a major public health challenge (*1*, *2*). The first clue in unraveling its etiologic agent was the discovery of an enigmatic serum protein named "Australia antigen" 60 years ago by Baruch Blumberg (*3*). In 1970, the Australia antigen was identified as a structural component of the hepatitis B virus (HBV) (*4*), which is the prototype member of the ancient *Hepadnaviridae*. That year, David Dane published electron micrographs of HBV 17- to 22-nm-diameter spherical and filamentous subviral particles (SVPs) and 44-nm-diameter infectious virions (the "Dane particle") (*5*). The virion is composed of a 36-nm-diameter icosahedral nucleocapsid core containing the DNA genome enveloped by a lipid membrane with embedded surface antigen proteins (HBsAg) (Fig. 1A). Its replication cycle is characterized by the production of Dane particles and a large number of noninfectious SVPs, which outnumber (even by 10,000-fold) the Dane particles (*6*) and are thought to act as decoys contributing to the immune tolerance and persistence of chronic infection (*7*).

HBsAg expression originates from a single open reading frame, encoding three isoforms of variable lengths at the N-terminal end (Fig. 1A): small (S-HBsAg), middle (M-HBsAg), and large (L-HBsAg) (*8*, *9*). S-HBsAg, the predominant isoform in infectious virions and subviral particles (*8*), contains the epitopes responsible for eliciting protective immunity from both natural infection and vaccination (*10*). It encompasses the conserved C-terminal S domain, characterized by a predicted structure of four transmembrane helices (*11*). S-HBsAg has the intrinsic ability to orchestrate the morphogenesis of subviral particles and the enveloping membrane of virions. By contrast, M-HBsAg extends with an additional glycosylated preS2 domain, and L-HBsAg further adds a preS1 domain to M-HBsAg (*8*). However, the proportion of L-HBsAg and M-HBsAg is considerably lower (only trace amounts) in subviral particles than in virions (*8*).

Presently, high-resolution structures for HBsAg or SVPs remain elusive, despite extensive studies utilizing molecular, structural, and biochemical approaches, as reviewed in (*12*, *13*). Some models at moderate resolution have been obtained by cryo–electron microscopy (cryo-EM) (*14*–*20*), atomic force microscopy (*21*), and x-ray scattering data (*22*). Proposed reconstructions from cryo-EM studies at 12- to 15-Å resolution suggest that recombinant 22-nm HBsAg SVPs, expressed in transgenic mice and yeast, exhibit octahedral symmetry (*14*, *16*), showing that HBsAg SVPs have fourfold symmetry (*14*). A recent study with a higher-resolution (6.3-Å) cryo-EM reconstruction on much smaller SVPs (16- to 17-nm diameter) from patient serum also reported octahedral symmetry (*18*). Short *et al.* found that the filament SVP had spike-like features projecting from the membrane, and in the plane of the membrane, the HBsAg proteins appear to be close-packed (*15*). Cao *et al.* scrutinized cryo-EM structures of 22-nm-diameter spherical SVPs, filament SVPs, and 44-nm-diameter Dane particles from sera of HBV carriers, at a resolution of ~30 Å (*17*), showing that HBsAg in the native spherical SVPs folds as protrusions on the surface, as do those on the native filamentous SVPs and Dane particles. Their results indicate that HBsAg is arranged in a universal, crystalline-like pattern on both native HBV viral and subviral particles, even though the whole particles may be irregularly organized.

Collectively, HBV viral and subviral particles are a complex biosystem, marked by numerous heterogeneities that might arise from some intrinsic properties inherent to the surface protein HBsAg. It is also a simple system, incorporating only lipids and a solitary HBsAg protein. The conundrum persists in comprehending how this simple system can give rise to a multitude of variable-size spheres or filaments while aligning with an internally organized regular icosahedron nucleocapsid.

Structure determination

We used mammalian cells, including a stably transfected Chinese hamster ovary (CHO) cell line and transiently transfected human embryonic kidney–293 (HEK293) cell lines, to express S-HBsAg for the production of SVP (Fig. 1B). This approach has proven sufficient for the formation of spherical particles that are indistinguishable from those formed naturally during human infection (*23*, *24*). The purified SVPs are rich in lipids (fig. S1) and can normally be recognized by neutralizing monoclonal antibodies (fig. S2). The particles showed a variety of morphologies, including filaments and spheres of various sizes (Fig. 1B and fig. S3). The gel filtration chromatography fractions (fig. S4A) with relatively uniform size were frozen into vitreous ice by plunge-freezing into liquid ethane. Cryo-EM data collection was conducted in super-resolution counting mode by use of a K3 detector mounted on a Titan Krios microscope at 1.06-Å pixel size. Cryo-EM micrographs and two-dimensional averaging (Fig. 1B) demonstrate the heterogeneity of the sample, with about 13 million particles about 22 nm in diameter of slightly varying sizes. Under the guidance of quantitative analysis of particle size and roundness [Fig. 1B and fig. S4, B and C; implementation details are provided in supplementary materials (SM), materials and methods], two groups of particles with diameters of 22 nm that are both closest to perfect spheres (fig. S4D) were iteratively enriched from the samples derived from the two

[1]Shanghai Institute for Advanced Immunochemical Studies and School of Life Science and Technology, ShanghaiTech University, Shanghai, China. [2]Shanghai Clinical Research and Trial Center, Shanghai, China. [3]National Facility for Protein Science, Shanghai Advanced Research Institute, Chinese Academy of Sciences (CAS), Shanghai, China. [4]Laboratory of Structural Biology, Tsinghua University, Beijing, China. [5]State Key Laboratory of Medicinal Chemical Biology, Frontiers Science Center for Cell Response, College of Life Sciences, and State Key Laboratory of Medicinal Chemical Biology, Nankai University, Tianjin, China. [6]National Laboratory of Biomacromolecules, CAS Center for Excellence in Biomacromolecules, Institute of Biophysics, CAS, Beijing, China. [7]University of Chinese Academy of Sciences, Beijing, China. [8]School of Chemistry and Molecular Biosciences, the University of Queensland, Brisbane, Australia. [9]National Institutes for Food and Drug Control, Beijing, China.
*Corresponding author. Email: wangq@shanghaitech.edu.cn (Q.W.); huzhy@nifdc.org.cn (Z.H.); raozh@tsinghua.edu.cn (Z.R.)
†These authors contributed equally to this work.

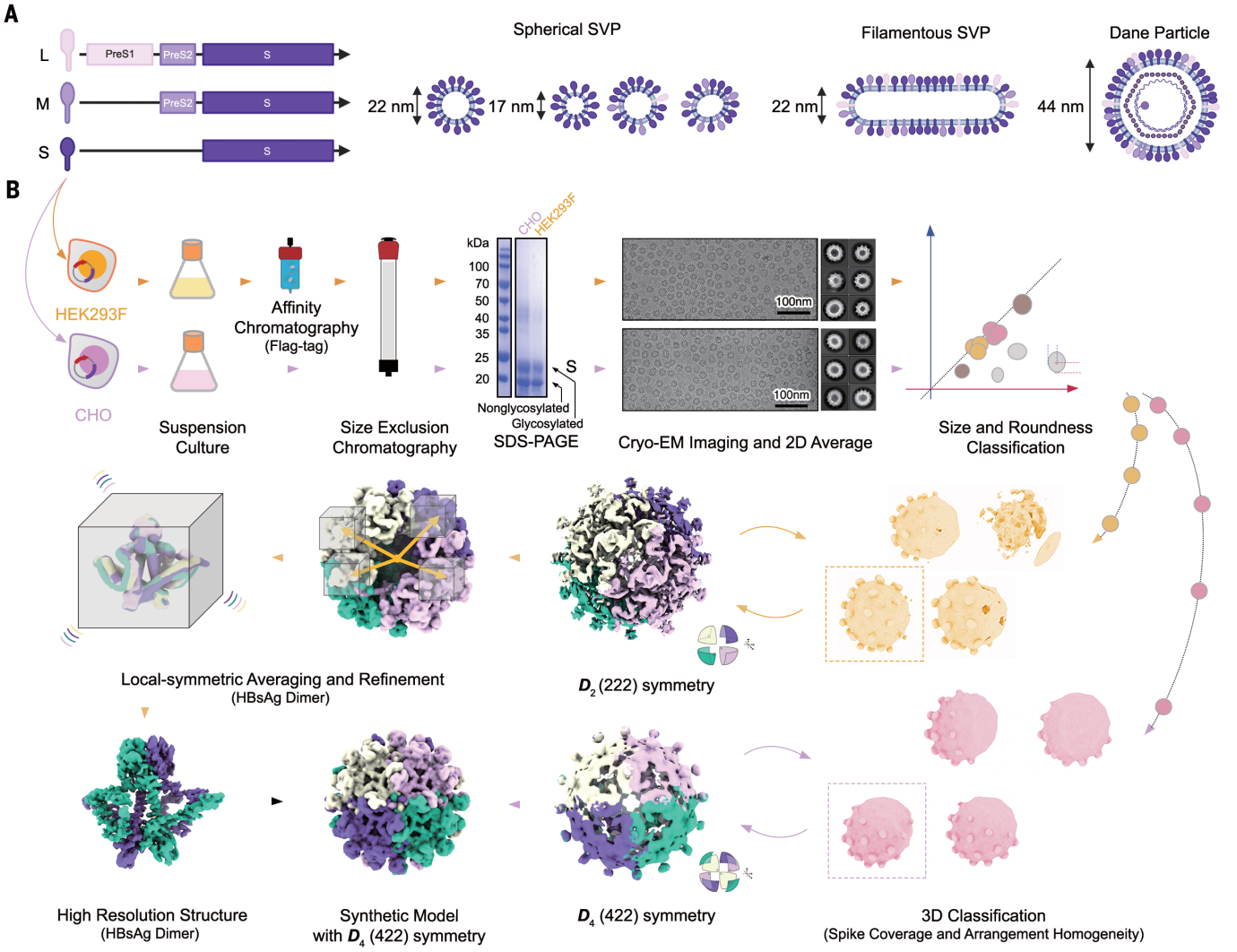

Fig. 1. Overall workflow of the study. (**A**) The HBV life cycle features 17- to 22-nm-diameter spherical and filamentous subviral particles (SVPs) and 44-nm-diameter infectious mature virions (the "Dane particle"). All types of viral particles are enveloped by the surface antigen protein HBsAg, which has three isoforms: S-HBsAg, M-HBsAg, and L-HBsAg. In this study, we focused on the 22-nm SVP particles composed entirely of S-HBsAg, prepared through recombinant expression. [Figure created with BioRender.com.] (**B**) A summary of sample preparation and cryo-EM data processing of HBsAg subviral particles from two different expression systems. In comparison with the traditional workflow, some special techniques were applied to deal with the extremely heterogeneous dataset (see details in SM, materials and methods).

expression systems, respectively (Fig. 1B). Data processing resulted in two reconstructions with $\boldsymbol{D}_2$ (222)– or $\boldsymbol{D}_4$ (422)–like quasisymmetric structure arrangements (Fig. 1B and fig. S5). After applying $\boldsymbol{D}_2$ or $\boldsymbol{D}_4$ symmetry, we obtained two maps at 6.6- and 8.5-Å resolutions, respectively (fig. S4E). For ease of reference, we named these two types of SVPs "D2" and "D4" particles. The map used for HBsAg structure interpretation is the 6.6-Å resolution map with $\boldsymbol{D}_2$ symmetry (113,000). Despite the modest resolution, the secondary structures, especially helices, are apparent in the maps (fig. S4F). In addition, there are fuzzy extra signals around the protein, presumably lipids (fig. S6). The distribution of the lipid signal is typical of a lipid bilayer. This is different from the observation in earlier studies of only lipid patches on the outer layer instead of a typical lipid bilayer for 17-nm-diameter particles (*18*). This difference may result from the relatively larger size of the particles in this study, which provides more space for lipid binding. Additionally, the lower symmetry imposed ($\boldsymbol{D}_2$) relative to octahedral ($\boldsymbol{O}$) symmetry during data processing allows more lipid signal to be retained. The SVP is a relatively dynamic membrane system, posing a challenge for higher-resolution, single-particle cryo-EM structure analysis. To further obtain the detailed structure of HBsAg, we developed an orientation transform script (*25*), which, combined with the local-symmetric averaging and refinement function in the cryo-EM software Scipion (with LocalRec plugin) and RELION, enabled us to fully utilize all the repeat units on the particle, yielding about 4.5 million subparticles (fig. S7). Through further averaging and refinement, a 3.7-Å near-atomic resolution structure was achieved, revealing side chains of the amino acid residues on the transmembrane helices (Fig. 1B and fig. S8, A and B). Aided by AlphaFold2 (*26*), we built a nearly complete model of the S-HBsAg (fig. S8C and table S1).

HBsAg structure

S-HBsAg protein is a 226–amino acid small-membrane protein but has an intrinsic ability to direct the formation of SVPs (*17*). Sequence analysis, structure prediction, and extensive mutagenesis analysis indicated that it likely

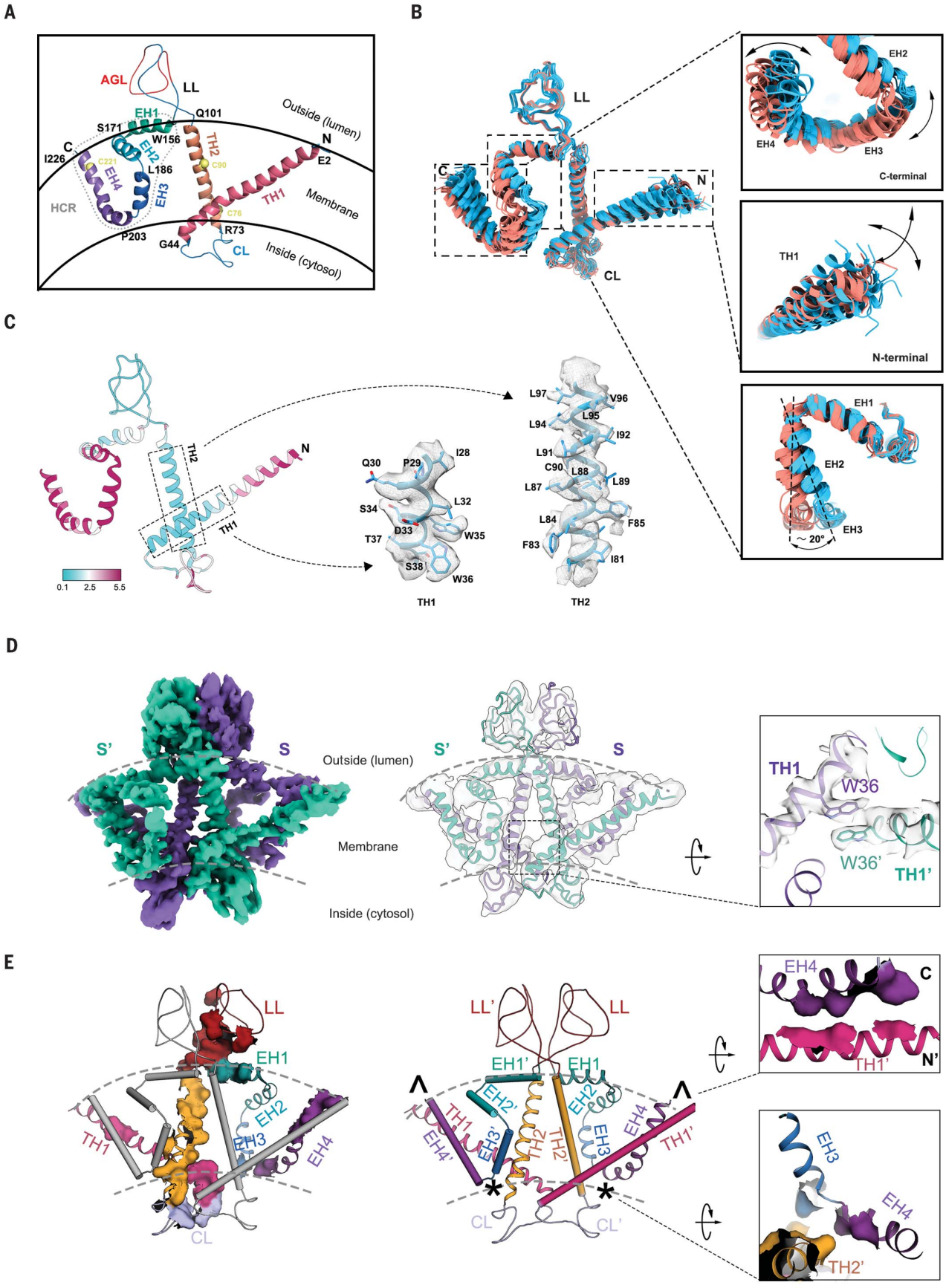

Fig. 2. The structure of HBsAg and dimerization. (**A**) The structure arrangement of HBsAg. From the N-terminal to the C-terminal, there are two transmembrane helices

(TH1 and TH2), four membrane-embedded alpha helices (H1 to EH4), and two loops [cytosolic loop (CL) and ER luminal loop (LL)] on each side of the membrane. The AGL is colored in red, and the HCR is labeled by the dashed box. (**B**) According to the high-resolution structural model of HBsAg, each HBsAg molecule within the asymmetric unit was individually modeled to fit its map features, resulting in a structural model for each molecule (see details in SM, materials and methods). This aligned conformation ensemble of HBsAg shows considerable structure variances in the three regions of the C-terminal, EH2-EH3, and N-terminal. The structure variances in EH2-EH3 can be coarsely classified into two states, which are colored respectively. (**C**) The conformational heterogeneity of the HBsAg, represented by the root mean square deviation (RMSD) level between aligned structures. Specifically, the cryo-EM map (contour level 0.3 and 0.022) and fitted atomic model of the structurally homogeneous region (part of TH1 and TH2, labeled by the dashed box) are shown, whose side chains are well resolved. (**D**) Cryo-EM maps of HBsAg dimer. (Left) The map (contour level 0.04) generated by localized reconstruction. (Middle) The map (contour level 0.3) from another localized reconstruction fitted with the atomic model. (Right) Symmetric-related tryptophan residues (W36) in the map (contour level 0.035) form the π-π stacking interaction. The model fitted into the cryo-EM map is shown, with two W36 side chains represented in sticks. Single-letter abbreviations for the amino acid residues are as follows: C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; I, Ile; L, Leu; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; and W, Trp. (**E**) The interactions between the two HBsAg protomers within one dimer. The interfaces on one protomer are shown in surface representation and colored. For clarity, the helices in one protomer are represented by cylinders, and the others are represented by ribbons. The two primary interactions are marked with symbols, and zoom views are shown in the right panel.

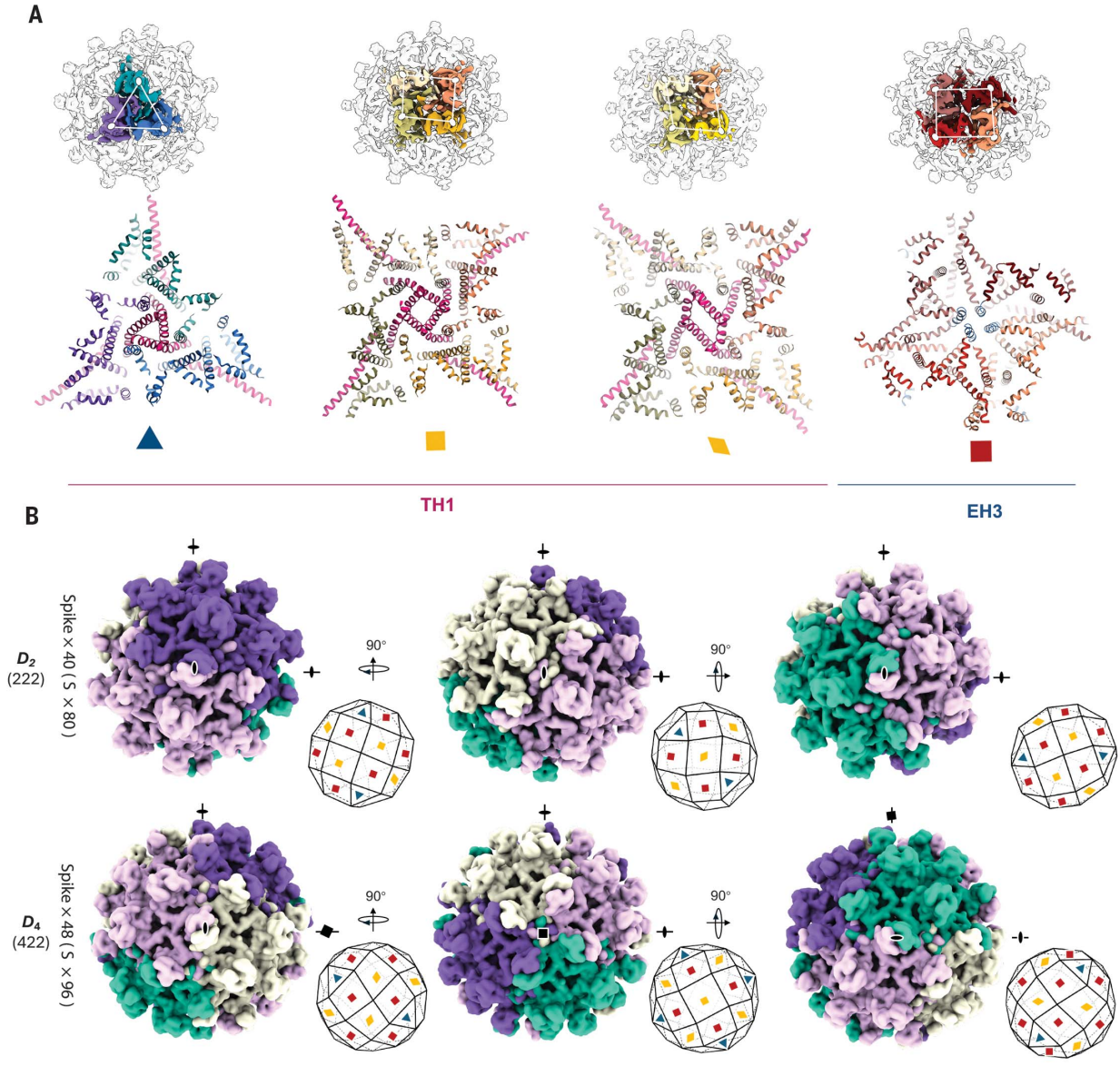

Fig. 3. Oligomerization of HBsAg. (**A**) The HBsAg dimers assemble into two types of unit blocks. The trimer and various tetramers are highlighted in the cryo-EM map of the D2 particle (top) and shown separately in a cartoon-style representation (bottom). One type of tetramer (yellow diamond and square) and the trimer (blue triangle) are primarily formed through interactions between four or three TH1s (highlighted in magenta), whereas the other type of tetramer (magenta square) is mainly formed through interactions between four EH3s (highlighted in blue). The diamond-like tetramer differs from the yellow square tetramer in that it resembles a twisted version of the square. (**B**) Three orthogonal views and stereographic perspective of the particles in $\boldsymbol{D}_2$ (222) and $\boldsymbol{D}_4$ (422) symmetry. The surface composition units of the stereographic perspective are indicated with the assembly block labels.

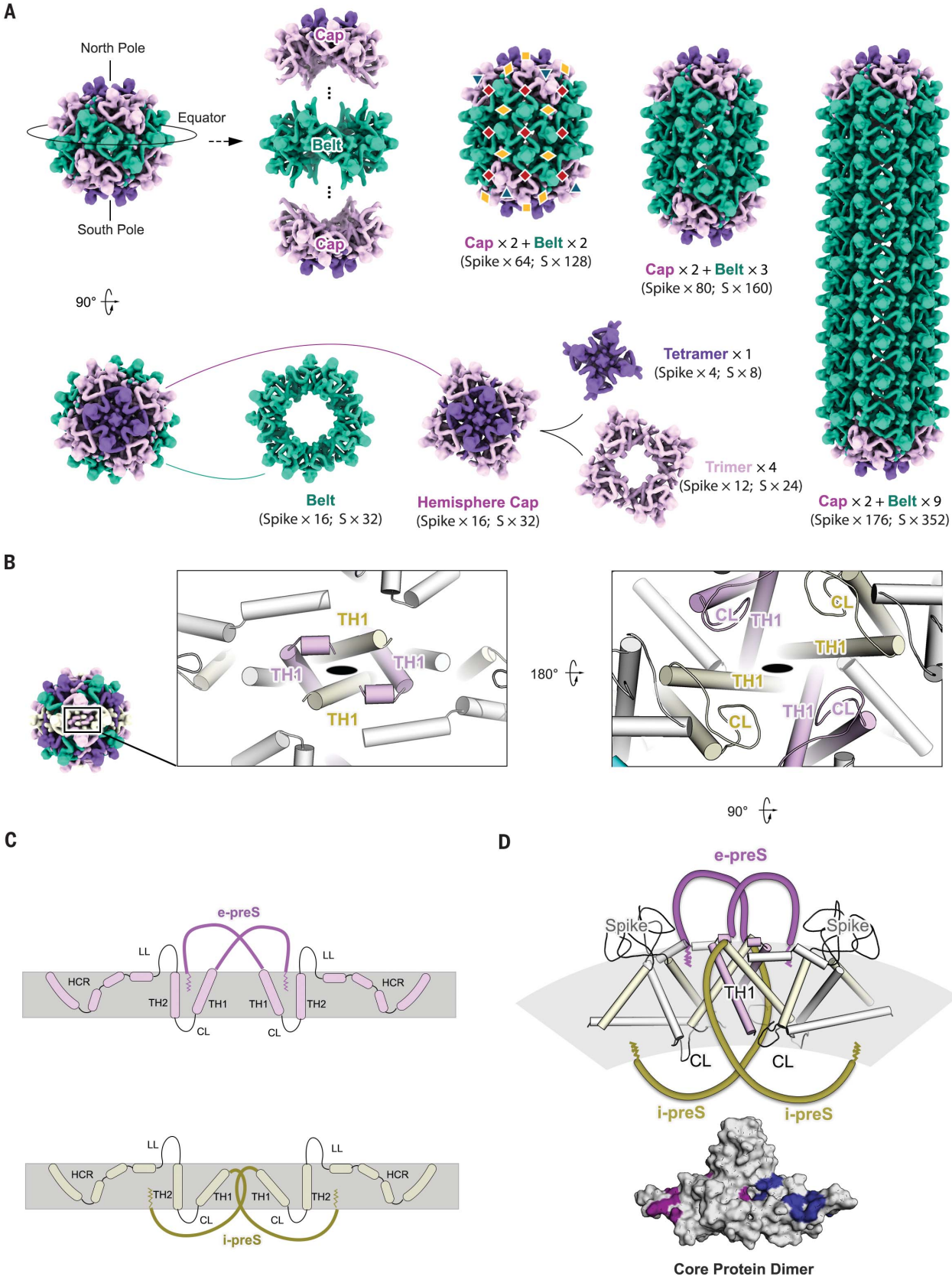

Fig. 4. Implications for filaments and virion assembly. (**A**) A possible filament-shaped SVP assembly inspired by the D4 particle. According to the D4 particle structure, the assembly could consist of two hemisphere caps around the C4 axis and one belt on the equator. By duplicating belts between two caps, a 22-nm-diameter filament particle is elongated (movie S1). Specifically, the unit blocks of two-belt filament are marked corresponding to Fig. 3A. (**B** to **D**) A possible virion

assembly. On the twofold axis of the D2 particle, two of the N terminus of HBsAg tend to adopt a conformation that preS can extend into the cytosol side (i-preS), while another two preS may be exposed to the lumen. The top view and the bottom view along with the twofold axis of the D2 particle are shown (B). The topology of dimerized HBsAg with preS highlighted (C). Side view of the twofold axis of D2 particle, showing the close location of the CL and i-preS that could enable the interaction with the core protein dimer. The model of core protein (PDB: 1QGT) (*34*) is shown in surface representation, and the amino acid residues in contact with the HBsAg are colored on each core protein protomer (D).

has a complex topology of more than four membrane-spanning and embedding helices (*11*). The structure that we observed now explains and revises the previous observations and predictions.

As illustrated in Fig. 2A, two N-terminal transmembrane alpha helices, TH1 and TH2, span residues 4 to 40 and 73 to 100, respectively. Four membrane-embedded alpha helices, EH1 to EH4, spanning residues 156 to 170, 173 to 185, 189 to 198, and 202 to 225, respectively, direct a distinctive topogenesis of the hydrophobic C-terminal region (HCR, amino acids 173 to 226) of S-HBsAg protein. The residues between TH1 and TH2 form an endoplasmic reticulum (ER) cytosolic loop (CL) inside the particle, and an ER luminal loop (LL), located between TH2 and the HCR, forms the antigenic loop (AGL) that protrudes from the particle surface. The spatial distribution of hydrophobic amino acids demonstrates the boundary of the lipid membrane (fig. S9). This pattern is consistent with a previously proposed model for the arrangement of proteins in SVP, whereby regions within the N-terminal and C-terminal parts of the two major HBsAg components are buried within the lipid matrix, and the antigenically essential residue 122 to 150 region (AGL) is exposed to the aqueous environment (*27*). The location of the HCR relative to the membrane has been elusive. Its C terminus is oriented toward the ER lumen (*28*), and a middle region (amino acids 196 to 201) is supposed to face the cytoplasm (*29*), suggesting that it has two transmembrane areas. Although the HCR does bend several times, it does not appear to cross the membrane fully but is more half-embedded in the membrane (Fig. 2A). When comparing all the structures of every S-HBsAg within one asymmetric unit patch, the HCR shows considerable structural heterogeneity (Fig. 2B), which will also account for the heterogeneity of the HBsAg oligomers (see below). This finding also explains why only some structurally homogeneous regions achieve higher resolution (Fig. 2C). HBsAg in the 17-nm particle (*18*) shows a similar overall topology but without heterogeneity (fig. S10). HBsAg is highly cross-linked by disulfide bonds in virions and SVPs. The high-resolution structure also corroborates the result of the previous alkylation studies on the 22-nm SVP secreted from cell lines (*30*), which show that C76, C90, and/or C221 carry the free sulfhydryl group(s) (Fig. 2A). Although the resolution is relatively lower on the spikes, the traced backbone of the AGL suggests three pairs of intramolecular disulfide bonds (C121-C124, C137-C149, and C139-C147) and one pair of intermolecular disulfide bonds between the two C107 from each protomer of HBsAg dimer (fig. S11).

HBsAg oligomerization

The two SVP structures resolved in this work only contain S-HBsAg, meaning that they are both solely formed through the oligomerizations of HBsAg. As previously demonstrated by biochemical data (*30*), HBsAg can form homodimers. Its dimer is proposed to be the major spike-like feature projecting from the membrane in a previous low-resolution cryo-EM study (*15*). In our structure, two S-HBsAg are tightly locked together (Fig. 2D). From a macroscopic perspective, it even seems that the dimerization maintains the topology and stability of this structure. TH1, CL, TH2, LL, and the HCR all contribute to the interface with more than 1300 $Å^2$ of buried surface area in total (Fig. 2E). Near the central dimer axis, CL and LL and their adjacent regions form interactions with their symmetry-related counterparts on each side of the membrane, respectively. Two symmetry-related tryptophans on TH1 form a π–π stacking interaction on the twofold axis (Fig. 2D). A HBsAg variant harboring a W36L mutation has been reported to exhibit a secretion-defective phenotype and universally induced ER stress pathways in hepatocytes (W36L indicates that tryptophan at position 36 was replaced with leucine) (*31*). In the far off-axis region, the luminal N-terminal leading of TH1 of one protomer contacts with the very C-terminal tail from the other protomer, and the head tip of its TH2 interacts with a near cytosolic turn between EH3 and EH4 in the HCR of the counter protomer, which appears to constrain the further stretching of EH3 to the cytosolic side, resulting in a kinked EH2-EH3 helix that looks like it should penetrate the membrane but doesn't (Fig. 2E).

TH1, which is the most distal region from the central axis of the dimer, not only participates in the interaction between the two dimeric protomers but also exhibits an unusual ability to mediate oligomerization. It coordinates the HBsAg dimers further into two different higher-order assemblies: trimers of dimers and tetramers of dimers (Fig. 3A). The trimerization of HBsAg dimers makes the three spikes show a regular triangle arrangement, whereas the spikes in the tetramers are more in a nonregular quadrilateral distribution. The EH3s at the side of the dimers mediate assembly into another type of tetramer, which seems more compact but also not strictly symmetric or even not homogeneous (Fig. 3A and movie S2).

SVP architecture

The two types of SVP with two different symmetries exhibit near-spherical shapes of a similar size, with maximum diameters of 220 Å (or 280 Å including spikes) (fig. S12A and movie S1). This differs from the octahedral particle (*18*), which has a diameter of 174 Å (or 224 Å including spikes) (fig. S13). The higher-resolution D2 particle clearly shows that its basic component unit is the homodimer of S-HBsAg protein, and each spike on the surface is formed by two oligomerized LLs from two S-HBsAg protomers (fig. S12B). Measured by the number of spikes, there are 80 S-HBsAg proteins (40 spikes) in the D2 particle, whereas the D4 particle has 96 S-HBsAg proteins (8 more spikes). The HBsAg dimers as a protomer form 8 trimers and 34 tetramers that densely pack in the D2 particle (Fig. 3B and movie S1). If L-HBsAg or M-HBsAg proteins are incorporated into this SVP, the preS1/S2 will be positioned at the oligomerization centers surrounded by the spikes (fig. S14A). Although the number and arrangement of the trimers and tetramers in the D4 particles differ from that in the D2 particles, they share the same set of unit blocks and a common dense-packing rule (movie S1). Therefore, under the constraints of spike distribution and arrangement on the D4 particle, we can build up a detailed synthetic structure model of the D4 particle with the component blocks from the D2 particle (Figs. 1B, 3B, and movie S1). Some excluded particles can also form a certain degree of local alignment (fig. S15A). The aligned region not only has the same patch units as D2 and D4 particles, but the overall spike distribution is also similar to one face of the D2 particle (fig. S15B). This suggests that the non-D2 and non-D4 SVPs, at least partially, share the same structural organization as the D2 or D4 particles. In previous analyses of samples sourced from diverse origins, albeit constrained by lower resolution at the time, subtle indications of trimeric (*17*, *19*, *20*) or tetrameric (*15*, *17*) oligomerization features of the spikes were discernible. When comparing the 17-nm octahedral particle (*18*) with D2 and D4 particles, there appears to be a common local assembly rule with D2 and D4 particles (fig. S16). However, the octahedral particle lacks the diamond-like tetramer assembly pattern found in the other two. This difference gives the SVP a lower curvature, larger radius, and greater flexibility over the entire sphere, which we observe in

the D2 and D4 particles and predict in filaments and Dane particles.

SVPs are assembled from only one protein, HBsAg, and as the basic unit of the entire particle structure, HBsAg shows several different oligomerization forms. This is reminiscent of the assembly of the HIV capsid, in which the capsid protein is oligomerized into both pentameric and hexamer forms to support the capsid forming an unusual cone-shaped structure (*32*). Unlike HIV capsid, where the capsid protein is packed into the cone surface as a component of either a pentamer or a hexamer unit, one HBsAg molecular in the SVPs can participate in the assembly of a dimer, trimer, and tetramer simultaneously (fig. S17A and movie S1). The structural flexibility of the HBsAg and the diverse interaction between them (fig. S17B) provide further assembly flexibility and variance. This adaptable structure arrangement is precisely functionally aligned with the compact HBV genome (*33*). Although it is not as complicated as HBsAg, the same adaptability has been shown in the HBV core protein, which oligomerizes into dimers, pentamers, and hexamers to form a regular icosahedral nucleocapsid (*34*).

Implications for filaments and virion assembly

S-HBsAg in the two SVP structures presented here demonstrates its ability to pack densely into a spherical envelope. It is reminiscent of a study that demonstrated that filaments also were assembled and released from cells even when only the S-HBsAg protein was expressed (*35*). In our recombinant expressed samples, in addition to a large number of spherical particles, filament-like particles also appeared (fig. S3, B and C). The width of these filament particles is similar to the diameter of spherical particles at 22 nm, despite their various lengths. It has been previously proposed that it is unlikely for the filament and spherical particles to have radically different packing structures because the ends of the tubes are closed off by hemispherical caps, which are equivalent to half an isometric spherical particle, without a discontinuity between the tube and the cap (*15*). These observations lead us to believe that the enveloping of the L-HBsAg and M-HBsAg containing filament-shaped SVPs and even virion particles are also mainly mediated by this shared S domain, whereas the preS1/S2 from L- and M-HBsAg mediate diverse functions in nucleocapsid binding and receptor recognition (*36*).

The diverse and flexible oligomerization forms of HBsAg, especially its symmetrical sphery distribution, exhibit unexpected potential for assembly extension. Specifically, if we take the D4 SVP as a combination of two hemispheres, the two tetramers at the north and south poles and the four adjacent trimers around them will appear to be two cap-like structures, and a belt-like region near the equator can be extended by stacking different copies, making the SVP into filaments with various lengths (Fig. 4A and movie S1). If L- and M-HBsAg are incorporated, the preS1/S2 will be positioned at the oligomerization centers surrounded by the AGL spikes (fig. S14B). This extension is solely based on the assembly rules found in the D4 particle (Figs. 3A and 4A) but does not introduce any new interaction modes, aligning well with the filament particle morphology seen in electron micrographs (fig. S18). By contrast, the D2 particle does not seem to have the same ability to expand (fig. S19).

How the HBsAg embedded in the envelope can align with an internally organized regular icosahedral nucleocapsid has been unclear but involves the unusual topology of L-HBsAg. During virion morphogenesis, M-HBsAg is dispensable for the production of infectious Dane particles (*37*, *38*), whereas the L-HBsAg is essential, and its preS region is believed to adopt a mixed topology, with some molecules having the preS region internal (i-preS), on the cytoplasmic/nucleocapsid facing side of the membrane, and others having it external (e-preS), on the luminal/outer surface of the membrane (*39*–*41*). Whereas the e-preS form exposed on the virion surface is believed to facilitate subsequent attachment and entry of the viral particle (*42*–*44*), a short conserved region (between residues 96 and 116) of the cytosolic i-preS1 and part of the cytosolic loop is essential for the interaction with nucleocapsid and subsequent virion envelopment (*45*–*48*).

22-nm SVPs are only half the size of the Dane particle in diameter (which is about 44 nm); thus, they cannot be directly used to interpret the structure of the capsid-enveloped Dane particle. However, the structure helps in understanding the dual topology of L-HBsAg. The N-terminal leading of TH1 penetrates the membrane into the luminal/outer surface and mediates the trimeric and two tetrameric oligomerizations of the HBsAg dimer. In the trimers, TH1 is unquestionably all exposed on the outside surface (e-preS). But in some diamond-like tetramers, the N terminuses of two TH1s are exposed (e-preS), and the other two tend to be buried in the membrane (i-preS) (Fig. 4B). If the two i-preS further extend into and across the membrane from here, they will be quite close to the four adjacent cytosolic loops on the other side of the membrane (Fig. 4, B and C). Previous comprehensive alanine mapping of the nucleocapsid, evaluating the mutants' impact on capsid formation and virion secretion, revealed a series of critical amino acids interacting with HBsAg (*49*, *50*). These residues on the solvent-exposed face of the nucleocapsid are symmetrically clustered on both sides of the nucleocapsid spike (Fig. 4D). The potential interaction of i-preS1 and part of the cytosolic loop with the nucleocapsid through these amino acids during virion envelopment appears plausible (Fig. 4D). Our study thus provides structural insight from the envelope side into the interaction between the nucleocapsid and envelope on the twofold axis.

Discussion

The question of how a protein molecule's properties and structural adaptability allow it to create a wide variety of particle sizes through dense packing on the membrane, which originated with Dane's pioneering electron micrograph half a century ago, has long been an enigma for pathologists, virologists, and molecular biologists. Our work provides long-anticipated structures of HBsAg and SVPs, offering near-atomic resolution insights into the diverse trimeric and tetrameric oligomerization of HBsAg dimer, leading to the assembly of 22-nm subviral particles with $\boldsymbol{D}_2$- or $\boldsymbol{D}_4$-like quasisymmetry. These findings challenge the conventional view that surface proteins are randomly distributed across the viral envelope without meaningful organization.

Our results and methodology may be useful for addressing the remaining crucial challenges in understanding HBV virion structure and assembly. These issues include structural characterization of irregular SVPs and the intact 44-nm Dane particle, especially the involvement of L-HBsAg and M-HBsAg components, comprehending the binding mode to the viral receptor, and deciphering the recognition patterns of neutralizing and non-neutralizing antibodies. Regarding the mixed topology of L-HBsAg, there are still various possibilities. Seitz *et al.* (*13*) have proposed a virion maturation step during which, at a certain stage, interactions between the capsid and i-preS subunits are lost. After detachment from the nucleocapsid, a topological switch of L-HBsAg occurs, allowing preS to translocate across the viral membrane and become exposed on the particle surface. Such possibilities require the structural study of a complete virion particle containing L-HBsAg to be verified.

REFERENCES AND NOTES

1. World Health Organization, *Global Progress Report on HIV, Viral Hepatitis and Sexually Transmitted Infections, 2021: Accountability for the Global Health Sector Strategies 2016–2021: Actions for Impact* (World Health Organization, 2021).
2. A. Kocher *et al.*, *Science* **374**, 182–188 (2021).
3. B. S. Blumberg, *Hepatitis B: The Hunt for a Killer Virus*. (Princeton Univ. Press, 2003).
4. B. S. Blumberg, A. I. Sutnick, W. T. London, *Am. J. Med.* **48**, 1–8 (1970).
5. D. S. Dane, C. H. Cameron, M. Briggs, *Lancet* **295**, 695–698 (1970).
6. D. Ganem, A. M. Prince, *N. Engl. J. Med.* **350**, 1118–1129 (2004).
7. G. E. Rydell, K. Prakash, H. Norder, M. Lindh, *Virology* **509**, 67–70 (2017).
8. K. H. Heermann *et al.*, *J. Virol.* **52**, 396–402 (1984).
9. C. Seeger, F. Zoulim, W. S. Mason, in *Fields Virology*, 7D. M. Knipe, P. M. Howley, Eds. (Lippincott, Williams & Wilkins, ed. 5, 2007), pp. 2977–3029.

10. A. Bertoletti, C. Ferrari, *J. Hepatol.* **64**, S71–S83 (2016).
11. H. J. Stirk, J. M. Thornton, C. R. Howard, *Intervirology* **33**, 148–158 (1992).
12. L. Berthier, O. Brass, G. Deleage, R. Terreux, *J. Mol. Graph. Model.* **98**, 107610 (2020).
13. S. Seitz, J. Habjanič, A. K. Schütz, R. Bartenschlager, *Annu. Rev. Virol.* **7**, 263–288 (2020).
14. R. J. Gilbert *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **102**, 14783–14788 (2005).
15. J. M. Short, S. Chen, A. M. Roseman, P. J. Butler, R. A. Crowther, *J. Mol. Biol.* **390**, 135–141 (2009).
16. A. M. Mulder *et al*., *PLOS ONE* **7**, e33235 (2012).
17. J. Cao *et al*., *Virus Res.* **259**, 90–96 (2019).
18. H. Liu *et al*., *Sci. Adv.* **8**, eabo4184 (2022).
19. K. A. Dryden *et al*., *Mol. Cell* **22**, 843–850 (2006).
20. S. Seitz, S. Urban, C. Antoni, B. Böttcher, *EMBO J.* **26**, 4160–4167 (2007).
21. P. E. Milhiet *et al*., *Biochimie* **93**, 254–259 (2011).
22. L. P. Aggerbeck, D. L. Peterson, *Virology* **141**, 155–161 (1985).
23. C. C. Liu, D. Yansura, A. D. Levinson, *DNA* **1**, 213–221 (1982).
24. O. Laub *et al*., *J. Virol.* **48**, 271–280 (1983).
25. T. Wang, Localized reconstruction of HBV SVP with elliptic estimation and Euler angles transformation, Zenodo, (2024); https://doi.org/10.5281/zenodo.11406516.
26. J. Jumper *et al*., *Nature* **596**, 583–589 (2021).
27. F. Gavilanes, J. M. Gonzalez-Ros, D. L. Peterson, *J. Biol. Chem.* **257**, 7770–7777 (1982).
28. B. E. Eble, V. R. Lingappa, D. Ganem, *Mol. Cell. Biol.* **6**, 1454–1463 (1986).
29. I. Komla-Soukha, C. Sureau, *J. Virol.* **80**, 4648–4655 (2006).
30. C. M. Mangold, F. Unckell, M. Werr, R. E. Streeck, *Arch. Virol.* **142**, 2257–2267 (1997).
31. I. K. Lee, S. A. Lee, H. Kim, Y. S. Won, B. J. Kim, *World J. Gastroenterol.* **21**, 6872–6883 (2015).
32. G. Zhao *et al*., *Nature* **497**, 643–646 (2013).
33. R. H. Miller, S. Kaneko, C. T. Chung, R. Girones, R. H. Purcell, *Hepatology* **9**, 322–327 (1989).
34. S. A. Wynne, R. A. Crowther, A. G. Leslie, *Mol. Cell* **3**, 771–780 (1999).
35. N. Chai *et al*., *J. Virol.* **82**, 7812–7817 (2008).
36. T. F. Baumert *et al*., *Curr. Opin. Virol.* **4**, 58–65 (2014).
37. V. Bruss, D. Ganem, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 1059–1063 (1991).
38. D. Fernholz *et al*., *Virology* **194**, 137–148 (1993).
39. V. Bruss, X. Lu, R. Thomssen, W. H. Gerlich, *EMBO J.* **13**, 2273–2279 (1994).
40. P. Ostapchuk, P. Hearing, D. Ganem, *EMBO J.* **13**, 1048–1057 (1994).
41. R. Prange, R. E. Streeck, *EMBO J.* **14**, 247–256 (1995).
42. P. Gripon, J. Le Seyec, S. Rumin, C. Guguen-Guillouzo, *Virology* **213**, 292–299 (1995).
43. J. Le Seyec, P. Chouteau, I. Cannie, C. Guguen-Guillouzo, P. Gripon, *J. Virol.* **73**, 2052–2057 (1999).
44. C. Lepère-Douard, M. Trotard, J. Le Seyec, P. Gripon, *J. Virol.* **83**, 11819–11829 (2009).
45. V. Bruss, *J. Virol.* **71**, 9350–9357 (1997).
46. F. Poisson, A. Severac, C. Hourioux, A. Goudeau, P. Roingeard, *Virology* **228**, 115–120 (1997).
47. J. Le Seyec, P. Chouteau, I. Cannie, C. Guguen-Guillouzo, P. Gripon, *J. Virol.* **72**, 5573–5578 (1998).
48. H. Löffler-Mary, J. Dumortier, C. Klentsch-Zimmer, R. Prange, *Virology* **270**, 358–367 (2000).
49. D. Ponsel, V. Bruss, *J. Virol.* **77**, 416–422 (2003).
50. H. Liu, J. Xi, J. Hu, *J. Virol.* **95**, e00230-21 (2021).

ACKNOWLEDGMENTS

We thank Y. Zhang from the Discovery Technology Platform and Analytical Chemistry Platform at the Shanghai Institute for Advanced Immunochemical Studies, ShanghaiTech University, for her technical support on surface plasmon resonance analysis. We thank L. Jiang and W. Wang from the Biomedical Big Data Platform at Shanghai Institute for Advanced Immunochemical Studies, ShanghaiTech University, for their support on data management and analysis. We thank Y. Han and P. Hao from the Multi-Omics Facility in the School of Life Science and Technology, ShanghaiTech University, for their support of the lipidomics analysis. We thank N.-S. Xia's team at Xiamen University, China, for providing an HBsAg conformational epitope recognition monoclonal antibody A2C1. We would especially like to thank the Electron Microscopy Facility of ShanghaiTech University, Shanghai Institute of Precision Medicine, and the National Facility for Protein Science (Shanghai) for their great support. This work was also supported by the Shanghai Frontiers Science Center for Biomacromolecules and Precision Medicine at ShanghaiTech University. **Funding:** This research was supported by Strategic Priority Research Program of the Chinese Academy of Sciences grants XDB08020200 (Z.R.) and XDB37020203 (Q.W.), Science and Technology Commission of Shanghai Municipality grant YDZX20203100001571 (Q.W.), National Program on Key Research Project of China grant 2017YFC0840300 (Z.R.), National Program on Key Research Project of China grant 2020YFA0707500 (Z.R.), National Natural Science Foundation of China grant 81520108019 (Z.R.), National Natural Science Foundation of China grant 813300237 (Z.R.), and National Natural Science Foundation of China grant 31971118 (Q.W.) **Author contributions:** Z.R., Z.H., and Q.W. conceived, initiated, and coordinated the project; L.C. and P.W. expressed the protein and purified the SVPs; T.W., A.M., and L.C. collected the cryo-EM data; T.W. and A.M. processed the cryo-EM data; A.M. built and refined the structure model; Q.W., T.W., A.M., P.W., L.C., and Z.R. wrote the manuscript. All authors discussed the experiments and results and read and approved the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** The cryo-EM maps have been deposited at the Electron Microscopy Data Bank under accession codes EMD-39395 (D2 map), EMD-39397 (D4 map), EMD-39396 (localized reconstructed map A0), and EMD-39404 (localized reconstructed map A). The atomic models have been deposited at the Protein Data Bank under accession codes 8YMJ (D2 map) and 8YMK (localized reconstructed map A0). Localized reconstruction of HBV SVP with elliptic estimation and Euler angles transformation have been deposited at Zenodo (*25*). All other data are available in the main text or the supplementary materials. The stably transfected CHO cell line, which expresses S-HBsAg for the production of SVP, and the purified SVPs derived from it, are provided by North China Pharmaceutical Genetech Biotechnology Co. under a material transfer agreement with the National Institutes for Food and Drug Control. We extend our gratitude for their generous sharing of materials and technical support.

SUPPLEMENTARY MATERIALS

Submitted 9 March 2024; accepted 12 August 2024
10.1126/science.adp1453