Proc. Natl. Acad. Sci. USA
Vol. 94, pp. 9556–9561, September 1997
Biochemistry

Localization of the C terminus of the assembly domain of hepatitis B virus capsid protein: Implications for morphogenesis and organization of encapsidated RNA

A. ZLOTNICK†, N. CHENG‡, S. J. STAHL†, J. F. CONWAY‡, A. C. STEVEN‡, AND P. T. WINGFIELD†§

†Protein Expression Laboratory, ‡Laboratory of Structural Biology, National Institute of Arthritis, Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, MD 20892

Communicated by Michael G. Rossmann, Purdue University, West Lafayette, IN, July 3, 1997 (received for review April 1, 1997)

ABSTRACT **The capsid protein of hepatitis B virus, consisting of an "assembly" domain (residues 1–149) and an RNA-binding "protamine" domain (residues 150–183), assembles from dimers into icosahedral capsids of two different sizes. The C terminus of the assembly domain (residues 140–149) functions as a morphogenetic switch, longer C termini favoring a higher proportion of the larger capsids, it also connects the protamine domain to the capsid shell. We now have defined the location of this peptide in capsids assembled *in vitro* by engineering a mutant assembly domain with a single cysteine at its C terminus (residue 150), labeling it with a gold cluster and visualizing the cluster by cryo-electron microscopy. The labeled protein is unimpaired in its ability to form capsids. Our density map reveals a single undecagold cluster under each fivefold and quasi-sixfold vertex, connected to sites at either end of the undersides of the dimers. Considering the geometry of the vertices, the C termini must be more crowded at the fivefolds. Thus, a bulky C terminus would be expected to favor formation of the larger (T = 4) capsids, which have a greater proportion of quasi-sixfolds. Capsids assembled by expressing the full-length protein in *Escherichia coli* package bacterial RNAs in amounts equivalent to the viral pregenome. Our density map of these capsids reveals a distinct inner shell of density—the RNA. The RNA is connected to the protein shell via the C-terminal linkers and also makes contact around the dimer axes.**

Hepatitis B virus (HBV) causes acute and chronic hepatitis and is an important cause of cirrhosis of the liver and hepatocellular carcinoma (for review see refs. 1 and 2). In its replication cycle, virion assembly begins with polymerization of the capsid protein (Cp) to enclose a complex of the 3.2-kilobase RNA pregenome and reverse transcriptase (3). Upon completion of capsid assembly, the pregenome is retrotranscribed (4, 5). The resulting DNA-containing nucleocapsid then acquires a lipoprotein envelope en route from the cell (6).

The HBV Cp, also known as core antigen or HBcAg, is a 183-residue protein with two domains: the assembly domain that forms the contiguous shell, and the protamine domain that is responsible for RNA packaging (7, 8) (Fig. 1*a*). Dimeric Cp (9, 10) spontaneously assembles into icosahedral capsids when expressed in heterologous systems (11) or from purified protein (10). These capsids are mixtures of particles with T = 3 and T = 4 icosahedral symmetry (10, 12, 13). T = 4 capsids predominate *in vivo* (14). Their structures have been determined in density maps calculated from cryo-electron micrographs (13–15). Both capsids are composed of dimers with the overall shape of an inverted capital T: the stem, 50 Å long, is bisected by the dimer interface, and its outer portion protrudes ≈25 Å outward from the contiguous shell as a spike; the crosspieces are ≈70 Å long and pack together to form the framework of the shell. Much of the substructure of this α-helical protein (10, 16) now has been delineated in cryomicroscopy-derived density maps at resolutions below 10 Å (17, 18).

Using Cp constructs progressively truncated from the C terminus, we found that residues 138–149 play an influential role in morphogenesis *in vitro* (15). About 95% of the capsids assembled from a 149-residue Cp (Cp149) have the T = 4 morphology; decreasing to ≈20% for Cp140 (see Fig. 1). Further truncating the protein by two residues renders it incapable of assembly (15, 19, 20). When full-length Cp183 is expressed in *Escherichia coli*, RNA-containing capsids are formed (11, 21), ≈90% T = 4. In our expression system, these capsids contain 25% RNA by weight (10), which corresponds to about 3,400 nucleotides for a T = 4 capsid—the same amount of RNA as the viral pregenome. However, the RNAs in the expressed capsids are bacterial in origin and small in size.

The location of the C terminus is of interest from two perspectives: (*i*) as an important determinant of capsid assembly; and (*ii*) for its role in the encapsidation and organization of the pregenome. We have addressed this question experimentally by labeling residue 150 of Cp (Fig. 1*a*) with an 11-gold atom cluster that has been used successfully in several previous studies (e.g., refs. 22–24). To obtain specific labeling, we used the maleimidyl derivative, which reacts with cysteines. The native assembly domain has three cysteines, at positions 48, 61, and 108. The first two can form disulfide bonds with the corresponding cysteines of the dyad-related monomer (16, 25), but are not required for capsid formation (26). Our experiments were performed with a 150-residue construct in which the three native cysteines were substituted by alanine, and an additional cysteine was appended to the assembly domain. We find that the terminal peptide of the assembly domain is at the ends of the Cp dimers; we also examined its relationship with encapsidated RNA in image reconstructions of RNA-containing capsids.

MATERIALS AND METHODS

Capsid Protein Constructs. The Cp*149 and Cp*150 proteins (for nomenclature see Fig. 1 legend) were produced in *E. coli* as previously described (15) using the expression vector pET11a and bacterial strain BL21 (DE3) (27). PCR methods (28) were used to change the cysteine codons to those encoding alanine and to add a cysteine codon at the carboxy terminus (Fig. 1*a*). The overexpressed protein was purified as described

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

0027-8424/97/949556-6$0.00/0
PNAS is available online at http://www.pnas.org.

Abbreviations: Au11, monomaleimidyl undecagold; Cp, capsid protein; HBV, hepatitis B virus.

§To whom reprint requests should be addressed at: Building 6B, Room B130, 6 Center Drive, MSC2755, National Institutes of Health, Bethesda, MD 20892-2755. e-mail: PELPW@helix.nih.gov.

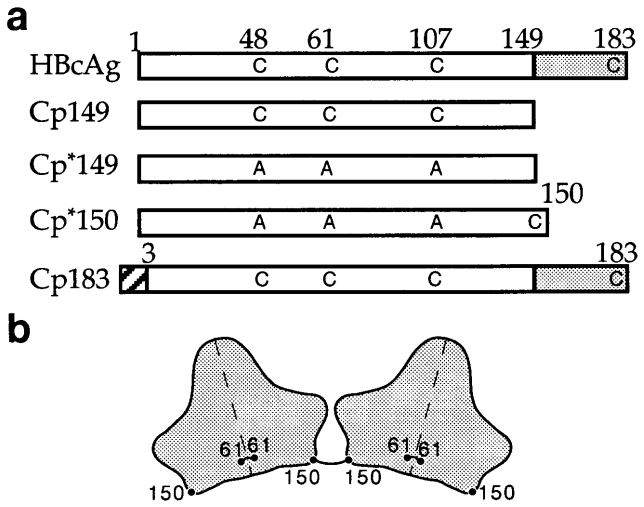

FIG. 1. Schematics of Cp sequences (*a*) and disulfide bonding (*b*). (*a*) The constructs are referred to as Cp followed by the number of the C-terminal residue, thus, Cp149 is a 149-residue protein. In Cp*149 the cysteines found in the native sequence, residues 47, 61, and 107, are mutated to alanine. Cp*150 has arginine 150 replaced by a C-terminal cysteine. The protamine domain of Cp183, residues 150–183, is shaded. The Cp183 construct has eight N-terminal residues (hatched box) derived from a cloning vector (29). (*b*) Shown are silhouettes of two Cp150 dimers connected by a disulfide bond between Cys-150 residues. Each molecule also contains an intradimer disulfide bond linking its two Cys-61 residues across the dyad axis (16, 25). The position of the latter bond has still not been fully defined (see refs. 17 and 18). The position of cysteine 150 also marks the beginning of the protamine domain in reconstructions of Cp183.

(15), except that for Cp*150, all buffers included 5 mM DTT. Purified Cp*150 dimer was stored in 100 mM sodium bicarbonate, pH 9.5 (storage buffer) with 5 mM DTT at −80°C. RNA-containing capsids of the Cp183 construct (29) were previously characterized (10).

Preparation and Biochemical Studies of Gold-Labeled Cp*150. Cp*150 capsids were assembled in 100 mM sodium phosphate, pH 6.9, 300 mM sodium chloride, by exchanging buffer in a PD10 column (Pharmacia). Identical results were obtained using 100 mM Hepes, pH 7.5, 300 mM sodium chloride. To prevent sulfhydryl-disulfide exchange on nonreducing SDS/PAGE, samples were alkylated with 20 mM iodoacetamide for 30 min before denaturation in SDS-containing gel-loading buffer. Capsids to be dissociated had their buffer exchanged with 100 mM sodium bicarbonate, pH 9.5, followed by addition of buffered urea to final concentrations of 3.5 M and <1 mg/ml protein. Assembly of some samples was characterized by sedimentation through 5–30% sucrose gradients (15).

Immediately before labeling, Cp*150 was incubated for 1 hr in fresh 50 mM DTT. The DTT was separated from the protein, and the buffer exchanged for 100 mM sodium phosphate, pH 6.9, using a PD10 column. The concentration of Cp*150 was determined by absorbance, using ε_{280} = 29,500 $M^{-1}\cdot cm^{-1}$. In a typical reaction, 12 nmols of reduced Cp*150 were mixed with 6 nmols of monomaleimidyl-undecagold (Au11) (30) (Nanoprobes, Stony Brook, NY) to a final volume of 300 μl and incubated for 24 hr at 4°C. The reactants were separated on a Sephacryl S-200 HR column (1.4 cm diameter × 6 cm), eluted in 0.5-ml fractions with 100 mM sodium phosphate, pH 6.9. Capsids eluted in the void volume at 3–4 ml, free dimers at 5–6 ml, and free Au11 at about 8 ml (Fig. 2).

Fractions were assayed for protein and Au11 content by absorbance spectroscopy. Au11 has absorbance maxima at 420 nm and 310 nm with extinction coefficients of 47,000 and

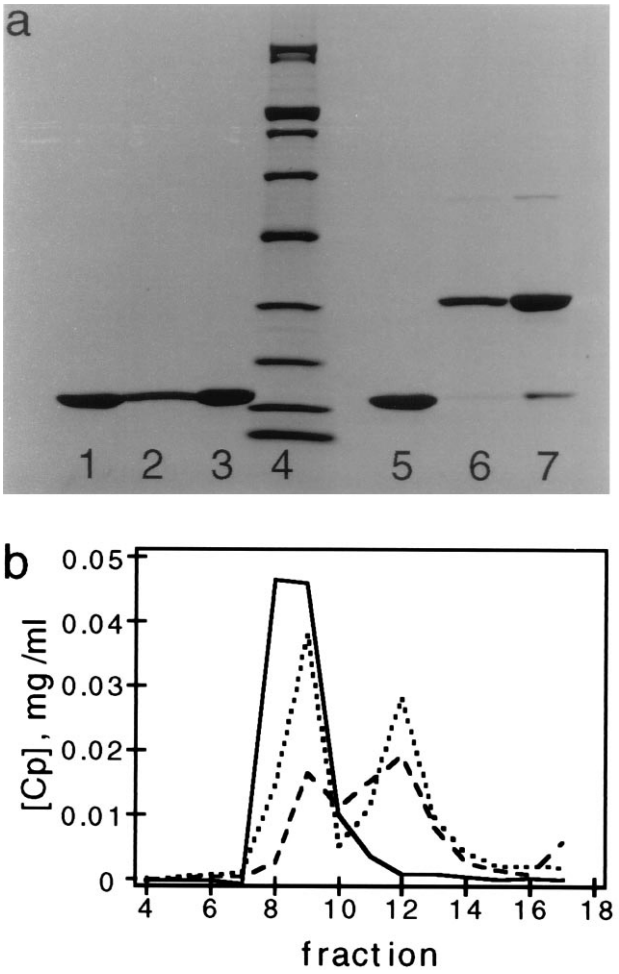

FIG. 2. Polymerized Cp*150 forms disulfide bonded dimers. (*a*) SDS/PAGE of Cp*150 from isolated capsids. Samples of Cp*150 were subjected to electrophoresis under reducing (lanes 1–3) and nonreducing conditions (lanes 5–7). Samples of Cp*150 polymerized at pH 9.5 (lanes 1 and 7) and at pH 7.5 (lanes 2 and 6). Lanes 3 and 5 are Cp*149, which has no cysteines. Molecular mass standards (200, 116, 97, 66, 45, 31, 21.5, 14.4, and 6.5 kDa) are in lane 4. (*b*) Size exclusion chromatography of capsid protein after exposure to 3.5 M urea. Samples are oxidized Cp*150 (solid line), Cp*150 with 130 mM DTT (dashed line), and Cp*149 (dotted line).

135,000 $M^{-1}\cdot cm^{-1}$, respectively. The extinction coefficient for free Au11 is 178,000 $M^{-1}\cdot cm^{-1}$ at 280 nm. The protein concentration in the fractions was determined using the same ε_{280} as unlabeled protein after subtracting the calculated Au11 absorbance.

Cryo-Electron Microscopy and Image Reconstruction. Capsid samples, adjusted to 1.5–2.5 mg/ml protein, were frozen in thin films of vitrified ice on holey carbon films (31, 32). Micrographs were recorded at a magnification of ×38,000 on a Philips CM20-FEG (Philips Electronic Instruments, Mahwah, NJ) fitted with a Gatan 626 cryoholder (Gatan, Pleasanton, CA) and operating at 120 keV. Pairs of micrographs at different defocus values were recorded using low-dose techniques. For gold-labeled capsids (Cp*150-Au11), the focal pair analyzed had first contrast transfer function zeros at spacings of (23.5 Å)$^{-1}$ and (18.0 Å)$^{-1}$, respectively; for RNA-filled Cp183 capsids, the zeros were at spacings of (25.5 Å)$^{-1}$ and (22.0 Å)$^{-1}$.

Micrographs were digitized on a Perkin–Elmer 1010MG scanner. Images of individual particles were extracted using a semiautomated procedure (33). Focal pairs of images were combined, and three-dimensional density maps were calcu-

lated as described (15, 34). Particle orientations were determined with the program EMPFT (35) using a 17-Å resolution density map of an empty T = 4 capsid as a starting model (15). Reconstructions of labeled Cp*150 (Cp*150-Au11) and Cp183 were calculated with data sets of 86 and 151 pairs of capsids, respectively. The resolution of the final maps were determined using the Fourier ring correlation method (33). The Cp147 structure shown for comparison with Cp*150-Au11 and Cp183 capsids was calculated from the 600-particle data set previously used for a 9-Å resolution reconstruction (17) but limited here to 17-Å resolution. The surface shaded maps were contoured to enclose a volume that was 100% of that expected for the protein mass.

RESULTS AND DISCUSSION

Capsids Assembled from Cp*150 and Cp*150-Au11. Purified Cp*149 and Cp*150 assemble into capsids under the same conditions as other Cp constructs (10, 15), with or without DTT. These capsids were indistinguishable by negative staining electron microscopy and sedimentation on sucrose gradients (data not shown). When reduced CP*150 capsids were stored without DTT for 2 days, >90% of the protein oxidized to form disulfide-bonded dimers (Fig. 2*a*). These bonds stabilize the quaternary structure of the capsid, as attested by the observation that oxidized Cp*150 capsids—unlike CP*149 capsids or reduced Cp*150 capsids—are resistant to dissociation by 3.5 M urea (Fig. 2*b*). Knowledge of the location of residue 150 (see below) indicates that this disulfide bond links two dimers (Fig. 1*b*) and is distinct from the intradimeric disulfide observed in Cp proteins with native cysteines (16, 25).

Generally, when Cp proteins are stored in a low ionic strength, high pH buffer they do not polymerize (10). However, when stored in this buffer without DTT, Cp*150 dimers assemble into capsids, as determined by negative stain electron microscopy and analytical ultracentrifugation. A high proportion of the protein in these capsids is disulfide-bonded (Fig. 2*a*). These data show that disulfide bond formation by Cp*150 can promote capsid assembly. Without disulfide formation, higher-order structures do not accumulate in storage buffer, i.e., the rate for dissociation is greater than the rate of association. Formation of these disulfide bonds stabilizes complexes against dissociation. Thus, under these conditions, Cp polymerization appears to involve an equilibrium between subunits, assembly intermediates, and capsids (36). We also note that, in capsids, the cysteine 150 residues from adjacent subunits must be close enough to one another to form a covalent bond, a distance of 4.6–7.4 Å between α carbons (37).

Undecagold Labeling of Cp*150. Assembled capsids of reduced Cp*150 did not react with Au11. Apparently, the C-terminal cysteine is inaccessible (see below). Accordingly, Cp*150 was labeled with Au11 at neutral pH and low ionic strength, where the sample is mainly free dimers (10). After reaction with Au11, >85% of the protein eluted in the void volume of a Sephacryl S-200 column, indicating that the protein was polymerized. Thus it appears that modification of Cp*150 with Au11 promotes polymerization. The Au11 has a single reactive maleimide group and cannot crosslink proteins. Though the molar ratio of Au11 to protein in the labeling reaction was varied from 1:1 to 1:5, a maximum of one label for every 5.3 Cp*150 monomers was observed, corresponding to about 45 Au11 clusters per T = 4 capsid.

Cryo-electron micrographs of labeled Cp*150 show a mixture of T = 3 and T = 4 capsids, typical for assembled Cp (Fig. 3*a*). Samples of labeled capsids were very similar to unlabeled ones, except that the proportion of T = 4 capsids was greater (91% vs. 75% T = 4, respectively). In micrographs taken close to focus (i.e. <1 μm defocus)—conditions that suppress the overall contrast of capsids but should enhance the visibility of the undecagold particles—individual clusters appeared to be marginally visible (cf. Fig. 3 *a* and *b*).

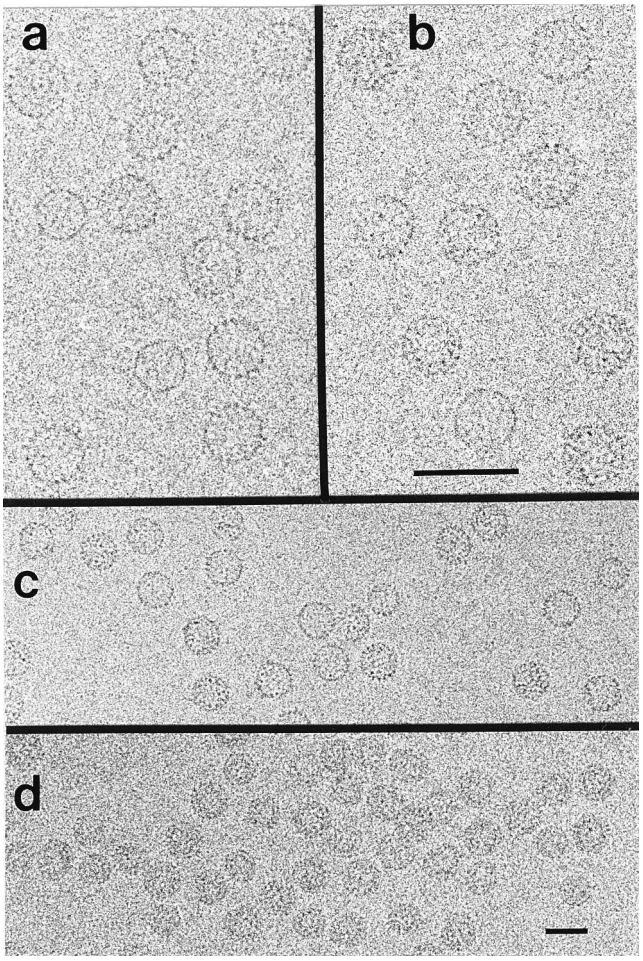

FIG. 3. Cryo-electron micrographs of unlabeled Cp149 capsids (*a*), Au11-labeled Cp*150 capsids (*b* and *c*), and RNA-filled Cp183 capsids (*d*). Views of unlabeled (*a*) and labeled (*b*) capsids are from closer to focus micrographs to facilitate comparison. Encapsidated RNA is visible as a dark ring within individual capsids in *d*. The first contrast zero function zeros are at spacings of 18.0, 18.0, 23.5, and 25.5 $(\text{Å})^{-1}$, respectively. (Bars = 50 nm.)

Localization of the C Terminus by Image Reconstruction. At a resolution of ≈20 Å, the outer surface of the Au11-labeled Cp*150 capsid is indistinguishable from those of unlabeled Cp147 and Cp183 capsids, (cf. Fig. 4 *Top*). In contrast, the interior view of the labeled Cp*150 capsid shows well defined densities attributable to the Au11 clusters directly beneath all fivefold and quasi-sixfold vertices (Fig. 4 *e* and *f*). These masses are ≈20 Å in diameter. Visualized in a central section through the map, the peak label-associated density is at least as strong as that of protein in the contiguous capsid (Fig. 4*f*): in comparison, only vestigial noise is seen in the corresponding region of the unlabeled capsid (cf. Fig. 4 *c* and *f*). Au11 is a ≈20 Å diameter molecule with an 8 Å diameter gold cluster at its center surrounded by an organic shell (30). Moreover, the density of gold is 19 g/cm^3 as compared with 1.3 for protein. As visualized, the label is larger and more diffuse than these numbers would suggest. However, the clusters visualized represent averages over the subunits around each fivefold or quasi-sixfold vertex and over all particles included in the reconstruction. Taking into account the measured level of labeling (≈45 clusters per T = 4 capsid) and the fact that the densities seen under the fivefold and quasi-sixfold axes are very similar, it follows that only one subunit per vertex is labeled. The apparent density of the label is broadened and further

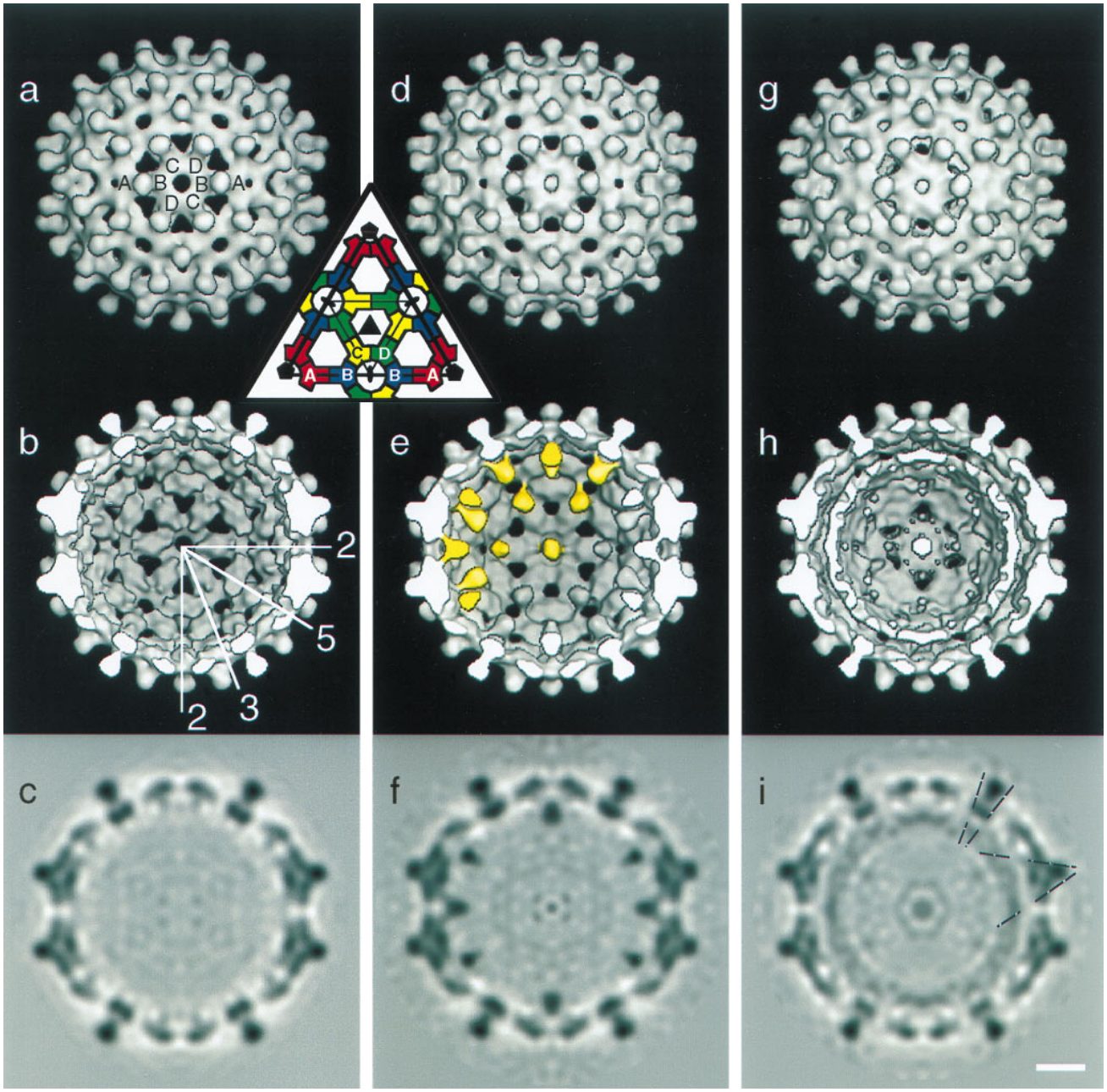

FIG. 4. Image reconstructions of T = 4 HBV capsids: empty Cp147 capsids at 17-Å resolution (*a–c*), C-terminally labeled Cp*150 capsids at 20 Å (*d–f*), and Cp183 capsids with RNA at 18.5 Å (*g–i*). Views are down an icosahedral twofold axis. (*Inset*) The arrangement of A, B, C, and D quasi-equivalent subunits on one icosahedral facet; the same labeling is used in *a*. The positions of some icosahedral symmetry axes, in the plane of the page, are shown in *b*. Some of the gold clusters in *e* are highlighted. Putative connections between capsid and RNA are identified by dashed lines in *i*. Surface shaded representations, external (*a*, *d*, and *g*) and cutaway (*b*, *e*, and *h*), are contoured so that 100% of the expected protein volume is enclosed. Gray scale maps (*c*, *f*, and *i*) are shown of the central section of each reconstruction. (Bar = 50 Å.)

attenuated by the limited resolution (20 Å) of the reconstruction.

As noted above, the 42 Au11 clusters observed in the reconstruction (T = 4 icosahedra have 12 fivefold and 30 quasi-sixfold vertices) agrees well with the 45 Au11 clusters per T = 4 capsid determined spectroscopically. Though less than 20% of the protein is labeled, there is one Au11 per vertex. We infer that labeling of one subunit enhances its propensity for assembly, and its label sterically excludes other labeled subunits from assembling into the ring forming around a vertex. The lack of Au11 reactivity for assembled Cp*150 particles most likely reflects the inaccessibility of the C terminus due to the large diameter of the label. The largest of the holes fenestrating the capsids is 24 Å by 18 Å (15) implying that it would be difficult for the 20 Å diameter Au11 molecule to diffuse through.

The Au11 labels are clearly bound to the ends of the Cp*150 dimers (Fig. 4*e*). A connection between Cp*150 and Au11 is observed for B subunits. This connection appears to emanate from the small protrusion first identified in reconstructions at 17 Å resolution (15). Our visualization of the arm of connecting density indicates that at least some the C-terminal residues of Cp*150 are icosahedrally ordered (Fig. 4*c*). The corresponding connections from Au11 to A, C, and D subunits are much weaker. The lack of connectivity to C and D subunits is due to a real preference for Au11-labeled subunits at the B position. This preference reflects the quasi-equivalence around HBV T = 4 quasi-sixfold vertices resulting in distinct

differences between the environment of the B subunit and that of the C and D subunits (15). The weak connectivity to C and D subunits is unlike the situation observed for A subunits where the density has been decreased by fivefold averaging (see below).

Why is the resolution of Cp*150-Au11 limited to 20 Å? Part of the answer is that because there is only one Au11 per vertex, labeled capsids are no longer strictly icosahedral. The reconstruction algorithm that we use assumes and imposes icosahedral symmetry (15, 17, 34). For example, only one of the five A subunits actually connects to a given Au11 cluster near a fivefold vertex; after averaging, the connecting density will be 20% of the density of nearby protein. In calculating a reconstruction, the handedness of each particle must be determined (38). For a T = 4 capsid the differences between the left-handed and right-handed enantiomorphs are necessarily subtle (12) and distinguishing between them is further hampered by the asymmetric distribution of Au11 clusters. The likely inclusion of some particles with the wrong hand would limit the resolution of the reconstruction. Nevertheless, our conclusions as to cluster location are unaffected.

Structure of RNA-Containing Capsids. In micrographs of Cp183 capsids, the encapsidated RNA is evident as a smaller ring within many of the particles (Fig. 3*d*). The observed double ring implies that encapsidated RNA forms an inner shell. If capsids were packed full of RNA, one would expect to see them as dense disks.

A density map of RNA containing particles was calculated to 18.5-Å resolution. Like the Au11-labeled particles, the external features of the T = 4 Cp183 capsid are as expected for a reconstruction at this resolution (Fig. 4*g*). In thin sections (e.g., Fig. 4*i*), an RNA shell is seen to be ≈30 Å thick and its average density is lower than that of the protein shell, implying that the distribution of the RNA is somewhat variable from particle to particle and between symmetry-related sites on a given particle. The highest density of RNA is located under fivefold and quasi-sixfold vertices and beneath the sides of the AB and CD spikes. There are also indications of low-density connections between protein and RNA (Fig. 4*i*, dashed lines).

CONCLUSIONS AND IMPLICATIONS

The Site Occupied by the C Terminus of the Assembly Domain. Residue 150 has been localized by image reconstruction of gold-labeled capsids. The appearance of the undecagold cluster in our map (Fig. 4 *e* and *f*) is broadened on account of the limited resolution (20 Å). Nevertheless, the position of its center of mass is defined to within a few Å units in three dimensions. For the B subunits, the connecting linker is sufficiently well ordered to be visualized directly (Fig. 4 *e* and *f*). We conclude that the labeled residue, cysteine 150, lies 10–15 Å along this linker from the cluster center of mass and, consequently, its location is also defined to within a few Å units (Fig. 5). We infer that the corresponding residues on A, C, and D subunits are similarly positioned. Thus, residue 150 has been localized to small protrusions at either end of the Cp dimer (15). This feature, as defined in our 9-Å resolution map (17), is shown and the C-terminal sites marked (Fig. 5).

Recently, image reconstructions of HBV Cp have been determined to unprecedented resolution (17, 18). The analysis of Böttcher *et al.* (18) tentatively assigned the C terminus to positions near the ends of the dimer and near the outer surface (figure 4 of ref. 18). A similar assignment was made in the model of Bringas (39). Our experimental localization of the C terminus confirms that it resides near the end of the arm, but on its underside.

Control of Assembly. Our earlier work revealed that residues 140–149 serve as a morphogenetic switch, longer proteins favoring T = 4 and shorter ones, T = 3 geometry (15). This behavior now can be explained in light of the position that the switch occupies. The C-termini of adjacent subunits crowd one another around fivefold and quasi-sixfold vertices. Viewed from the inside, the crowding is more acute around the fivefold vertices, suggesting that the more open geometry of the T = 4 capsid would be preferred for proteins with a bulkier C terminus (see also ref. 18). This argument is consistent with the greater proportion of T = 4 capsids with gold-labeled versus unlabeled capsids (this work).

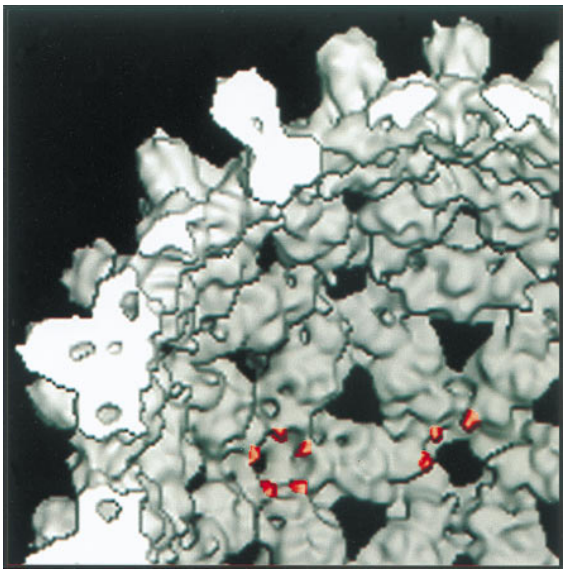

FIG. 5. The C terminus of the assembly domain in a 9-Å resolution reconstruction. Some C termini are highlighted. The map is contoured at 100% of expected protein density, a lower contour level than used to trace the protein chain (17).

Other observations imply that the C termini also may influence assembly in more subtle ways. For instance, binding Au11 to Cp*150 induces assembly, though Au11 cannot crosslink subunits nor, because of its organic shell, coordinate C-terminal cysteines. Binding Au11 may cause small changes in the molecular surfaces, near the C termini, that dock together when dimers polymerize and, in this way, stimulate the assembly process. If the binding of the reverse transcriptase-RNA complex to the Cp (5) were to transmit a similar effect, this could enhance the cooperativity of nucleocapsid assembly.

Packaging of RNA and its Interactions with the Capsid. The protamine domain (residues 150–183) is required for packaging RNA (8, 21) and should greatly influence the organization of encapsidated RNA (cf., ref. 40). This expectation is confirmed in our density map (Fig. 4*i*), in which RNA-associated density is greatest under fivefold and quasi-sixfold vertices, and presumably represents the binding of RNA to protamine domains, secured to the contiguous shell via the C termini of the assembly domains. Otherwise, the RNA is not closely associated with the capsid inner surface except at sites underlying both AB and CD dimer interfaces (Fig. 4*i*). This latter contact is consistent with that observed by Kenney *et al.* (14) in their reconstruction of duck hepatitis virus capsids. This site on the assembly domain with an apparent affinity for nucleic acid may coincide with that previously inferred by biochemical methods (41).

Compared with many other viruses, the packing density of RNA in HBV capsids is decidedly loose (cf., ref. 31), the RNA being confined to an inner shell of density. We make the caveat that the organization of the native RNA pregenome may differ from that of the bacterially derived RNA in the virus-like particles that we have studied. However, we note that HBV capsids isolated from human liver also appear to be hollow in negative stain electron micrographs (42). A similar configu-

ration of encapsidated nucleic acid has been documented in image reconstructions of cauliflower mosaic virus (43), a pararetrovirus, that—like HBV—has a DNA genome with an RNA replicative intermediate (44).

A General Method for Mapping Polypeptide Chains at Moderate Resolution. As reconstructions calculated from cryo-electron micrographs continue to improve in resolution to 10 Å and beyond (17, 18, 45), the need arises to be able to determine the overall fold of polypeptide chain(s) within them. In general, higher resolution is required for unambiguous chain tracing by the traditional x-ray crystallographic method of recognizing bulky side-chains and correlating this information with the path followed by the backbone. We suggest that the approach followed in this study—of incorporating cysteines at defined sites by site-specific mutagenesis, labeling them with gold clusters, and determining their locations by cryo-electron microscopy—may prove to be useful for this purpose.

We acknowledge the superb technical contributions of Mr. Joshua Kaufman and Mr. Ira Palmer in the preparation of genetic constructs and purification of protein. We thank Drs. Frank Booy, Benes Trus, David Belnap, and Jim Hainfeld for helpful discussions and Tim Baker for software. A.Z. was supported by a National Research Council-National Institute of Arthritis and Musculoskeletal and Skin Diseases fellowship.

1. Hollinger, F. B. (1996) in *Fields Virology*, eds. Fields, B. N., Knipe, D. M., Howley, P. M., Chanock, R. M., Melnick, J. L., Monath, T. P., Roizman, B. & Straus, S. E. (Lippincott, Philadelphia), 3rd Ed, pp. 2738–2808.
2. Nassal, M. & Schaller, H. (1993) *Trends Microbiol.* **1,** 221–228.
3. Bartenschlager, R., Junker-Niepmann, M. & Schaller, H. (1990) *J. Virol.* **64,** 5324–5332.
4. Summers, J. & Mason, W. S. (1982) *Cell* **29,** 403–415.
5. Bartenschlager, R. & Schaller, H. (1992) *EMBO J.* **11,** 3413–3420.
6. Gerelsaikhan, T., Tavis, J. E. & Bruss, V. (1996) *J. Virol.* **70,** 4269–4274.
7. Petit, M. A. & Pilot, J. (1985) *J. Virol.* **53,** 543–551.
8. Nassal, M. (1992) *J. Virol.* **66,** 4107–4116.
9. Zhou, S. & Standring, D. N. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 10046–10050.
10. Wingfield, P. T., Stahl, S. J., Williams, R. W. & Steven, A. C. (1995) *Biochemistry* **34,** 4919–4932.
11. Cohen, B. J. & Richmond, J. E. (1982) *Nature (London)* **296,** 677–679.
12. Caspar, D. L. D. & Klug, A. (1962) *Cold Spring Harbor Symp. Quant. Biol*, 1–24.
13. Crowther, R. A., Kiselev, N. A., Böttcher, B., Berriman, J. A., Borisova, G. P., Ose, V. & Pumpens, P. (1994) *Cell* **77,** 943–950.
14. Kenney, J. M., von Bonsdorff, C. H., Nassal, M. & Fuller, S. D. (1995) *Structure* **3,** 1009–1019.
15. Zlotnick, A., Cheng, N., Conway, J. F., Booy, F. P., Steven, A. C., Stahl, S. J. & Wingfield, P. T. (1996) *Biochemistry* **35,** 7412–7421.
16. Zheng, J., Schodel, F. & Peterson, D. L. (1992) *J. Biol. Chem.* **267,** 9422–9429.
17. Conway, J. F., Cheng, N., Zlotnick, A., Wingfield, P. T., Stahl, S. J. & Steven, A. C. (1997) *Nature (London)* **386,** 91–94.
18. Böttcher, B., Wynne, S. A. & Crowther, R. A. (1997) *Nature (London)* **386,** 88–91.
19. Beames, B. & Lanford, R. E. (1993) *Virology* **194,** 597–607.
20. Birnbaum, F. & Nassal, M. (1990) *J. Virol.* **64,** 3319–3330.
21. Pasek, M., Goto, T., Gilbert, W., Zink, B., Schaller, H., MacKay, P., Leadbetter, G. & Murray, K. (1979) *Nature (London)* **282,** 575–579.
22. Milligan, R. A., Whittaker, M. & Safer, D. (1990) *Nature (London)* **348,** 217–221.
23. Crum, J., Gruys, K. J. & Frey, T. G. (1994) *Biochemistry* **33,** 13719–13726.
24. Hainfeld, J. F. (1987) *Science* **236,** 450–453.
25. Nassal, M., Rieger, A. & Steinau, O. (1992) *J. Mol. Biol.* **225,** 1013–1025.
26. Nassal, M. (1992) *Virology* **190,** 499–505.
27. Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. (1990) *Methods Enzymol.* **185,** 60–89.
28. Higuchi, R., Krummel, B. & Saiki, R. K. (1988) *Nucleic Acids Res.* **16,** 7351–7367.
29. Stahl, S., MacKay, P., Magazin, M., Bruce, S. A. & Murray, K. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 1606–1610.
30. Bartlett, P. A., Bauer, B. & Singer, S. J. (1978) *J. Am. Chem. Soc.* **100,** 5085–5092.
31. Booy, F. P., Newcomb, W. W., Trus, B. L., Brown, J. C., Baker, T. S. & Steven, A. C. (1991) *Cell* **64,** 1007–1015.
32. Booy, F. P. (1993) in *Viral Fusion Mechanisms*, ed. Bentz, J. (CRC, Boca Raton, FL), pp. 21–54.
33. Conway, J. F., Trus, B. L., Booy, F. P., Newcomb, W. W., Brown, J. C. & Steven, A. C. (1993) *J. Struct. Biol.* **111,** 222–233.
34. Crowther, R. A. (1971) *Philos. Trans. R. Soc. London Ser. B* **261,** 221–230.
35. Baker, T. S. & Cheng, R. H. (1996) *J. Struct. Biol.* **116,** 120–130.
36. Zlotnick, A. (1994) *J. Mol. Biol.* **241,** 59–67.
37. Thornton, J. M. (1981) *J. Mol. Biol* **151,** 261–287.
38. Baker, T. S., Drak, J. & Bina, M. (1989) *Biophys. J.* **55,** 243–253.
39. Bringas, R. (1997) *J. Struct. Biol.* **118,** 189–196.
40. Cheng, R. H., Reddy, V. S., Olson, N. H., Fisher, A., J., Baker, T. S. & Johnson, J. E. (1994) *Structure* **2,** 271–282.
41. Matsuda, K., Satoh, S. & Ohori, H. (1988) *J. Virol.* **62,** 3517–3521.
42. Onodera, S., Ohori, H., Yamaki, M. & Ishid, N. (1982) *J. Med. Virol.* **10,** 147–155.
43. Cheng, R. H., Olson, N. H. & Baker, T. S. (1992) *Virology* **186,** 655–668.
44. Mesnard, J. M. & Carriere, C. (1995) *Virology* **213,** 1–6.
45. Trus, B. L., Roden, R. B. S., Greenstone, H. L., Vrhel, M., Schiller, J. T. & Booy, F. P. (1997) *Nat. Struct. Biol.* **4,** 411–418.