Articles

Hepatitis Core Antigen Produced in *Escherichia coli*: Subunit Composition, Conformational Analysis, and *In Vitro* Capsid Assembly†

Paul T. Wingfield,*,‡ Stephen J. Stahl,‡ Robert W. Williams,§ and Alasdair C. Steven||

Protein Expression Laboratory, Office of the Director, National Institutes of Health, Building 6B, Room 1B130, Bethesda, Maryland 20892-2775, Department of Biochemistry, Uniform Services University of the Health Sciences, 4301 Jones Bridge Road, Bethesda, Maryland 20814-4799, and Laboratory of Structural Biology Research, National Institutes of Arthritis, Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, Maryland 20892-2755

Received August 3, 1994; Revised Manuscript Received December 6, 1994⊗

ABSTRACT: The production and biochemical and physicochemical analysis are described of recombinant-produced hepatitis B virus core antigen (HBcAg capsid) and the corresponding particle produced by a deletion mutant missing the C-terminal 39 residues (HBeAg). Conditions for producing HBeAg from HBcAg capsids by *in vitro* proteolysis are also described. The morphology and masses of these capsids were determined by scanning transmission electron microscopy. Both HBcAg and HBeAg capsids comprise two size classes that correspond to icosahedral lattices with triangulation numbers (T) of 3 and 4, containing 180 and 240 subunits per capsid, respectively. This dimorphism was confirmed by sedimentation equilibrium and sedimentation velocity measurements on a Beckman Optima XL-A analytical ultracentrifuge. More than 60% of HBcAg capsids were $T = 4$, whereas only 15−20% of HBeAg capsids were of this size class: the remainder, in each case, were $T = 3$. Circular dichroism and Raman spectroscopy were used to determine the overall secondary structures of HBcAg and HBeAg capsids. Both have high α-helical contents, implying that this capsid protein does not conform to the canonical β-barrel motif seen for all plant and animal icosahedral viral capsids solved to date. We suggest that the C-terminal domain of HBcAg has a random coil conformation. *In vitro* dissociation of HBeAg capsids under relatively mild conditions yielded stable dimers. The reassociation of HBeAg dimers into capsids appears to be driven by hydrophobic processes at neutral pH. Capsid assembly is accompanied by little change in subunit conformation as judged by circular dichroism and fluorescence spectroscopy. The thermal stability of HBcAg capsids was compared calorimetrically with that of *in vitro* assembled HBeAg capsids. Both have melting temperatures >90 °C, implying that the C-terminal region makes little difference to the thermal stability of HBcAg: nevertheless, we discuss its possible role in facilitating disassembly and the release of viral nucleic acid.

Hepatitis B virus (HBV)[1] is a member of the hepadnavirus family which infects only humans and higher primates. The intact virus (Dane particle) contains an outer lipoprotein envelope and an inner nucleocapsid which is composed of a double stranded DNA (3.2 kb), a reverse transcriptase, and a capsid coat protein. HBV replication is related to that of retroviruses with some important differences, for example, extracellular virions contain DNA rather than RNA and their integration into the host genome is not an obligatory step in the life cycle (review: Nassal & Schaller, 1993).

The nucleocapsid protein or core antigen (HBcAg) of HBV is a 22 kDa protein that can be expressed in *Escherichia coli* and other host systems where it assembles into capsids. Recombinant-produced capsids have similar morphological and immunological properties to the authentic antigen isolated from infected human hepatocytes (review: Murray, 1987). HBcAg has a C-terminal region rich in arginine residues. This histone-like region was originally postulated to be involved in binding the viral nucleic acid (Pasek *et al.*, 1980) and may also have a role in nuclear transport (Yeh *et al.*, 1990).

Related to the core antigen is a 17 kDa protein found in the sera of patients with active HBV infection. This protein is a truncated version of HBcAg missing 34 residues from the C-terminus (Takahashi *et al.*, 1983) and is serologically defined as the "e-antigen" (HBeAg). It has been shown that serum HBeAg is derived from a second core translation product, namely, precore protein, which is HBcAg plus a 29 amino acid N-terminal extension (Strandring *et al.*, 1988, and references therein). C-Terminal deletion mutants of

† This work was supported in part by the AIDS Targeted Antiviral Program of the Office of the Director of the National Institutes of Health (P.T.W., S.J.S., and A.C.S.) and by USUHS Grant R071BV (R.W.W.).

‡ Protein Expression Laboratory, National Institutes of Health.

§ Uniform Services University of the Health Sciences.

|| National Institutes of Arthritis, Musculoskeletal and Skin Diseases, National Institutes of Health.

⊗ Abstract published in *Advance ACS Abstracts,* March 1, 1995.

[1] Abbreviations: CD, circular dichroism; DSC, differential scanning calorimetry; DTNB, 5,5′-dithiobis(2-nitrobenzoic acid); Gdn-HCl, guanidine hydrochloride; HBV, hepatitis B virus; *N*-AcTrpNH$_2$, *N*-acetyltryptophanamide; PMSF, phenylmethanesulfonyl fluoride; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; STEM, scanning transmission electron microscopy; UV, ultraviolet; T, triangulation number of a viral capsid (equal to the number of identical proteins within each of the 60 icosahedral asymmetric units comprising the capsid).

0006-2960/95/0434-4919$09.00/0

HBcAg that mimic the sequence of HBeAg have been expressed in *E. coli* by several groups. The synthesized protein assembles into capsid-like structures, indicating that the C-terminal region is not critical for assembly (Milich *et al.*, 1988; Gallina *et al.*, 1989).

Despite the availability of large quantities of recombinant-produced HBV nucleocapsid proteins that appear biochemically and immunologically similar to their native counterparts, little information is available on the protein structure or on how it assembles into capsids. Some progress on this topic has been made using an *in vivo* model system (Zhou & Strandring, 1992), but our knowledge is still relatively scanty. On the basis of negatively stained electron micrographs (see, e.g., Onodera *et al.*, 1982), most investigators have assumed icosahedral symmetry with a triangulation number (Caspar & Klug, 1962) $T = 3$ with 180 subunits, for the organization of native or recombinant HBcAg capsids. Very recently, Crowther *et al.* (1994) have used cryo-electron microscopy and image reconstruction to identify two size classes among recombinant HBcAg capsids expressed in *E. coli*. These classes correspond to $T = 4$ and $T = 3$, respectively. We also have addressed this question in detail and reached similar conclusions using two complementary mass determination techniques, namely, scanning transmission electron microscopy (STEM) and analytical ultracentrifugation. With these methods, we show that both HBcAg capsids formed intracellularly and HBeAg capsids assembled *in vitro* from purified dimeric protein exhibit the same $T = 3$ and $T = 4$ dimorphism, albeit with different proportions of the two capsid species.

Although the availability of natural HBcAg capsids for investigation, usually isolated from infected human liver, is severely limited, recombinant-produced HBV protein offers an abundant system in which to study the *in vitro* assembly of an icosahedral capsid. In order to establish a basic framework for detailed studies of this kind, we have undertaken to generate the necessary physiochemical and structural information on HBV capsid proteins. We describe the production of recombinant HBcAg and HBeAg in *E. coli*. The conversion of HBcAg capsids into HBeAg capsids by proteolytic processing is also described, as are *in vitro* dissociation and reassembly of HBeAg. Far-UV CD and Raman spectroscopy were used to determine the overall secondary structure of intact capsids and dissociated capsid protein. A variety of spectroscopic, hydrodynamic, and thermodynamic techniques were applied to characterize the proteins and their assembly properties.

MATERIALS AND METHODS

Expression of HBcAg and Truncated HBcAg Polypeptides in E. coli. Plasmid pR1-11, which directs the synthesis of HBcAg, was described previously (Stahl *et al.*, 1982). A truncated version of HBcAg missing 39 residues from the C-terminus was expressed from a derivative of ptacHpaII (Stahl & Murray, 1989). In addition to the first 8 amino acids of β-galactosidase and amino acid residues 3−144 of HBcAg, the polypeptide expressed by ptacHpaII contains 10 amino acids at the carboxy-terminus which are due to readthrough into the pBR322 region of the plasmid. The last 8 of these amino acids were eliminated by cutting ptacHpaII with *Hin*dIII and filling out the ends using the Klenow fragment of DNA polymerase. This introduces a termination codon in frame with the HBcAg coding sequence such that only 2 extraneous amino acid residues are present at the carboxy-terminus of the truncated HBcAg polypeptide (ptacHBc144).

Purification of HBcAg Capsids. Buffers used in purification were as follows: buffer A, 200 mM sodium phosphate, pH 7.5, containing 5 mM benzamidine hydrochloride and 5 mM EDTA; buffer B, 50 mM Tris-HCl, pH 7.8, containing 5% (w/v) sucrose, 5 mM EDTA, and 1 mM sodium azide. All operations were at 4 °C. HBcAg protein in column fractions was located by SDS−PAGE.

Cells (150 g wet mass) were resuspended with 350 mL of buffer A and broken with a French pressure cell (SLM Aminco). The suspension was briefly sonicated and centrifuged at 8000*g* for 1 h and the supernatant recentrifuged at 125000*g* for 2 h. The pelleted HBcAg was resuspended with 35 mL of buffer B and applied to a column (90 × 5.0-cm diameter) of Sepharose CL-4B (Pharmacia) equilibrated with buffer B. The peak containing HBcAg was pooled (430 mL) and applied to a column (10 × 5.0-cm diameter) of Blue Sepharose CL-6B (Pharmacia) equilibrated with buffer B. The HBcAg protein was collected from the flow-through (385 mL) and concentrated 8-fold by ultrafiltration using Diaflo XM 100 membranes (Amicon). Concentrated protein was applied (5−10 mL/tube) to discontinuous sucrose density gradients formed in polycarbonate tubes (maximum volume 70 mL) by layering 10 mL of 60% sucrose (w/v), 20 mL of 50% sucrose (w/v), and 20 mL of 40% (w/v) sucrose. Sucrose solutions were buffered with 50 mM Tris-HCl, pH 7.8. The tubes were centrifuged in a Beckman rotor type 45 Ti at 111000*g* for 7 h at 4 °C and then fractionated from the bottom using a peristaltic pump. The HBcAg was located just within the 50% (w/v) sucrose zone, well separated from a green colored proteinaceous zone (NADPH-sulfite reductase) located near the top of the 40% (w/v) sucrose zone. HBcAg (110 mL) was dialyzed against buffer B, concentrated by ultrafiltration to about 2−3 mg/mL, and stored at −80 °C. About 90 mg of purified protein was recovered.

Preparation of HBeAg. Purification was the same as described for HBcAg up to the sucrose density gradient stage which was omitted. The HBeAg protein was pelleted by centrifugation (125000*g* for 2 h) and resuspended with 100 mM Na_2CO_3-$NaHCO_3$, pH 9.5, containing 5.0 M urea. The protein was chromatographed on a column of Sephacryl S-200 (Pharmacia) equilibrated with the resuspension buffer minus the urea. The low molecular weight protein peak was collected ($K_{av} = 0.30$) and either stored at −80 °C or dialyzed overnight against 50 mM Tris-HCl, pH 7.0, containing 0.15 M NaCl. The dialysate was concentrated using a Diaflo XM 100 membrane and then chromatographed on a Sepharose CL-4B column equilibrated with dialysis buffer.

Preparation of Tryptic Digested HBcAg. HBcAg (40 mL at 2.2 mg/mL) in 50 mM Tris-HCl, pH 8.0, containing 1 mM $CaCl_2$, 5% (w/v) sucrose and 4.5 M urea was incubated with 2.5% (w/w) bovine trypsin (Worthington) for 1 h at 23 °C. Bovine RNase A (Sigma) was added (50 μg/mL), followed by incubation for a further 30 min at 23 °C. PMSF was added to 1 mM, the pH was adjusted to pH 9.5 with NaOH, and the solution (40 mL) applied to a Sephacryl S-200 column (5-cm diameter × 95 cm) equilibrated with 100 mM Na_2CO_3-$NaHCO_3$, pH 9.5. The digested HBcAg was pooled on the basis of elution position and SDS−PAGE

analysis. About 66 mg of protein was obtained (typically >75% recovery).

Isokinetic Sucrose Density Centrifugation. Isokinetic sucrose density gradients of 15.0−31.5% (w/w) in 100 mM Tris-HCl, pH 7.8, 0.15 M NaCl, and 1.0 mM sodium azide were prepared in Beckman type SW 41 polyallomer tubes according to McCarty *et al.* (1974). Samples (50−200 μL) were layered on the gradients and centrifuged at 198000*g* for 2.5 h at 4 °C. The gradients were drained by displacement upward at about 1 mL/min using 60% (w/w) sucrose. The absorbance at 280 nm was monitored, and 0.33−0.34 mL fractions were collected. Sedimentation coefficients were estimated as described by McCarty *et al.* (1974).

CsCl Gradient Centrifugation. HBcAg solutions (50−100 μL at 5 mg/mL) were mixed with 15 g of 52% (w/v) CsCl in 50 mM Tris-HCl, pH 7.8, and the mixtures were transferred to Beckman type SW 41 polyallomer tubes. Centrifugation was at 110000*g* for 64 h at 4 °C. The tube contents were pumped out downward through a UV monitor and fractionated. Fractions of 250 μL were collected. The refractive index of each fraction was measured using an Abbe-3L refractometer (Milton Roy).

Protein Determination. Protein concentrations were estimated by measuring the absorption in a 1 cm path length cell using a double-beam, diode array Hewlett-Packard 8450A UV/vis spectrophotometer. For HBcAg, a molar extinction coefficient at 260 nm of 125 mM^{-1} cm^{-1} was used. This value was calculated using the summed protein and nucleic acid molar extinction coefficients at 260 nm, assuming the latter is 20% by weight. This value was confirmed both by model building using amino acids and yeast tRNA and by quantitative amino acid analysis. The HBeAg concentration was measured at 280 nm using a calculated molar absorbance coefficient of 28.1 mM^{-1} cm^{-1}. The molar absorption coefficient was calculated according to the method of Wetlaufer (1962). Light scattering in polymeric samples was corrected for as described in the legend to Figure 1.

Chemical Analysis. Amino acid analysis, N-terminal sequence analysis, and sulfhydryl determinations were carried out as previously described (Wingfield *et al.*, 1986). The C-terminal analysis of HBeAg was performed by mass spectrometry. The principle of the analysis by combined gas−liquid chromatography/mass spectroscopy (GLC-MS), and full details of acylation, permethylation, reagent and sample preparation, and machine operating parameters have been described previously (Rose *et al.*, 1983).

Analytical Ultracentrifugation. Analytical ultracentrifugation was carried out using a Beckman Optima XL-A analytical ultracentrifuge with an An-60 Ti rotor and standard double-sector centerpiece cells. For equilibrium measurements on capsid preparations, centrifugations (43 h at 20 °C) were at 1800 rpm (HBcAg) and 2800 or 2000 rpm (HBeAg). Sedimentation velocity measurements were at 12 000 rpm (HBcAg) and 16 000 rpm (HBeAg) for 2−3 h at 20 °C with data collection every 10 min. Samples were in 50 mM Tris-HCl, pH 7.0, containing 0.15 M NaCl (ϱ = 1.00494 g/mL). Equilibrium measurements on the low molecular weight HBeAg protein were made in 100 mM Na_2CO_3-$NaHCO_3$, pH 9.5 (ϱ = 1.0085 g/mL). Centrifugation at 15 000 rpm was for 15 h at 20 °C. Sedimentation velocity measurements on the low molecular weight HBeAg were made in pH 9.5 buffer at 45 000 rpm at 20 °C. Data were analyzed using both the standard Optima XL-A data analysis software (version 3.0 for DOS) and the Beckman-Origin software (version 2.0 for Windows). Sedimentation velocity data were also analyzed by the method of Stafford (1992). Protein partial specific volumes were calculated from amino acid compositions (Cohn & Edsall, 1943).

Fluorescence Measurements. A Perkin Elmer LS-5 luminescence spectrometer interfaced to a Perkin Elmer 3700 data station was used. The excitation wavelength was 295 nm, and both the excitation and emission bandwidths were 2.5 nm. Spectra were recorded at 22 °C. Further details are found in the legend to Figure 7 and footnotes to Table 4.

Circular Dichroism. Spectra were recorded using a Jasco J-720 spectropolarimeter. Measurements in the near-ultraviolet region (240−340 nm) were made using protein solutions in 50 mM sodium phosphate, pH 7.0, or in 100 mM Na_2CO_3-$NaHCO_3$, pH 9.5, and a 1 cm path length cell. Measurements in the far-ultraviolet region (180−260 nm) were made in the same buffers using a 0.01 cm path length cell. For both near- and far-UV CD, protein solutions having 1.0−1.8 OD at 260 nm (HBcAg) or 280 nm (HBeAg) were used with a bandwidth of 1 nm. Solutions were filtered with a Millex-GV 0.22-μm filter (Millipore) and degassed prior to use. Secondary structures were estimated using the methods of Perczel *et al.* (1992), Sreerama and Woody (1993), and Chang *et al.* (1976).

Vibrational Spectroscopy. The Raman instrumentation and the calibration method have been described elsewhere (Weaver & Williams, 1990). The spectrometer was calibrated daily using the 546.074 nm (18312.3 cm^{-1}) Hg vapor line. The laser was operated at 514.5 nm with 100 mW of light power focused at the sample to a beam with a diameter of about 0.2 mm. The sample, about 5 μL of pelleted HBeAg particles, was contained in a melting point capillary tube thermostated at 20 °C.

Differential Scanning Calorimetry. Differential scanning calorimetry (DSC) was carried out with a Microcal MC-2 instrument (MicroCal Inc., Amherst, MA). The volume of the calorimeter cell was 1.25 mL. A heating rate of 90 °C/h was used. Samples (1.5 mL at 0.5−1.0 mg/mL) were degassed and equilibrated at 15 °C for 1 h before scanning. Data were collected between 20 and 100 °C and analyzed using Origin version 2.9 for Windows (MicroCal Inc.). The thermogram for HBeAg at pH 9.5 was fitted assuming an independent non-two-state transition (Model 2: non-2-state zero ΔC_p, of the Origin software).

Scanning Transmission Electron Microscopy. Purified HBeAg and HBcAg capsids were adsorbed to thin carbon films according to the "wet film" method (Wall *et al.*, 1985) and freeze-dried at a constant sublimation rate, essentially as described by Wall (1979). The specimens were ~2 mg/mL protein in 50 mM Tris-HCl and 150 mM NaCl, pH 7.0. After adsorption, the grids were washed ten times with 50 mM ammonium acetate, pH 7.0. Tobacco mosaic virus was used as internal mass standard. Specimens were maintained at −150 °C during observation on the STEM at the Brookhaven National Laboratory (Wall & Hainfield, 1986). Dark-field micrographs were recorded in digital form, each of 512 × 512 pixels, with a 2 nm raster step between pixels. Mass measurements were carried out by means of the PIC image analysis program (Trus *et al.*, 1992), running on a VAX 4000-500 computer (Digital Equipment Corp., Maynard, MA) and making use of a Gould IP9000 image

processor (Vicom, Fremont, CA). Radial reconstructions were performed as described by Steven *et al.* (1984). The analysis was performed on sets of the best preserved particles by visual criteria, i.e., symmetrical particles surrounded by clean background. Their centers were located with subpixel precision by cross-correlation methods. For each particle, a one-dimensional trace was calculated, representing a scan along its azimuthally averaged mass projection, with its origin at the particle center. At this step, the sampling rate was increased from 2 nm per pixel to 0.667 nm per pixel without interpolation-induced loss of resolution, by Vernier binning (Steven *et al.*, 1984). The projection traces were averaged over 10 particles (HBeAg, small); 9 particles (HBeAg, large); 14 particles (HBcAg, small); and 9 particles (HBcAg, large). The background contributions from the carbon support film were subtracted, and the radial density profiles were calculated. Points close to the origin are susceptible to anomalous fluctuations from even low levels of residual noise (Steven *et al.*, 1984; Duda *et al.*, 1985); accordingly, calculated densities for radii of less than 2 nm were disregarded.

RESULTS

Purification and Characterization of Particulate HBcAg and HBeAg. HBcAg expressed in *E. coli* accumulates in a particulate state referred to herein as HBcAg capsids. HBcAg capsids were isolated from *E. coli* as nucleoprotein complexes (absorbance ratio 280/260 ~ 0.63; Figure 1A). The nucleic acid associated with the HBcAg capsids could be partially digested with RNase A; gel analysis of a complete phenolic extract of nucleic acids indicated predominantly low molecular weight RNA, probably *E. coli* tRNA. The capsids formed a single band on CsCl gradients with a density of 1.38−1.39 g/mL. This density corresponds to about 20% nucleic acid, a value similar to that estimated by UV/vis spectroscopy. The particles also gave one band on SDS−PAGE with an estimated molecular weight of about 22 000 (the M_r derived from the DNA coding sequence is 21 970; Pasek *et al.*, 1979).

It has been previously shown that removal of the basic C-terminal 40 residues of recombinant HBcAg does not appear to influence capsid assembly (Milich *et al.*, 1988; Gallina *et al.*, 1989). Both the HBcAg deletion mutant and the tryptic processed HBcAg (see below) will be referred to as HBeAg (Milich *et et.*, 1988). HBeAg capsids did not contain nucleic acid as judged by their UV absorption (Figure 1B); however, they contained entrapped protein which was manifest as background staining when analyzed by SDS−PAGE. This background was not observed in purified HBcAg preparations. The entrapped protein was readily removed by taking advantage of the self-assembly properties of HBeAg. Thus, protein was depolymerized under nondenaturing conditions, purified by gel filtration as a low molecular weight species, and then reassembled into particles by adjusting the solvent pH.

HBeAg was also produced by proteolytic digestion of intact HBcAg capsids (see Materials and Methods). The tryptic cleavage site was determined by mass spectroscopy to be between arginine residues 150 and 151, producing the C-terminal sequence: (etc.)...TTVVR. Amino acid analysis gave a good correlation with an amino acid composition expected for protein processed at this site. The C-terminal

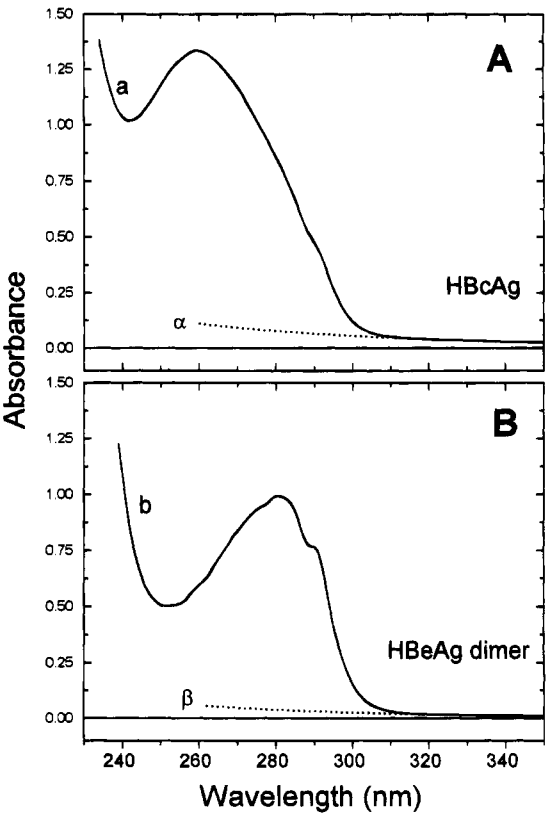

FIGURE 1: Absorbance and light scattering of HBeAg capsids and dimeric HBeAg. Panel A: HBeAg capsids in 50 mM Tris-HCl, pH 7.5, and 150 mM NaCl. Panel B: HBeAg in 100 mM Na_2CO_3-$NaHCO_3$, pH 9.5. Absorbance (spectra a and b) was measured in a in 1 cm path length cell, and light scattering (plots α and β) was derived from optical densities from 320 to 400 nm and extrapolation of the log A vs log λ plot to lower wavelengths. Note: The spectrum of HBeAg capsids at pH 7 was the same as that of dimeric HBeAg except that the scattering plot was similar to that shown by HBcAg (plot α).

sequence of the native and nonparticulate HBeAg protein isolated from human serum, namely, TTVV, is one residue shorter (Takahashi *et al.*, 1983).

Molecular Weight and Subunit Composition. Gel filtrations of HBcAg and HBeAg capsids on Sepharose CL-4B gave single peaks with fairly symmetrical appearances. Sedimentation velocity analyses of fractions from the ascending slope, the peak apex, and the descending slope of the peaks indicated size heterogeneity. For HBcAg capsids, the front and apex fractions contained mostly 96S species (Figure 2A) and the descending slope fraction a mixture of 96S and 82S material (Figure 2B). These two species could be resolved preparatively on isokinetic sucrose density gradients and were observed to have similar protein:nucleic acid ratios as judged by UV/vis spectroscopy (data not shown). The sedimentation coefficients estimated by this technique were similar to those determined by analytical ultracentrifugation. The 96S material constituted about 60−70% of the mixture.

HBeAg capsids were also resolved into two components, the ascending slope fraction containing 65S and 44S species (Figure 2C) and the peak and descending slope fractions, predominately the 44S species (Figure 2D). The 44S species constituted about 80% of the mixture.

The molecular weights of fractions from the Sepharose CL-4B chromatography of HBcAg capsids were determined

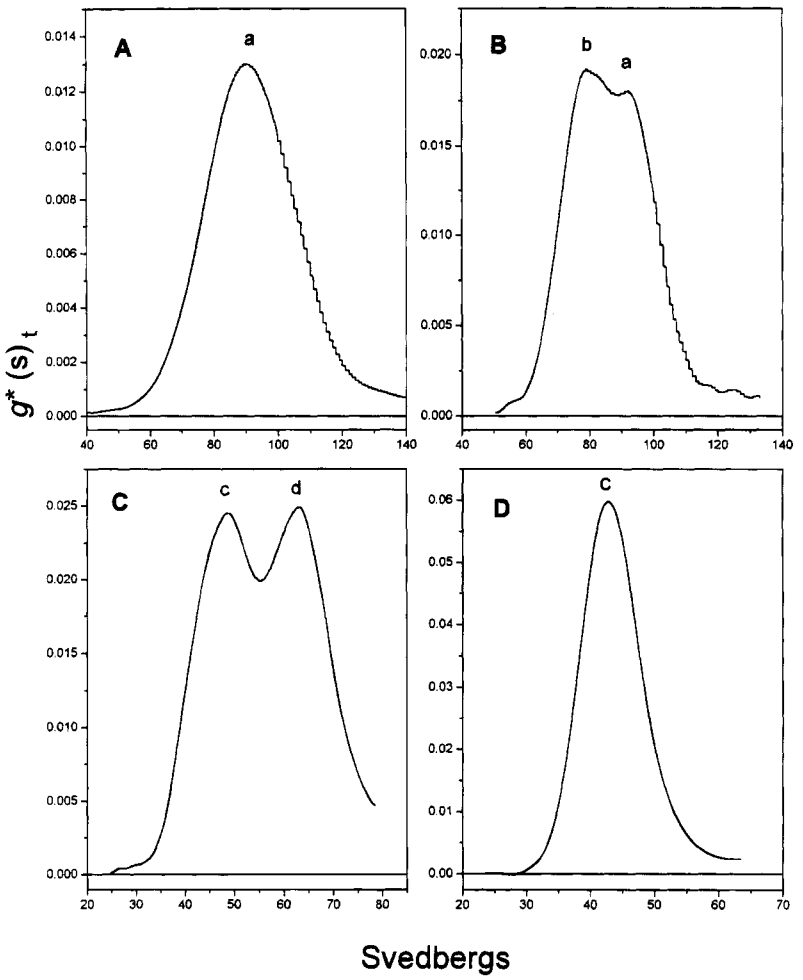

FIGURE 2: Sedimentation velocity of HBV capsid proteins. Abscissa: Svedberg units (10^{-13} s). Ordinate: The apparent sedimentation coefficient distribution coefficient (uncorrected for diffusion effects) is indicated by $g^*(s)_t$, where t = time. For further detail see Stafford (1992). Panel A: HBcAg front or apex fractions from Sepharose CL-4B chromatography (see text for details). Panel B: As in (A), descending slope fraction. Panel C: HBeAg front peak from Sepharose CL-4B chromatography. Panel D: as in (C), apex or descending slope fraction. Peaks labeled a, b, c, and d correspond to the 96-, 82-, 44-, and 65-S species described in Table 1.

by sedimentation equilibrium. The resulting apparent masses ranged from 5.2 to 6.1 × 10^6 Da. For the mass calculations, a partial specific volume of 0.692 mL/g was used; this value is based upon an assumed 20% weight contribution from RNA (partial specific volumes of 0.53 and 0.733 mL/g were used for constituent RNA and protein, respectively). A 5% higher RNA content (i.e., 25%) would reduce the estimated mass by about 3%, and vice versa for a 5% lower (i.e., 15%) RNA content. The equilibrium profile for HBcAg protein from the ascending peak fraction is shown in Figure 3A.

The protein mass contribution to the capsid ranged from 4.16 to 4.88 × 10^6 Da. If we take into account the nucleic acid content and assume that the protein:nucleic ratios are constant in the various size classes, then there are between 189 and 222 subunits per capsid (Table 1). The lower subunit numbers are close to the expected complements for icosahedral particles with triangulation numbers of $T = 3$ (180 copies), and the higher numbers are suggestive of $T = 4$ (240 copies), respectively (Caspar & Klug, 1962).

HBeAg capsids were analyzed in a similar manner as HBcAg, the apparent molecular weight range for the various Sepharose CL-4B fractions being (2.9−4.2) × 10^6, and the peak fraction having a mass of 3.1 × 10^6 Da (Figure 3B). In this case, the calculation of the protein mass was not complicated by the presence of nucleic acid (a partial specific volume of 0.743 mL/g was used). The larger and smaller particle mass values correspond to 242 and 167 subunits per particle, respectively. In contrast to HBcAg capsids, the majority of HBeAg capsids appear to have $T = 3$ symmetry. The results from the combined centrifugational study are summarized in Table 1.

In the above mass determinations, the equilibrium data for capsids were fitted assuming a single species was present; hence, the derived molecular weights are apparent values. When the data were fitted assuming a $T = 4/T = 3$ mixture, the summed proportions of the two species were equal to the average total mass determined from the single species model. For example, the apparent mass of HBcAg capsids in Figure 3A is 6.1 × 10^6 Da. When the data are fitted to a two-species model, the $T = 4$ ($M_r = 6.59 \times 10^6$) and $T = 3$ ($M_r = 4.94 \times 10^6$) capsids contribute about 65% and 35%, respectively.

STEM allows mass determination of macromolecular complexes on an individual basis by dark-field imaging of unstained specimens (Wall, 1979; Wall & Hainfeld, 1986; Thomas *et al.*, 1994). This is useful for investigating the variability within populations of particles that are heterodisperse as with HBcAg and HBeAg capsid preparations. STEM micrographs were recorded for purified fractions of both HBcAg and HBeAg capsids (Figure 4A). Most of the

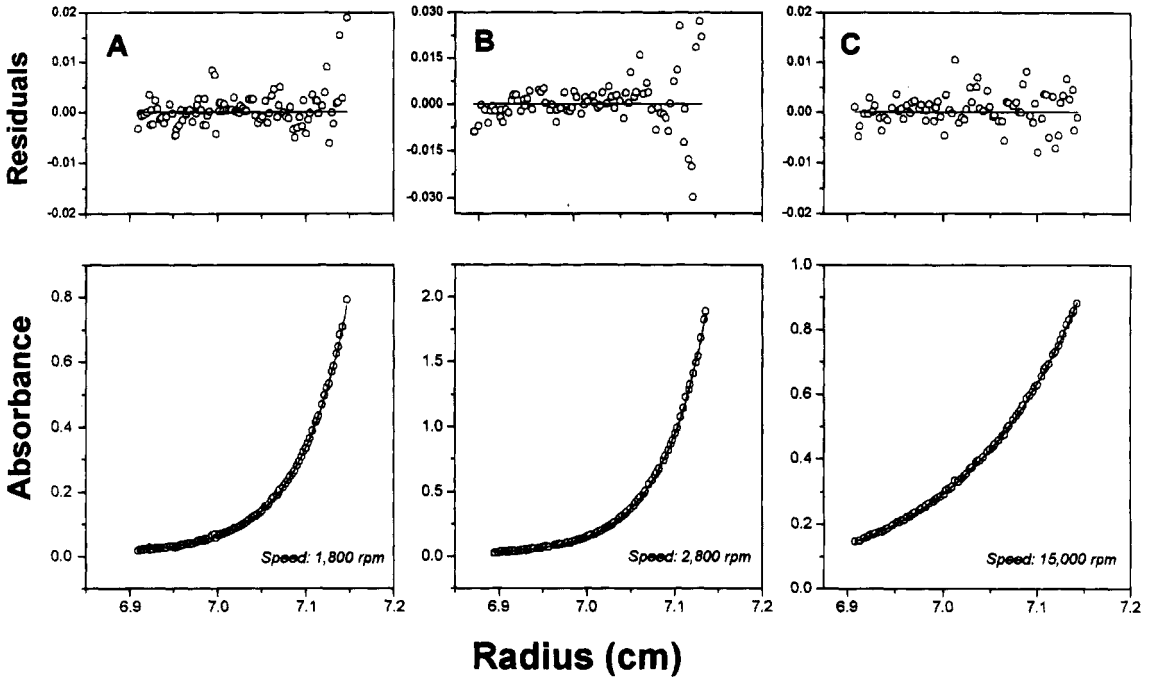

FIGURE 3: Sedimentation equilibrium of HBV capsid proteins. The bottom panels show the absorbance gradients at 290 nm in analytical centrifuge cells at sedimentation equilibrium. The solid line is the fit to a single ideal sedimenting species. The top panels indicate the residuals (subtraction of absorbance of fit minus absorbance of data) as a function of radius. Panels A: HBcAg, ascending slope fraction from Sepharose CL-4B chromatography (see text for details). Panels B: HBeAg, peak fraction from Sepharose CL-4B. Panels C: Peak fraction of HBeAg chromatographed at pH 9.5 on Sephacryl S-200. The molecular weights calculated from these data are given in Table 1.

Table 1: Analytical Ultracentrifugation of HBV Capsid Proteins[a]

preparation	apparent molecular weight	svedbergs (10^{-13} s)	subunits
HBcAg			
A	6.1×10^6	96.0	222
B	5.5×10^6	96.0	200
C	5.2×10^6	96.0, 82.0	189
HBeAg			
A	4.2×10^6	44.0, 65.0	242
B	3.1×10^6	44.0	178
C	2.9×10^6	44.0	167
HBeAg			
D	34 600	2.48	2

[a] Molecular weights were determined by sedimentation equilibrium. For the capsid proteins, A, B, and C refer to ascending slope, peak apex, and descending slope fractions of protein gel filtrated on Sepharose CL-4B at pH 7.0. D refers to the peak fraction from gel filtration of HBeAg on Sephacryl S-200 at pH 9.5. The equilibrium profiles for HBcAg (A), HBeAg (B), and D are shown in Figure 3, panels A, B, and C, respectively. Subunit compositions were determined by dividing the apparent molecular weight of the preparations, after subtracting any nucleic acid contribution, by the molecular weight of the subunits. For the HBcAg capsids, a nucleic acid content of 20% was assumed. Monomer molecular weights of 21 970 (HBcAg) and 17 370 (HBeAg) were derived from the respective DNA coding sequences. The sedimentation coefficients ($s_{20,w}$) were determined by the transport method (Schachman, 1959). The boundaries corresponding to the respective $T = 4$ capsids were fairly well resolved from the slower moving $T = 3$ boundaries.

particles in these fields appear well preserved structurally (i.e., not flattened), so that measurements of their diameters can be made, in addition to mass determinations. These parameters were plotted as bivariate distributions (Figure 4B). In all three experiments, the data segregate diagonally into two discrete populations. There is some variability within each population, but this dispersion is sufficiently small that the two populations are clearly resolved. The separation is clear along the mass axis; along the diameter axis, the separation is marginal but contributes to a more conclusive separation between the two populations in two dimensions.

HBeAg capsids average about 26 nm in diameter, and 3.17 $\times$ 10^6 Da in mass, corresponding to 182 $\pm$ 10 copies of the truncated capsid protein. About 15% of the particles have a higher average mass of 4.14 $\times$ 10^6 Da (corresponding to 238 $\pm$ 11 copies of the capsid protein). These copy numbers correspond to the expected complements for icosahedral particles with triangulation numbers of $T = 3$ and 4, respectively. The diameters of such particles would expect to bear the ratio $(4/3)^{0.5} = 1.15$, and this figure is in good agreement with the observed data (see Table 2).

The micrographs of HBcAg capsids show denser particles (cf. Figure 4A, a and b). Bivariate distributions of mass and diameter were compiled for HBcAg capsids, taken from different fractions after gel filtration on Sepharose CL-4B (Figure 4B, a and b). Both experiments show the same two populations, but differ in how the particles partition between these populations. Fractions from the descending slope of the peak (Figure 4B, a) contain particles one-third of which are the larger species; and in fractions from the ascending slope of the peak (Figure 4B, b), about two-thirds of the particles fall into this category. The outer diameters of these two kinds of particles are not significantly different from those of the two kinds of HBeAg capsids: however, their masses are systematically higher, by about 55% (Figure 4B, c; Table 2). The higher mass reflects the protein mass difference, the presence of nucleic acid (see below), and possibly some trapped protein in the capsids.

The HBV nucleocapsid contains circular DNA of 3.2 kilobases in length ($M_r \sim 1.6 \times 10^6$) which would contribute about 24% and 30% by weight for $T = 4$ and $T = 3$ particles, respectively. By comparison, the recombinant HBcAg $T =$

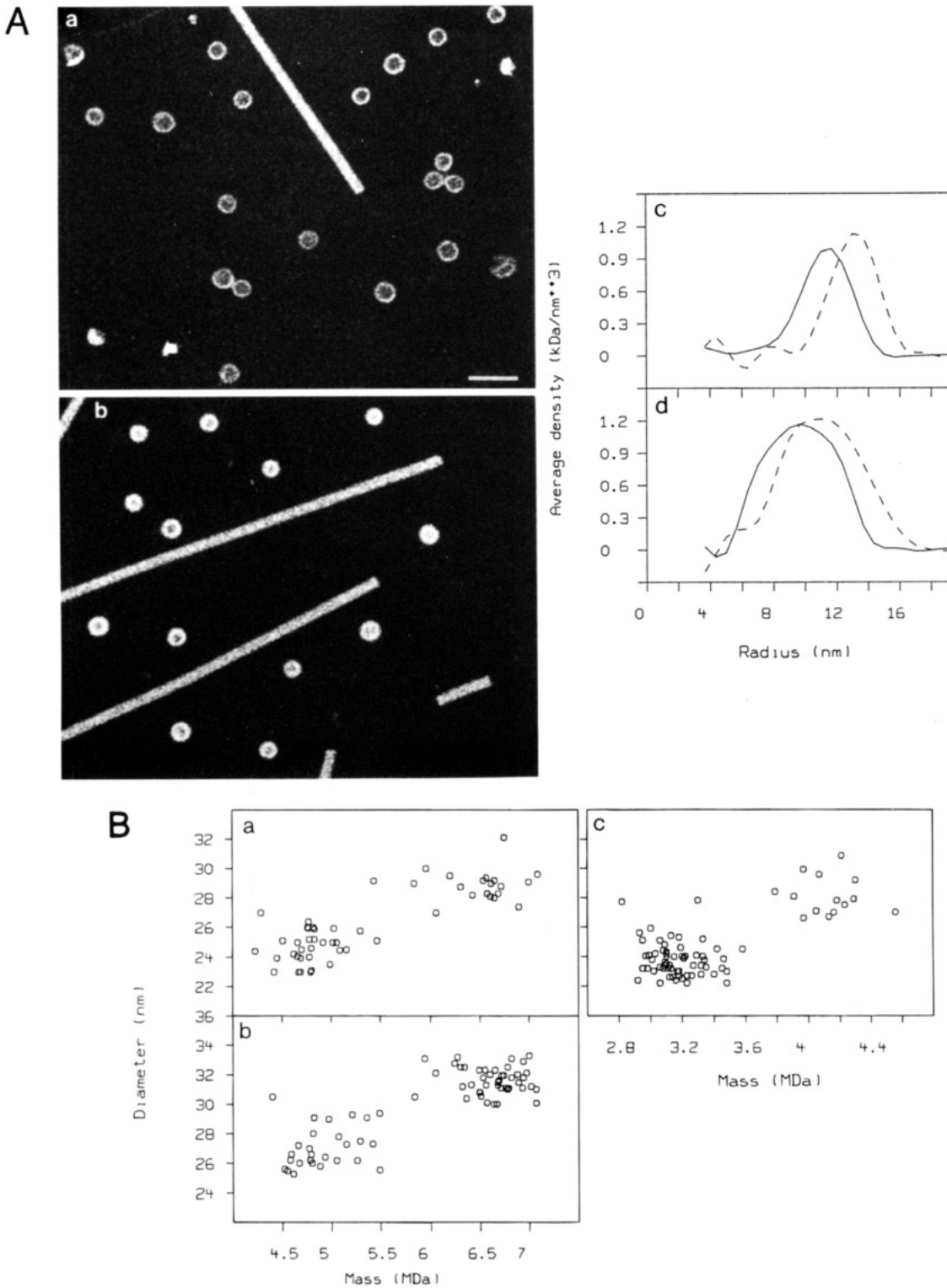

FIGURE 4: (Panel A) STEM of HBV nucleocapsids. Dark-field scanning transmission electron micrographs of unstained freeze-dried specimens of (a) HBeAg particles and (b) HBcAg particles. Both fields contain particles of two different sizes. The rod-like particles are tobacco mosaic virions, used as an internal standard. The two micrographs were normalized so that the same density is conveyed by the same gray level in each case. Bar = 50 nm. Radial profiles of spherically averaged density were calculated for both size classes in each case: (c) and (d), respectively. (Panel B) Size distributions of HBV nucleocapsids. Two-dimensional distributions of measurements of mass and diameter made from dark-field STEM micrographs of HBcAg (a and b) and HBeAg particles (c). In all three cases, the data segregate diagonally into two subpopulations.

4 and $T = 3$ capsids contain lower but similar amounts of *E. coli* RNA, namely, 19% and 20%, respectively (Table 2).

Radial Density Profiles of Distinct Species of Particles. STEM micrographs show projections of mass density, and the visual appearance of HBeAg and HBcAg capsids suggests that both are hollow and that HBeAg capsids are thinner-walled than HBcAg capsids (cf. Figures 4A, a and b). This property was investigated in a more quantitative way by computationally reconstructing the average radial density profile for each of the four species of particles (Steven *et al.*, 1984). To suppress noise, the data were

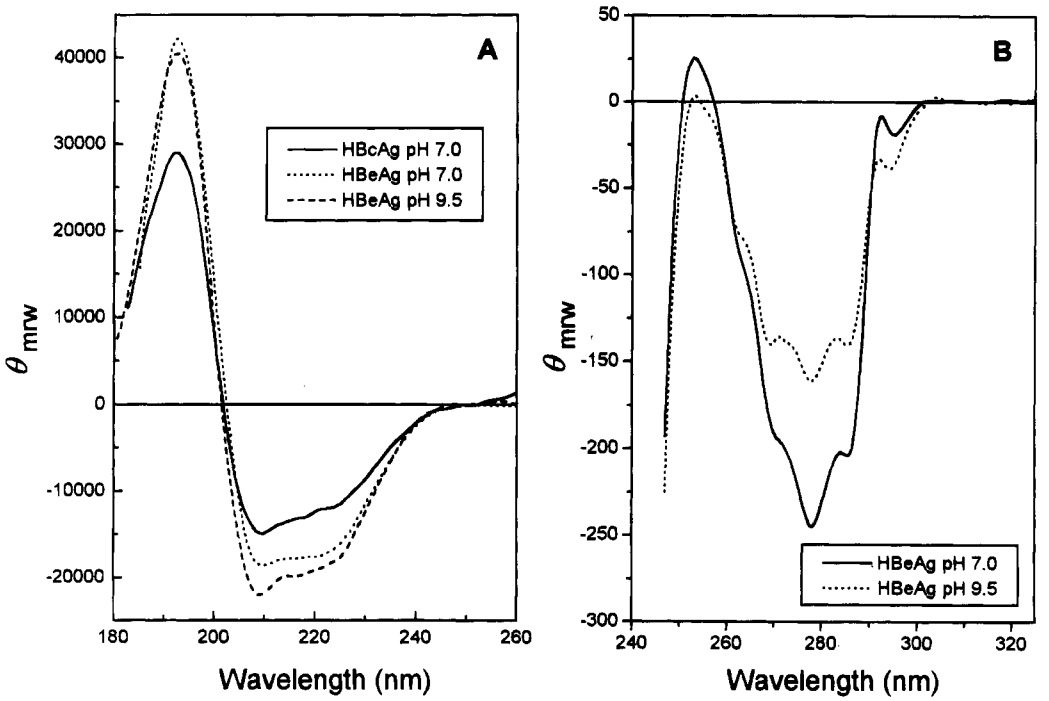

FIGURE 5: Circular dichroism of HBV nucleocapsids. Panel A: Far-UV CD spectra of nucleocapsid proteins. Panel B: Near-UV CD spectra of HBeAg proteins. The ordinate (θ_{mrw}) is the mean residue ellipticity and has units of deg·cm^2·dmol^{-1}.

Table 2: Mass and Dimensions of HBV Capsids Measured from Dark-Field STEM Micrographs[a]

	mass (MDa)				diam.[b]			thickness		
	mean	SD	SEM	*N*	mean	SD	diam.[c]	50%	25%	subunits
HBcAg										
small	4.86	0.30	0.04	60	27.0	1.4	26.3	6.5	8.4	177
large	6.60	0.31	0.04	63	31.4	1.0	29.8	6.9	8.3	240
HBeAg										
small	3.17	0.17	0.02	63	25.8	1.1	26.2	3.5	4.6	182
large	4.14	0.19	0.05	16	30.1	1.3	30.2	3.5	4.9	238

[a] Capsid diameters and wall thickness are expressed in nm. [b] Diameters were measured by using a cursor to manually define the edges of particles displayed on a graphics monitor. [c] Diameters were determined from the averaged radial profiles (Figure 4A, c and d), defining the edge as the outermost point at which the density profile fell to 50% of maximum. The thickness was determined from the radial profiles as the space between the two points at which the density profile fell to 50% (respectively, 25%) of maximum above background. The number of subunits was determined as described in the footnote to Table 1.

azimuthally averaged on each particle, and the resulting projection traces averaged over many particles of each type. The resulting profiles are compared in Figure 4A, c and d. With HBeAg capsids of both sizes, the radial density profiles are symmetrical and of similar width, ~3.5 nm at half-height, and 4.5−5.0 nm at 25%-of-maximum. The only difference between these profiles is a relative displacement by ~2 nm to higher radii, suggesting that essentially the same surface lattice is folded into shells of two different sizes.

The two kinds of HBcAg capsids also have similar radial distributions of density. Their peak densities are very similar to those of their respective HBeAg counterparts. However, their walls are thicker by 3−4 nm. As their outer radial limits match those of the corresponding HBeAg capsids, it appears that the inner surfaces of HBcAg capsid are lined with nucleic acid that contributes most of the additional mass (Table 2).

Depolymerization of Capsids. HBcAg capsids are extremely stable structures, requiring strongly denaturing conditions such as heating in the presence of 8 M Gdn-HCl and reducing agent to dissociate the protein subunits completely. On the other hand, the nucleic acid free HBeAg capsids were largely depolymerized at pH 9.5 by nondenaturing concentrations of urea (1.5−2.0 M). Depolymerized protein was separated from polymeric material by gel filtration and subjected to analytical ultracentrifugation. A molecular weight of 36 000 was determined indicating a dimeric structure (Figure 3C). The dimers were stable at pH 9.5, showing no tendency for association or dissociation. Sedimentation velocity analysis of the dimers ($s_{20,w}$ = 2.48 S) indicated they have an elongated shape with a frictional coefficient of about 1.3.

HBeAg capsids derived by proteolysis of HBcAg or by direct expression had indistinguishable biochemical and physicochemical properties.

Disulfide Analysis. As in earlier reports (Gallina *et al.*, 1989), our HBcAg preparations contained intermolecular disulfide linkages. Monomeric protein was not observed by SDS−PAGE of HBcAg capsids under nonreducing conditions: namely, protein pretreated sequentially with iodoacetamide (to prevent artifactual disulfide exchange) and SDS in the absence of reductant. The protein migrated as dimers or failed to enter the gel. The relative proportions of dimers and higher multimers varied from preparation to preparation, the latter increasing with storage time. Predominantly monomeric, and to a lesser degree, dimeric protein was observed in nonreducing SDS−PAGE of freshly prepared tryptic processed HBcAg protein in which the C-terminal region was removed. Thus, the C-terminal cysteine residue participates in many of the intermolecular disulfide linkages, as has been previously suggested (Gallina *et al.*, 1989; Nassal *et al.*, 1992).

In processed protein, there was a time-dependent increase in the amount of intermolecular disulfide-linked dimer, as

Table 3: Secondary Structure of HBV Capsid Proteins[a]

protein	method	secondary structure (%) α-helix	β-sheet	β-turn	other
HBcAg-capsid	A	39	17	27	17
	B	49	5	9	36
	C	42	35	3	21
HBeAg-capsid	A	53	13	21	13
	B	55	11	17	18
	C	62	16	7	27
	D	71	7	13	9
HBeAg-dimer	A	58	7	22	13
	B	55	13	13	19
	C	62	6	7	27

[a] A, B, and C were estimated from far-UV CD spectra by the respective methods of Sreerama and Woody (1993), Perczel *et al.* (1992), and Chang *et al.* (1978). D was estimated from the Raman spectrum of HBeAg capsids (Figure 6) by the method of Williams (1986).

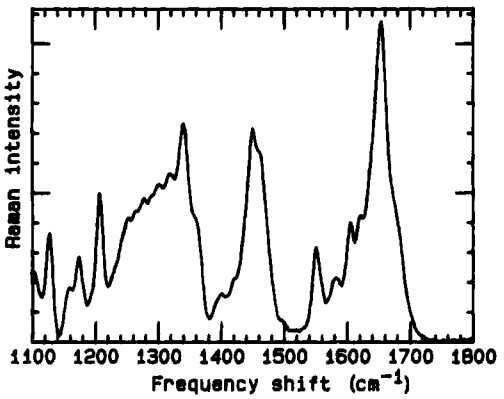

FIGURE 6: Raman spectrum of HBeAg nucleocapsids. The amide I band between 1620 and 1710 cm^{-1} is relatively narrow, and the maximum is near 1650 cm^{-1}. These characteristics are typical of helical proteins. The amide III region between 1220 and 1320 cm^{-1} is also typical of helical proteins, showing more intensity near 1300 cm^{-1} (the helix region) than at 1240 cm^{-1} (the β-sheet and unordered structure region). The secondary structure content calculated from the amide I band is 71% helix (this includes all types of helix) and 7% β-sheet.

judged by nonreducing SDS−PAGE. At pH 7.0 where the protein is polymeric, the covalent dimer content increased from about 10% to 90% over 48 h. The rate of intermolecular disulfide bond formation was similar for dimeric protein at pH 9.5.

Secondary Structure and Conformational Analysis. The overall conformation of HBcAg and HBeAg capsids was measured by far-UV CD analysis (Figure 5A). Both particles exhibited spectra typical of α-helix-containing proteins with minima at 220 and 209 nm and a strong positive peak at 192 nm. Secondary structure analysis of the capsids by several methods indicated 43−60% α-helix and a low content of β-sheet (Table 3). HBeAg capsids had about 15% higher α-helix content than HBcAg capsids, indicating that the C-terminal region in HBcAg capsids is probably nonhelical. Hence, removal of 17% of the protein by weight (33 residues) increases the overall helical content of the deletion mutant by a proportional amount. Dimeric protein from depolymerized HBeAg capsids has essentially the same overall conformation as the parent capsids (Table 3).

The secondary structure of the HBeAg capsids was also determined by Raman spectroscopy (Figure 6). Analysis of the spectrum using non-negative least squares fits (Williams, 1986) indicated 71% α-helix and 7% β-sheet (see Table 3). The Raman amide I spectrum was very similar to that of hemerythrin.

The near-UV CD spectrum of HBcAg capsid was dominated by the RNA content, as indicated by a fairly broad positive peak centered at 264 nm (data not shown). The spectrum was similar to that of *E. coli* tRNA (Okabe *et al.*, 1977). The spectra of HBeAg capsids and dimeric protein were qualitatively similar, with a weak negative tryptophan band at 295 nm and strong negative bands between 269 and 286 nm probably derived from tyrosine. For HBeAg, the band intensity of the main negative peak at 278 nm was dependent on the amount of polymer present. Hence, the ellipticity of HBeAg capsid protein at 278 nm was >1.5-fold higher than that of dimeric protein.

Denaturation of either HBeAg capsids or dimers with 5 M Gdn-HCl resulted in loss of the near-UV CD signal, indicating that the spectra bands were derived from asymmetrically oriented aromatic groups as is normally found in native protein (Strickland, 1974).

Dissociation and Denaturation Properties: Spectroscopic Observations. HBeAg capsids or trypsinized HBcAg capsids were dissociated into dimers with 1.5−2.0 M urea at pH 9.5 as determined by sedimentation velocity or gel filtration. Urea at these concentrations had no effect on either the near- or far-UV CD spectra of the dimers, indicating that depolymerization occurred without protein denaturation.

The fluorescence emission maximum (λ_{max}) of either HBcAg or HBeAg capsids was at 326 nm (Figure 7 and Table 4). The λ_{max} was considerably blue-shifted relative to that of free solvent-exposed tryptophan (λ_{max} of *N*-Ac$TrpNH_2$ = 353 nm), suggesting that in capsids the 4 tryptophan residues are on average shielded from the solvent (Burnstein *et al.*, 1973).

Denaturation of HBcAg with 5 M Gdn-HCl was indicated by a 14−15 nm red shift in λ_{max} (Kronman & Holmes, 1971). There was also a 55−60% increase in emission intensity. Similar results were also observed with the HBeAg capsids, although the red shift was 5 nm longer, indicating more extensive unfolding or solvent exposure of protein tryptophans (Figure 7).

The increase in fluorescence emission by both HBcAg and HBeAg capsids upon unfolding suggests quenching of buried tryptophan(s) in the folded state. In this situation, protein denaturation moves the quenching group(s) from the neighboring tryptophan. For buried and unquenched tryptophans, there is usually a decrease in emission intensity with increased solvent exposure (Kronman & Holmes, 1971). Sulfhydryl and especially disulfide groups are strong quenchers of indole fluorescence (see, e.g., Kikuchi *et al.*, 1986). It is noteworthy that Trp-61 is adjacent to a cysteine residue (Figure 2) which may be involved in disulfide bond formation (Zheng *et al.*, 1992).

The above results indicate that the tryptophan environments in HBcAg and HBeAg are similar and further suggest that the C-terminal region, present in HBcAg and absent in HBeAg, does not significantly effect the overall conformation of capsid subunits. (It should be noted that there are no aromatic residues in the first 51 residues of the C-terminal sequence.)

The tryptophan emission of HBeAg dimer was about 30% lower and red-shifted by 4−5 nm compared to HBeAg (or HBcAg) capsids (Table 4). Denaturation of HBeAg dimers

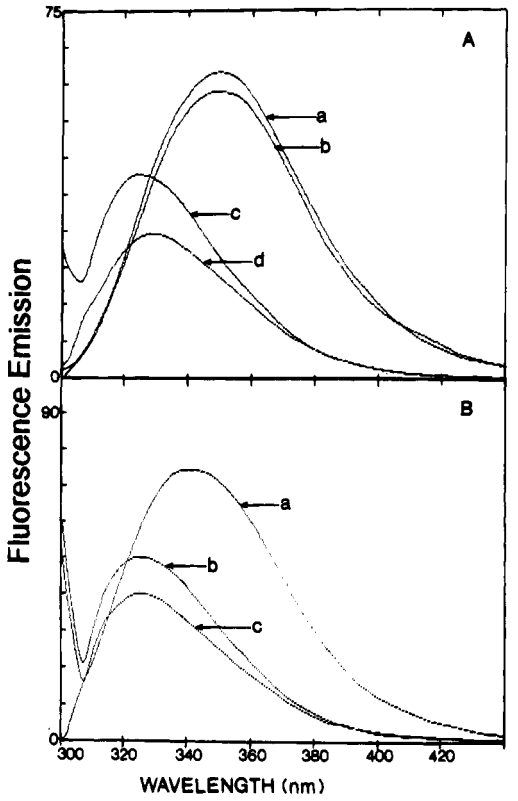

FIGURE 7: Fluorescence emission spectra of HBV nucleocapsid proteins. The top panel (A) refers to HBeAg proteins (10 μg/mL): (a) in 100 mM Na_2CO_3-$NaHCO_3$, pH 9.5, containing 6 M Gdn-HCl; (b) in 100 mM Tris-HCl, pH 7.0, containing 6 M Gdn-HCl; (c) in 100 mM Tris-HCl, pH 7.0, and (d) in 100 mM Na_2CO_3-$NaHCO_3$, pH 9.5. The respective polymeric and dimeric state of HBeAg used in (c) and (d) was established by analytical ultracentrifugation. The bottom panel refers to HBcAg (10 μg/mL): (a) in buffer plus 5 M Gdn-HCl; (b) in buffer alone, and (c) in buffer plus 2 M Gdn-HCl. The buffer for HBcAg was 100 mM Tris-HCl, pH 7.5. The ordinate scales (fluorescence emission) are in arbitrary units. The excitation wavelength was 295 nm. Quantitative aspects of the data are presented in Table 4.

Table 4: Fluorescence Emission of HBV Capsid Proteins[a]

protein	addition to buffer	fluorescence emission	
		λ_{max} (nm)	intensity
HBcAg-capsid	none	326	846
	2 M Gdn-HCl	326	630
	5 M Gdn-HCl	341	1313
HBeAg-capsid	none	325	840
	2 M Gdn-HCl	329	688
	5 M Gdn-HCl	345	1168
HBeAg-dimer	none	329	606
	2 M Gdn-HCl	337	884
	5 M Gdn-HCl	347	1339
N-AcTrpNH$_2$	none	356	945
	5 M Gdn-HCl	358	1159

[a] The excitation wavelength was 295 nm. Fluorescence emission intensities are in arbitrary units standardized to 2.8 μM protein. The *N*-AcTrpNH$_2$ concentration was equivalent to that of 2.8 μM HBcAg protein. Spectra were recorded at 22 °C after 1 h of incubation at the same temperature. The buffers correspond to those described in the legend to Figure 7.

with 5 M Gdn-HCl normalized the spectrum to that of unfolded HBeAg capsid (Figure 7). The additions of lower concentrations of Gdn-HCl (e.g., 2 M) to either HBcAg or HBeAg capsids gave spectra resembling the HBeAg dimer, suggesting capsid dissociation with little or no protein denaturation (Figure 7 and Table 4). The higher fluorescence

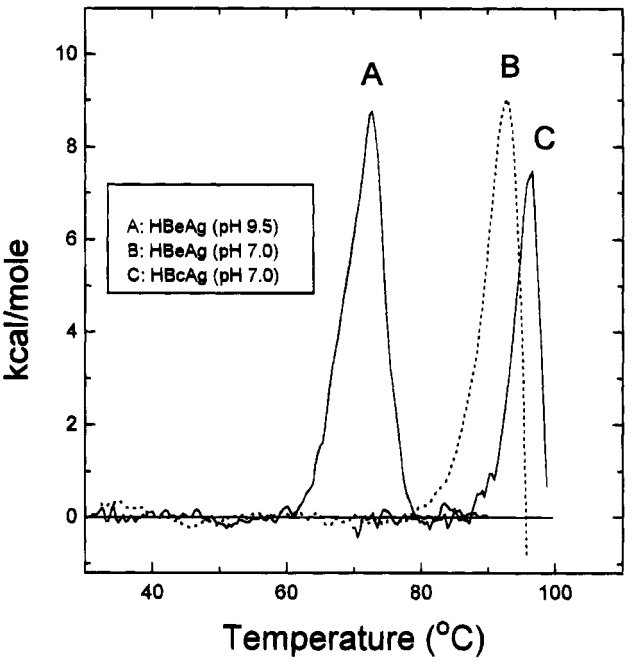

FIGURE 8: Differential scanning calorimetry of HBV nucleocapsid proteins. Thermograms of (A) HBeAg in pH 9.5 buffer; (B) HBeAg in pH 7.0 buffer; and (C) HBcAg in pH 7.5 buffer. Buffer base lines were subtracted and the the data normalized for concentration. Best-fitted base lines constructed for the normalized curves were also subtracted.

intensity of HBeAg capsids compared to dimers may be due to immobilization and solvent shielding in the former of partially exposed tryptophan residue(s) close to, or at, interfacial regions.

Capsid Thermal Stability. The thermal denaturation of HBcAg is characterized by an endothermic transition with an apparent T_m of 97 °C (Figure 8C). Protein aggregation occurs before the true T_m is reached as indicated by a large exothermic transition at >97 °C. HBcAg heated in the presence of 5 mM dithiothreitol melted about 4−6 °C lower. HBeAg capsids were also hyper-stable structures exhibiting a single transition with an apparent T_m of about 93 °C (Figure 8). In contrast to HBcAg, inclusion of dithiothreitol had no effect on the T_m observed for the denaturation of HBeAg capsids. HBeAg capsids also showed a large exothermic transition at or just beyond the transition temperature. As no endotherms were observed on rescanning either kind of capsid, their thermal denaturation appears to be an irreversible process.

Denaturation of HBeAg dimers was characterized by a single transition with a T_m of 72 °C (Figure 8). The DSC peak was asymmetric, with the advancing edge being broader and less sharp than the trailing edge. The ΔH_{cal} (calorimetric heat exchange) for the transition was 63 kcal/mol, and the ratio $\Delta H/\Delta H_v$ was 0.53, where ΔH_v is the van't Hoff enthalpy. A ratio value of about 0.5 corresponds to a dimeric protein undergoing a single coupled transition on unfolding (review: Sturtevant, 1987). We conclude that capsid dissociation and the unfolding of dimeric HBeAg occur simultaneously.

In the presence of dithiothreitol, the T_m of HBeAg dimers was reduced by 5 °C to 67 °C (data not shown). HBeAg dimers are, thus, partially stabilized against thermal denaturation by the presence of intermolecular disulfide linkage(s).

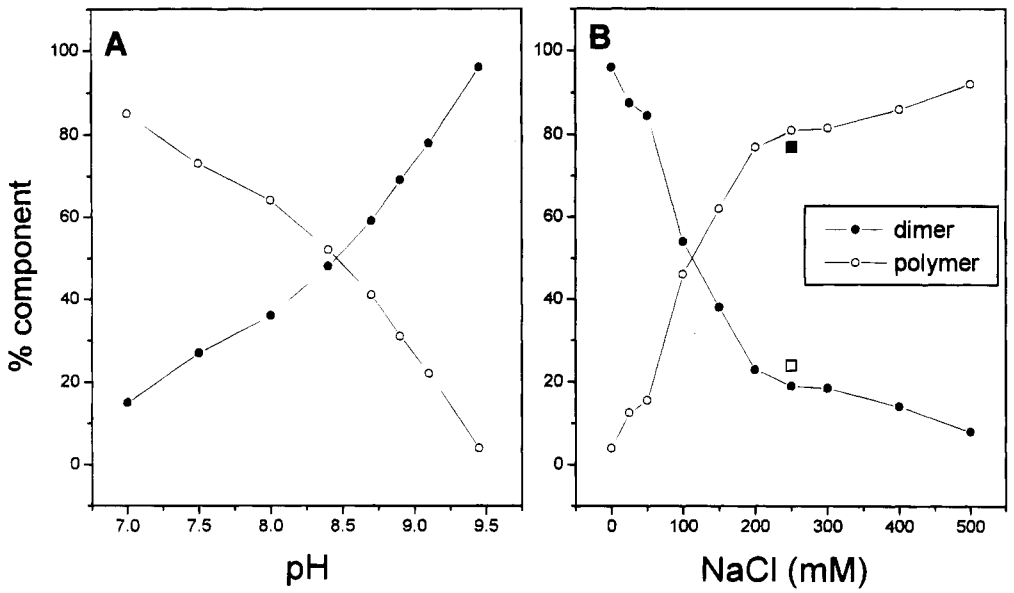

FIGURE 9: In vitro assembly of dimeric HBeAg into capsids: pH and salt dependency. Panel A: The pH dependency of dimer assembly into capsids was determined by dialyzing HBeAg (0.5 mg/mL) in pH 9.5 buffer against buffers of the indicated pH which contained 250 mM NaCl for 12−14 h at 4 °C. Panel B: The salt dependency of capsid assembly was determined by rapidly exchanging HBeAg in pH 9.5 buffer to 50 mM Tris-HCl, pH 7.0, using a PD-10 column (Pharmacia) of Sephadex G-25 (exchange time 2−3 min). The protein was diluted to 0.5 mg/mL in 50 mM Tris-HCl, pH 7.0, containing the indicated concentrations of NaCl and incubated for either 0.5 h (at 22 °C) or 14 h (at 4 °C). In one experiment, indicated by filled and open squares, the protein was diluted to 50 μg/mL in 250 mM NaCl. For both pH and salt dependency studies, the concentration of dimer and polymer was determined by sedimentation velocity.

The difference in stability between HBeAg capsids and dimers (measured by monitoring the increase in fluorescence at 340 nm) was also apparent when the midpoints of the Gdn-HCl-induced unfolding transitions (C_m) were compared, namely, 4 and 1.8 M for capsid and dimer, respectively (data not shown).

Polymerization of Dimeric Protein. As described above, HBeAg capsids consist of 44S and 65S species. These polymers can be easily resolved from dimeric protein (2.48S) by sedimentation velocity analysis. Measurement of the height (UV absorbance) of the sedimenting boundaries allows the concentrations of the various species to be determined.

HBeAg capsids were dissociated into dimers by treatment with 2 M urea at pH 9.5 and gel filtrated in buffer alone at the same pH. The low molecular weight fraction (dimeric protein) was used for assembly experiments. Initially, the pH dependency of self assembly was examined. Polymerization increased progressively with decrease in pH and was complete at pH 7.0 (Figure 9A). A moderately high ionic strength was critical for assembly. In the experiment shown in Figure 9B, dimeric protein was rapidly transferred from pH 9.5 buffer into 50 mM Tris-HCl, pH 7.0, by gel filtration, and then NaCl was added over the concentration range shown. In the absence of NaCl, dimeric protein failed to polymerize even at pH 7. Polymerization was complete at NaCl concentrations >250 mM. Sedimentation velocity measurements were made after incubations for either 0.5 h (at 22 °C) or 14 h (at 4 °C) with similar results.

As expected for an assembly process, protein concentration was an important factor. For example, in the above experiment, if the protein concentration was reduced by a factor of 10 (to 50 μg/mL), then at 250 mM NaCl only 20% of the protein polymerized as compared to 80% at 0.5 mg/mL (Figure 9B).

Capsids assembled under all the experimental conditions described above appeared to be composed of predominantly $T = 3$ particles, similar to those prepared by the large scale dialysis procedure used for preparative purposes (see Materials and Methods section). Furthermore, assembly intermediates have not been observed; i.e., only unassembled dimers and capsids are detected by the sedimentation velocity analysis.

DISCUSSION

Dimorphism ($T = 3$ and $T = 4$ Variants) of both HBcAg and HBeAg Capsids. Native hepatitis B virus core particles from human liver visualized by negative staining have been reported to have diameters of 27−30 nm and an icosahedral appearance (e.g., Onodera *et al.*, 1982). Similar dimensions and morphologies have also been reported for recombinant capsids produced in *E. coli* (the present and previous studies); in yeast (Yamaguchi *et al.*, 1988); in insect cells with a baculovirus expression vector (Hilditch *et al.*, 1990); and in a *Xenopus* oocyte system (Zhou & Strandring, 1991). The subunit composition and triangulation number of the supposedly icosahedral capsids were not conclusively determined, although most workers have assumed a triangulation number of $T = 3$ (Onodera *et al.*, 1982).

During preparation of this paper, a study of recombinant HBcAg capsids by cryo-electron microscopy and image reconstruction was published by Crowther *et al.* (1994), who reported capsid dimorphism with surface lattices of both $T = 3$ and 4. Our analysis, by different methods, confirms and extends this observation. Using STEM and analytical ultracentrifugation, methods which yield the masses of macromolecular assemblies, we have demonstrated the same $T = 3/T = 4$ dimorphism, whether full-length HBcAg is assembled intracellularly in *E. coli*, or whether the truncated form, HBeAg, is assembled *in vitro* from purified, soluble protein. For HBcAg, the proportion of the two classes varied slightly from preparation to preparation, but usually >60% corresponded to $T = 4$. HBeAg capsids assembled *in vitro*

mostly assume the $T = 3$ lattice, with 15−20% in the $T = 4$ form.

In the only other physicochemical analysis of recombinant-produced HBV capsids of which we are aware, Hilditch *et al.* (1990) studied HBcAg from a baculovirus expression system, obtaining a sedimentation coefficient of 82.5 S and a calculated molecular weight of $(5.8-6.3) \times 10^6$. From these data, the authors concluded a copy number of about 300, which is incompatible with $T = 3$ (180 subunits) or $T = 4$ (240 subunits). Although their findings are in part similar to our own, there is inconsistency in that, unlike our capsids from *E. coli* which are about 20% (w/w) nucleic acid, the capsids from insect cells were reported to be nucleic acid free. We suspect that the discrepancy may reflect loss of nucleic acid during purification in CsCl gradients, as we have observed nucleic acid depletion during this type of analysis, and Onodera *et al.* (1982) observed the breakdown of authentic HBcAg capsids in this medium and cautioned against its use.

It is not yet known whether native HBcAg capsids, located in the virus core or as uncoated capsids in hepatocytes, exhibit the same $T = 3/T = 4$ dimorphism. The difference in diameter between these two variants is slight and cannot be determined with confidence by conventional electron microscopic methods.

Location and Role of the Carboxy-Terminal Peptide. This highly basic peptide, which is absent from the truncated protein, HBeAg, is not required for assembly, although if present, it does influence assembly in two ways. It enhances the stability of the capsid, as evidenced by increased thermal stability (Figure 8); and it appears to favor the formation of the larger ($T = 4$) species. The presence of this peptide correlates with the higher mass of the core particles, contributed by an additional layer of density, ~4 nm thick, lining the inner surface of the capsid (cf. Figure 4A, c and d). We infer that this layer of density represents nucleic acid molecules bound to the basic carboxy-terminal peptide (cf. Crowther *et al.*, 1994).

Although the HBcAg C-terminal peptide is normally located within the capsid shell, under certain circumstances, this may not be true. For instance, chimeric HBcAg proteins, in which this peptide was replaced by certain epitopes, were found to assemble into capsid-like structures in which these epitopes were surface-accessible (Yoshikawa *et al.*, 1993, and references therein). The location of the C-terminus may thus be determined by whether it can bind nucleic acid or not. Nonbinding sequences, as with the aforementioned chimeras, could become surface-exposed by various mechanisms, including accessibility through pores (see above), lattice lesions, or conformational changes (Steven *et al.*, 1991). The native C-terminus may thus become surface-exposed when the binding of nucleic acid is weakened or prevented. We have described *in vitro* conditions that lead to nucleic acid depletion, and a similar situation may arise *in vivo*, for example, by phosphorylation of serine residues in the C-terminus. It appears that the capacity of HBcAg to bind DNA is reduced by phosphorylation (Machida *et al.*, 1991, and references therein).

A Novel, α-Helical, Conformation for a Viral Capsid Protein. HBcAg is predominantly (~50%) α-helical. HBeAg, either particulate or in a dissociated state, has an even higher content of α-helix (55−70%) and very little β-sheet (Table 3). Our confidence that the CD analysis of HBeAg particles was not significantly affected by light scattering effects was enhanced by obtaining very similar results with Raman spectroscopy. The CD spectrum of native HBcAg capsids (Yamaki *et al.*, 1982) is qualitatively similar to that of our recombinant capsids over the range analyzed: 200−240 nm; but the molar ellipticity values were about 2-fold lower with a consequently lower estimate for α-helix content (25%). This discrepancy may reflect a real difference between native and recombinant capsids or—more likely—it may be due to protein concentration differences (i.e., the native protein content may have been overestimated).

We suggest that the higher α-helical content of HBeAg capsids is due to absence of the nonhelical C-terminal region. Its sequence contains three copies of the SPRR peptide which is closely related to the nucleic acid binding motif (SPKK) found in some histones (Suzuki, 1989). Domains composed of multiple copies of this motif are likely to have random coil conformations (Suzuki, 1989). Moreover, in known virus structures, the polybasic nucleic acid binding region is usually disordered (see, for example, Choi *et al.*, 1991).

The high resolution structures of more than a dozen icosahedral viruses have been determined (reviews: Rossman & Johnson, 1989; Stuart, 1993). All but one of these capsid proteins are based on the canonical "jelly roll" fold, consisting of an eight-stranded antiparallel β-barrel. The sole exception is bacteriophage MS2, whose subunits assume a different fold, albeit one that also has a high content of β-sheet (~50%) (Valegård, *et al.*, 1990). The low β-sheet content of HBV capsid proteins, taken in conjunction with their molecular weights, implies that they do not exhibit either of these conformations and must assume some other fold not previously encountered in a viral capsid protein. Burns *et al.* (1992) have commented that the secondary structure of the p27 capsid protein of simian immunodeficiency virus (44% α-helix) is also incompatible with the jelly roll fold. Interestingly, the CD spectrum of this protein is very similar to that of HBcAg. However, it has not yet been established that p27 actually forms an icosahedral shell.

We note, however, that an icosahedral shell composed of predominantly α-helical subunits is not without precedent: the enzyme heavy riboflavin synthase (HRS) from *Bacillus subtilis* has an overall content of 51% α-helix and forms a $T = 1$ icosahedron (Ladenstein *et al.*, 1988).

Regulation of Assembly. Our initial investigations into the assembly process of HBeAg capsids have centered on the self-association properties of dimeric HBeAg obtained from dissociation of HBeAg capsids or from proteolytically processed HBcAg capsids. The small difference in primary sequence made no difference to the results obtained. Dimeric protein rapidly polymerizes to capsid-like structures at neutral pH in buffer containing moderate salt concentrations. There appears to be no requirement for specific cations or anions, although we have taken no precaution to exclude trace amounts of any particular one. Hence, the driving force for assembly at neutral pH appears hydrophobic, and the pH dependency of the process indicates the requirement for protonation of various side chain(s). This dimer-based pathway accords well with the visualization of subunit dimers in assembled capsids (Crowther *et al.*, 1994).

Although we have not yet made a thorough investigation of the effects of protein concentration on assembly (see Figure 9B), it is likely that a rate-limiting step is the initial formation of nucleation centers, or oligomeric precursors,

the formation of which is concentration-dependent (Rossmann, 1984). Recent work on the *in vivo* assembly of HBcAg using a *Xenopus* oocyte model system indicates that protein concentrations of 0.7−0.8 mM are required before assembly of the capsid occurs (Seifer *et al.*, 1993).

Stabilization and Disassembly. HBcAg capsids are extremely stable with a T_m close to 100 °C (Figure 8), and very high concentrations of chaotropic agents plus reductant are required for complete disassembly. This stability derives in part from intermolecular disulfide bonding involving the C-terminal cysteine residues: in the presence of reductant, the T_m falls to ~90 °C (data not shown). HBeAg capsids at neutral pH also have a T_m of 90−92 °C. This difference in stability can be explained in part by differences in the extent of disulfide cross-linking.

Zheng *et al.* (1992) have shown that disulfide bond formation is not essential for the assembly of HBcAg capsid. This is consistent with our observations that disulfide bond cross-linking is a maturing process: this scenario allows that protein initially assembled in the reducing environment of the *E. coli* cytoplasm is probably fully reduced, but on isolation and storage, oxidation occurs, especially involving the highly reactive C-terminal cysteines. These residues are clearly implicated by the fact that freshly prepared HBcAg, when its C-terminus is removed by trypsin digestion and analyzed by SDS−PAGE under nonreducing conditions, is mostly monomeric. When polymeric HBeAg produced by tryptic cleavage at pH 7 (see below) is analyzed in the same way, there is a gradual increase in the amount of dimer (data not shown). These dimers originate from covalent intermolecular disulfide linkage formation, presumably involving Cys-48 or Cys-61 (Zheng *et al.*, 1992).

HBcAg capsids are also resistant to many proteases although this situation can be dramatically reversed by removal of nucleic acid. Nucleic acid appears to restrict access to the C-terminal cleavage site, as capsids depleted of nucleic acid are exquisitely sensitive to trypsin which rapidly cleaves [within minutes using 1% (w/w) enzyme] between arginyl residues 150 and 151. This *in vitro* cleavage product is one residue longer (Arg-150) than HBeAg isolated from human serum (Takahashi *et al.*, 1983). It has been shown that serum HBeAg is derived from a second translation product, namely, "precore protein", which is HBcAg plus a 29 amino acid N-terminal extension (Strandring *et al.*, 1988). The extra amino acids contain a signal sequence that targets the precore into the secretory pathway. Nevertheless, we suggest that, *in vivo*, the primary processing in the C-terminal region is by a serine protease, probably at the same site observed *in vitro*, and that arginine-150 is removed by a secondary cleavage, either specific or artifactual.

Our data indicate that both the polybasic C-termini and the nucleic acid are located at the inner surface of the capsid (see above). In recombinant HBcAg, nucleic acid can be partially processed with exogenous nuclease such as RNase and is sensitive to urea at concentrations that do not disrupt the capsid. One explanation is that the HBcAg capsid may be permeable, allowing access to proteases and the release of nucleic acid. HBcAg capsids appear to have holes of ~2 nm diameter through the capsid shell (Crowther *et al.*, 1994).

How does such a stable structure as HBcAg uncoat *in vivo*? The *in vitro* observations discussed above may cast some light on this paradox. We conjecture that the polybasic C-terminal region of HBcAg may be critical in this regard. Mechanisms such as phosphorylation of serine residues or nucleic acid processing may lead to the release of nucleic acid and/or the exposure of proteolytically sensitive sites in the C-terminus. Processing may occur at the inner surface of the capsid, or as discussed above, the C-terminus may translocate to a more exposed location. Whatever the exact *in vivo* scenario, we have seen that, *in vitro*, removal of the C-terminus compromises the stability of the capsid which may then be dissociated under relatively mild solvent conditions.

ACKNOWLEDGMENT

We thank Drs. M. Simon and J. Wall (Brookhaven National Laboratory) for provision of the STEM data. We would also like to thank Dr. Manoj Misra for making the initial observations on the HBV nucleocapsids by negative staining electron microscopy, Dr. Keith Rose (University of Geneva) for performing the C-terminal sequence analysis, and Mr. Ira Palmer for technical assistance.

REFERENCES

Burns, N. R., Craig, S., Lee, S. R., Richardson, M. H., Stenner, N., Adams, S. E., Kingsman, S. M., & Kingsman, A. J. (1992) *J. Mol. Biol. 216*, 207−211.

Burnstein, E. A., Vedenkima, N. S., & Ivkova, M. N. (1973) *Photochem. Photobiol. 18*, 263−279.

Casper, D., & Klug, A. (1962) *Cold Spring Harbor Symp. Quant. Biol. 217*, 1−24.

Chang, C. T., Wu, C.-S., & Yang, J. T. (1978) *Anal. Biochem. 91*, 13−31.

Choi, H.-K., Tong, L., Minor, W., Dumas, P., Boege, U., Rossmann, M. G., & Wengler, G. (1991) *Nature 354*, 37−43.

Cohn, E. J., & Edsall, J. T. (1943) *Proteins, Amino Acids and Peptides*, Van Nostrand-Reinhold, Princeton, NJ.

Crowther, R. A., Kiselev, N. A., Böttcher, B., Berriman, J. A., Borisova, G. P., Ose, V., & Pumpens, P. (1994) *Cell 77*, 943−950.

Duda, R. L., Wall, J. S., Hainfeld, J. F., Sweet, R. M., & Eiserling, F. A. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 5550−5554.

Gallina, A., Bonellei, F., Zentilin, L., Rindi, G., Muttini, M., & Milanesi, G. (1989) *J. Virol. 63*, 4645−4652.

Hildrich, C. M., Rogers, L. J., & Bishop, D. H. L. (1990) *J. Gen. Virol. 71*, 2755−2759.

Kikuchi, H., Goto, Y., & Hamaguchi, K. (1986) *Biochemistry 25*, 2009−2013.

Kronman, M. J., & Holmes, L. G. (1971) *Photochem. Photobiol. 14*, 113−134.

Ladenstein, R., Schneider, M., Huber, R., Bartunik, H.-D., Wilson, K., Schott, K., & Bracher, A. (1988) *J. Mol. Biol. 203*, 1045−1070.

Machida, A., Ohnuma, H., Tsuda, F., Yoshikawa, A., Hoshi, Y., Tanaka, T., Kishimoto, S., Akahane, Y., Miyakawa, Y., & Mayumi, M. (1991) *J. Virol. 65*, 6024−6030.

McCarty, K. S., Jr., Vollmer, R. T., & McCarty, K. S. (1974) *Anal. Biochem. 61*, 165−183.

Milich, D., McLachlan, A., Stahl, S., Wingfield, P., Thorton, G. B., Hughes, J. L., & Jones, J. E. (1988) *J. Immunol. 141*, 3617−3624.

Murray, K. (1987) *Proc. R. Soc. London B 230*, 107−146.

Nassal, M., & Schaller, H. (1993) *Trends Microbiol. 1*, 221−228.

Nassal, M., Rieger, A., & Steinau, O. (1992) *J. Mol. Biol. 225*, 1013−1025.

Okabe, N., Serikawa, M., & Tomita, K. (1977) *J. Biochem. (Tokyo) 81*, 871−877.

Onodera, S., Ohori, H., & Ishida, N. (1982) *J. Med. Virol. 10*, 147−155.

Pasek, M., Goto, T., Gilbert, W., Zink, B., Schaller, H., MacKay, P., Leadbetter, G., & Murray, K. (1979) *Nature (London) 282*, 575−579.

Perczel, A., Park, K., & Fasman, G. D. (1992) *Anal. Biochem. 203*, 83−93.
Rose, K., Simona, M. G., & Offord, R. E. (1983) *Biochem. J. 215*, 261−272.
Rossmann, M. G. (1984) *Virology 134*, 1−11.
Rossmann, M. G., & Johnson, J. E. (1989) *Annu. Rev. Biochem. 58*, 533−537.
Schachman, H. K. (1959) *Ultracentrifugation in Biochemistry*, Academic Press, New York.
Seifer, M., Zhou, S., & Standring, D. N. (1993) *J. Virol. 67*, 249−257.
Sreerama, N., & Woody, R. W. (1993) *Anal. Biochem. 209*, 32−44.
Stafford, W. F. (1992) *Anal. Biochem. 203*, 295−301.
Stahl, S. J., & Murray, K. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 6283−6287.
Stahl, S. J., MacKay, P., Magazin, M., Bruce, S. A., & Murray, K. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 1606−1610.
Steven, A. C., Hainfeld, J. F., Trus, B. L., Steinert, P. M., & Wall, J. S. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 6363−6367.
Steven, A. C., Bauer, A. C., Bisher, M. E., Robey, F. A., & Black, L. W. (1991) *J. Struct. Biol. 106*, 221−236.
Strandring, D. N., Ou, J-H., Masiarz, F. R., & Rutter, W. J. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 8405−8409.
Strickland, E. H. (1974) *Crit. Rev. Biochem. 2*, 113−175.
Stuart, D. (1993) *Curr. Opin. Struct. Biol. 3*, 167−174.
Sturtevant, J. M. (1987) *Annu. Rev. Phys. Chem. 38*, 463−488.
Suzuki, S. (1989) *EMBO J. 8*, 797−804.
Takahashi, K., Machida, A., Funatsu, G., Nomura, M., Usuda, S., Miyakawa, Y., & Mayumi, M. (1983) *J. Immunol. 130*, 2903−2907.
Thomas, D., Schultz, P., Steven, A. C., & Wall, J. S. (1994) *Biol. Cell 80*, 181−192.
Trus, B. L., Unser, M., Pun, T., & Steven, A. C. (1992) *Scanning Microsc. Suppl. 6*, 441−451.
Valegåard, K., Liljas, L., Fridborg, K., & Unge, T. (1990) *Nature 345*, 36−41.
Wall, J. S. (1979) in *Mass Measurements with the Electron Microscope. Introduction to Analytical Electron Microscopy* (Hren, J. J., Goldstein, J. I., & Joy, D. C., Eds.) pp 333−342, Plenum Publishing Corp., New York.
Wall, J. S., & Hainfeld, J. F. (1986) *Annu. Rev. Biophys. Biophys. Chem. 15*, 355−376.
Wall, J. S., Hainfeld, J. F., & Cheung, K. (1985) in *Proceedings of the 43rd Annual Meeting, EMSA*, pp 716−717, San Francisco Press, San Francisco.
Weaver, J., & Williams, R. W. (1990) *Biopolymers* 30, 393−598.
Wetlaufer, D. B. (1962) *Adv. Protein Chem. 17*, 303−390.
Williams, R. W. (1986) *Methods Enzymol. 130*, 311−331.
Wingfield, P. T., Payton, M., Travernir, J., Barnes, M., Shaw, A., Rose, K., Simona, M. G., Demcuk, S., Williamson, K., & Dayer, J.-M. (1986) *Eur. J. Biochem. 160*, 491−497.
Yamaguchi, M., Hirano, T., Sugahara, K., Mizokami, H., Araki, M., & Matsubara, K. (1988) *Eur. J. Cell. Biol. 47*, 138−143.
Yamaki, M., Ohori, H., Onodera, S., Ishida, N., & Maeda, H. (1982) *Biochim. Biophys. Acta 706*, 165−173.
Yeh, C.-T., Liaw, Y.-F., & Ou, J.-H. (1990) *J. Virol. 64*, 6141−6147.
Yoshikawa, A., Tanaka, T., Hoshi, Y., Kato, N., Tachibana, K., Iizuka, H., Machida, A., Okamoto, H., Yamasaki, M., Miyakawa, Y., & Mayumi, M. (1993) *J. Virol. 67*, 6064−6070.
Zheng, J., Schodel, F., & Peterson, D. L. (1992) *J. Biol. Chem. 267*, 9422−9429.
Zhou, S., & Strandring, D. N. (1991) *J. Virol. 65*, 3319−3330.
Zhou, S., & Strandring, D. N. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 10046−10050.

BI941775L