Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Crystallization and X-ray analysis of the *T* = 4 particle of hepatitis B capsid protein with an N-terminal extension

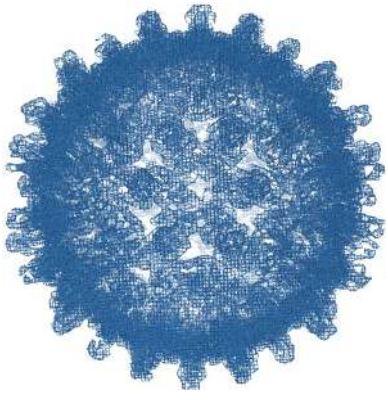
Wen Siang Tan,[a] Iain W. McNae,[b] Kok Lian Ho[b] and Malcolm D. Walkinshaw[b]*

[a]Department of Microbiology, Faculty of Biotechnology and Biomolecular Sciences, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia, and [b]Institute of Structural and Molecular Biology, School of Biological Sciences, University of Edinburgh, Michael Swann Building, King's Buildings, Mayfield Road, Edinburgh EH9 3JR, Scotland

Correspondence e-mail: m.walkinshaw@ed.ac.uk

Received 23 November 2006
Accepted 10 July 2007

PDB Reference: hepatitis B capsid protein, 2qij, r2qijsf.



Hepatitis B core (HBc) particles have been extensively exploited as carriers for foreign immunological epitopes in the development of multicomponent vaccines and diagnostic reagents. Crystals of the $T = 4$ HBc particle were grown in PEG 20 000, ammonium sulfate and various types of alcohols. A temperature jump from 277 or 283 to 290 K was found to enhance crystal growth. A crystal grown using MPD as a cryoprotectant diffracted X-rays to 7.7 Å resolution and data were collected to 99.6% completeness at 8.9 Å. The crystal belongs to space group $P2_12_12_1$, with unit-cell parameters $a = 352.3$, $b = 465.5$, $c = 645.0$ Å. The electron-density map reveals a protrusion that is consistent with the N-terminus extending out from the surface of the capsid. The structure presented here supports the idea that N-terminal insertions can be exploited in the development of diagnostic reagents, multicomponent vaccines and delivery vehicles into mammalian cells.

1. Introduction

Hepatitis B virus (HBV) poses a major public health problem worldwide, particularly in Southeast Asia and Africa, despite the presence of effective vaccines. Currently, there are more than 350 million chronic carriers worldwide, of which one million die each year (Jung & Pape, 2002). The virus has a partially double-stranded DNA of about 3.2 kbp which is protected by a capsid formed by multiple copies of a single core protein (HBcAg). The capsid is enveloped by a lipid bilayer containing three forms of related surface antigens (HBsAg) known as long (L), medium (M) and short (S) polypeptides (Heermann *et al.*, 1984).

HBcAg contains 183 or 185 residues with a C-terminal region of about 40 residues rich in arginine and is believed to interact with the viral genome (Pasek *et al.*, 1979). HBcAg can be readily synthesized in *Escherichia coli* (Burrell *et al.*, 1979), where it assembles into icosahedral core (HBc) particles containing 180 or 240 monomers with triangulation number $T = 3$ and $T = 4$, respectively (Crowther *et al.*, 1994). The yields of truncated HBcAg derivatives lacking the Arg-rich C-terminal region are substantially higher than that of the full-length protein (Stahl & Murray, 1989; Tan *et al.*, 2003).

For the past two decades, HBc particles have been extensively exploited as a carrier for foreign epitopes (for reviews, see Murray & Shiau, 1999; Pumpens & Grens, 2001). Hundreds of HBcAg derivatives containing insertions of foreign sequences either at the N- or C-terminus or in the major immunodominant region (positions 78–83) have been produced in *E. coli* and these also assemble to yield particles, demonstrating the potential of HBcAg as a molecular carrier in multicomponent vaccine development. An N-terminal insertion which replaced the first two residues of HBcAg with the N-terminal eight residues of β-galactosidase and a linker sequence of three amino acids provided the first example of such modified particles using an *E. coli* expression system (Stahl *et al.*, 1982). The three-dimensional structure of HBc particles at 3.3 Å deduced from

X-ray diffraction analysis (Wynne *et al.*, 1999) revealed that the N-terminus is located on the external surface of the capsid. Using cryoelectron microscopy and difference imaging at 11 Å, Conway *et al.* (1998) demonstrated that an extraneous octapeptide fused to the N-terminus is localized near the position at which the spike protrusion enters the contiguous shell. However, peptides fused to the N-terminal ends may affect the assembly of the capsid unless a flexible linker sequence is included immediately preceding the N-terminus of HBcAg (Clarke *et al.*, 1987; Murray & Shiau, 1999). The C-terminus of HBcAg is located in the interior of the capsid (Zlotnick *et al.*, 1997; Wynne *et al.*, 1999; Watts *et al.*, 2002), therefore fusion of polypeptide to this end is likely to be buried within the capsid and has a lower immunogenicity.

Here, we report the crystallization of the $T = 4$ capsid of truncated HBcAg harbouring an N-terminal extension of 11 residues. An initial low-resolution data set (8.9 Å) has been obtained from a single crystal and analysed. Subsequent structure solution clearly indicates the presence of the N-terminal extension and shows it to be positioned on the exterior surface of the particle.

2. Materials and methods

2.1. Plasmid and bacteria strain

Plasmid pR1-11E (Stewart, 1993) carrying the coding sequence for the truncated HBcAg (residues 3–148; Pasek *et al.*, 1979) preceded by the eight N-terminal residues of β-galactosidase (MTMITDSL) and a tripeptide linker (EFH) was obtained from Professor K. Murray (University of Edinburgh). The plasmid is under the control of the *tac* promoter and was introduced into *E. coli* strain W3110IQ (Bachmann, 1972). Fig. 1 shows the amino-acid sequence of the N-fusion truncated HBcAg (N-HBcAgΔ) produced by the plasmid pR1-11E, compared with that of strain HBcΔ-CW, the structure of which has been solved at 3.3 Å resolution (Wynne *et al.*, 1999).

2.2. Purification of HBcAg

Purification of $T = 4$ particles of HBcAg took place as described in Tan *et al.* (2003).

```
MTMITDSLEF   HIDPYKEFGA   TVELLSFLPS     30
*********M   D

DFFPSVRDLL   DTAAALYRDA   LESPEHCSPH     60
             S     E

HTALRQAILC   WGDLMTLATW   VGTNLEDPAS     90
             E           N

RDLVVSYVNT   NVGLKFRQLL   WFHISCLTFG    120
     N       M   I

RETVLEYLVS   FGVWIRTPPA   YRPPNAPILS    150


TLPETTV-                               157
       V
```

Figure 1
Amino-acid sequence comparison of N-fusion truncated HBcAg (N-HBcAgΔ) and HBcΔ-CW. The amino-acid sequence of N-HBcAgΔ is shown. The N-terminal extension is marked by asterisks. Amino acids 1–8, N-terminus of β-galactosidase; amino acids 9–11, linker; amino acids 12–157, HBcAg. The differences in amino-acid sequence in variant CW (HBcΔ-CW) are shown underneath the sequence.

2.3. Crystallization

Crystallization trials were performed by the hanging-drop vapour-diffusion method using Linbro-style 24-well plates and siliconized cover slips. The drops contained 3 µl of $T = 4$ particles of HBcAg and 3 µl mother liquor and were equilibrated against 1 ml mother liquor.

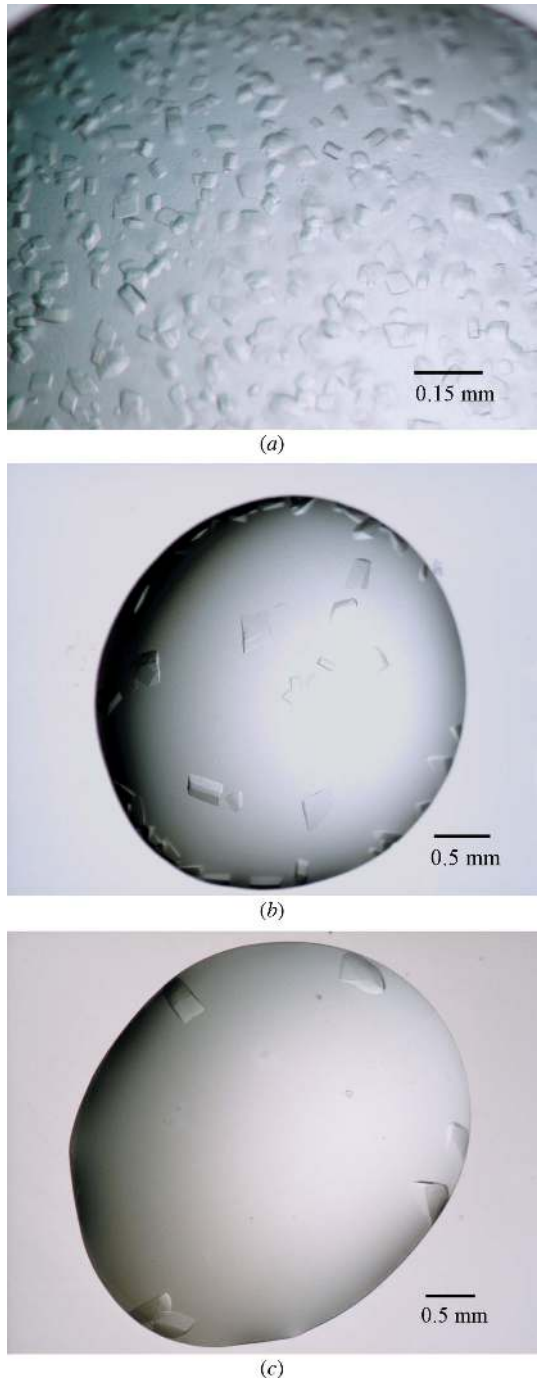


Figure 2
Crystals of $T = 4$ particles of N-HBcAgΔ. (*a*) Tiny rhombohedral crystals were grown by adding 3 µl protein solution (17.5 mg ml^{-1}) to an equal volume of mother liquor containing 0.1 *M* MES pH 6.5, 0.4 *M* ammonium sulfate, 4%(*w*/*v*) PEG 20 000 and 20%(*v*/*v*) butanediol at 294 K. (*b*) A temperature jump from 277 to 290 K produced larger crystals. The conditions were as in (*a*) except that 10 mg ml^{-1} protein was used and the crystal drop was kept at 277 K for 5 d and then transferred to 290 K. (*c*) Crystals grown under the same conditions as (*b*), except that butanediol was replaced with 12% MPD.

Table 1
Effects of alcohol on the growth of crystals.

Alcohol	Crystal morphology	Diffraction resolution (Å), source
MPD	Crystals grew in 11–14%(*v*/*v*) MPD and are well faceted.	7.7, SRS and ESRF
(2*R*,3*R*)-(−)-2,3-Butanediol	Crystals grew in 15–25%(*v*/*v*) butanediol. Well faceted crystals grew in 15–20%(*v*/*v*) butanediol.	8, SRS and ESRF
(±)-2,3-Butanediol	Crystals were observed in 15–25%(*v*/*v*) butanediol. Well faceted crystals grew in 15–20%(*v*/*v*) butanediol.	8, SRS and ESRF
Glycerol	No crystals were observed in the presence of glycerol.	No data were collected
Propan-2-ol	Heavy precipitation from 5–50%(*v*/*v*) propan-2-ol.	No data were collected
Ethanol	Crystals grew in 10–25%(*v*/*v*) ethanol; well faceted crystals were only observed in 10%(*v*/*v*) ethanol.	No diffraction observed at SRS
Methanol	Microcrystals grew in 10–30%(*v*/*v*) methanol and crystals grew in 40%(*v*/*v*) methanol. Not all were well faceted.	No data were collected

Crystallization conditions were initially screened based on the conditions used for HBcΔ-CW: 0.1 *M* MES pH 6.5, 0.4 *M* ammonium sulfate, 4%(*w*/*v*) PEG 20 000 and 20%(*v*/*v*) butanediol (Wynne *et al.*, 1999). The conditions were refined further by altering the protein concentration (1–22 mg ml^{-1}), pH (6.0–7.0), ammonium sulfate concentration, PEG 20 000 concentration and temperature (277, 283, 290 and 293 K). The growth of crystals in the presence of cryoprotectants [ethanol, propan-2-ol, methanol, (2*R*,3*R*)-(−)-2,3-butanediol (Aldrich), (±)-2,3-butanediol (mixture of isomers; Aldrich) and 2-methylpentane-2,4-diol (MPD, Fisher Scientific)] was studied.

2.4. Data collection and structure analysis

Crystals were mounted in cryo-loops (Hampton Research) and flash-frozen in liquid nitrogen; the crystallization solution acted as a cryoprotectant. Data collection was performed at 100 K at station 14.1, SRS Daresbury; the wavelength was 1.488 Å. A total of 100.5° of data were recorded with an oscillation angle of 0.3°. Processing was performed using the programs *MOSFLM* and *SCALA* (Leslie, 1992). A self-rotation function was performed using the program *MOLREP* (Vagin & Teplyakov, 1997). The initial structure was solved using the program *Phaser* (McCoy *et al.*, 2005) and the previously solved HBc particle (PDB code 1qgt) was used as the search model; the entire particle generated from these coordinates was used as the search model. Initial phases were obtained using *Phaser* and density modification was performed using the program *DM* (Cowtan, 1994) utilizing 60-fold icosahedral averaging. Analysis was performed using the program *Coot* (Emsley & Cowtan, 2004). All programs were part of the *CCP*4 suite of programs (Collaborative Computational Project, Number 4, 1994).

3. Results and discussion

The conditions used to produce the plate-like crystals of HBcΔ-CW protein described by Wynne *et al.* (1999) gave rise to tiny rhombohedral crystals of N-HBcAgΔ with approximate dimensions 0.05 × 0.05 × 0.03 mm (Fig. 2*a*). These crystals grew in approximately 3 d, but no diffraction was observed using a rotating-anode X-ray source. In order to increase the size of the crystals, the temperature and protein concentration were refined. The same size and shape of crystals were observed at 290 K with 15–22 mg ml^{-1} protein, but the size of crystals increased significantly (0.1 × 0.1 × 0.06 mm) at a lower protein concentration (10–12.5 mg ml^{-1}). Precipitation appeared at 277 and 283 K but no crystals formed. Surprisingly, when the crystallization trays were transferred to 290 K, the precipitates dissolved and crystals grew within 2 d. Different shapes of crystals formed (Fig. 2*b*) at protein concentrations between 7.5 and 22 mg ml^{-1}, with the largest crystals produced between 7.5 and 10 mg ml^{-1}. These crystals grew to maximum dimensions of 0.5 × 0.5 × 0.3 mm and remained stable for at least three months. The frozen

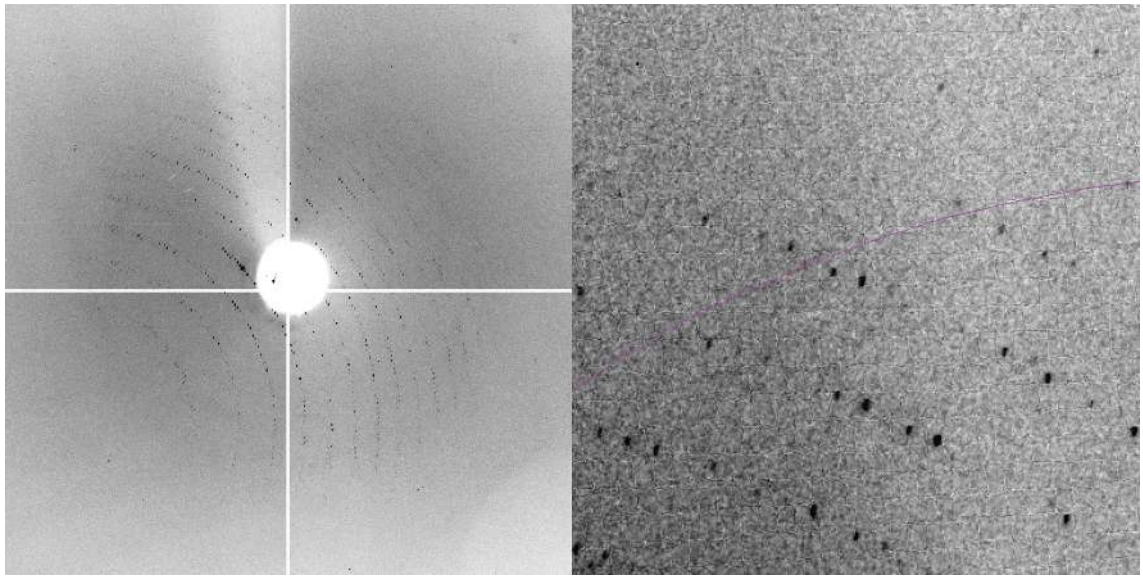


Figure 3
Diffraction image and a close-up of the highest resolution reflections. The purple line indicates 8.9 Å resolution.

crystals were found to diffract X-rays to about 8 Å at the SRS, Daresbury.

Attempts to improve resolution were carried out by soaking crystals in the mother liquor containing cryoprotectants [30%(*w*/*v*) glucose, 30%(*w*/*v*) sucrose, 30%(*w*/*v*) glycerol, 15%(*w*/*v*) PEG 4000, 20, 30 and 50%(*v*/*v*) butanediol, 20 and 50%(*v*/*v*) MPD] and oils (sunflower oil, mineral oil, paraffin oil, silicon oil and baby oil). The crystals appeared to be stable in these soaking solutions, but they diffracted X-rays poorly. As an alternative approach, crystallization trays were set up in the presence of different alcohols, kept at 277 K for 5 d and transferred to 290 K. Table 1 summarizes the effects of alcohol on the growth of crystals.

Well faceted crystals grown in 15–25%(*v*/*v*) butanediol diffracted X-rays to about 8 Å at the SRS. Replacement of the butanediol with 11–14%(*v*/*v*) MPD also produced well faceted crystals (Fig. 2*c*). Data were obtained from a single crystal grown in MPD and initial diffraction extended to a resolution of about 7.9 Å (Fig. 3). Radiation damage was clear early in data collection and the final data were obtained to a resolution of 8.9 Å. The crystal mosaicity was 0.2° and is similar to that observed in other virus crystals. Statistics of data collection and merging are shown in Table 2. The crystal was found to be orthorhombic, with unit-cell parameters a = 352.3, b = 465.5, c = 645.0 Å, and differs from those reported by Wynne *et al.* (1999) and Zlotnick *et al.* (1999), which had $C2$ symmetry. Clear systematic absences for $h00$ and $0k0$ axes were present; however, owing to the weak nature of the data the presence of systematic absences for the $00l$ axis was ambiguous. As a consequence, an attempt was made to solve the structure in both the $P2_12_12_1$ and $P2_12_12$ space groups. The Matthews coefficient for a whole core particle within the asymmetric unit is 7.8 Å^3 Da^{-1}, with a solvent content of 84%. This is similar to

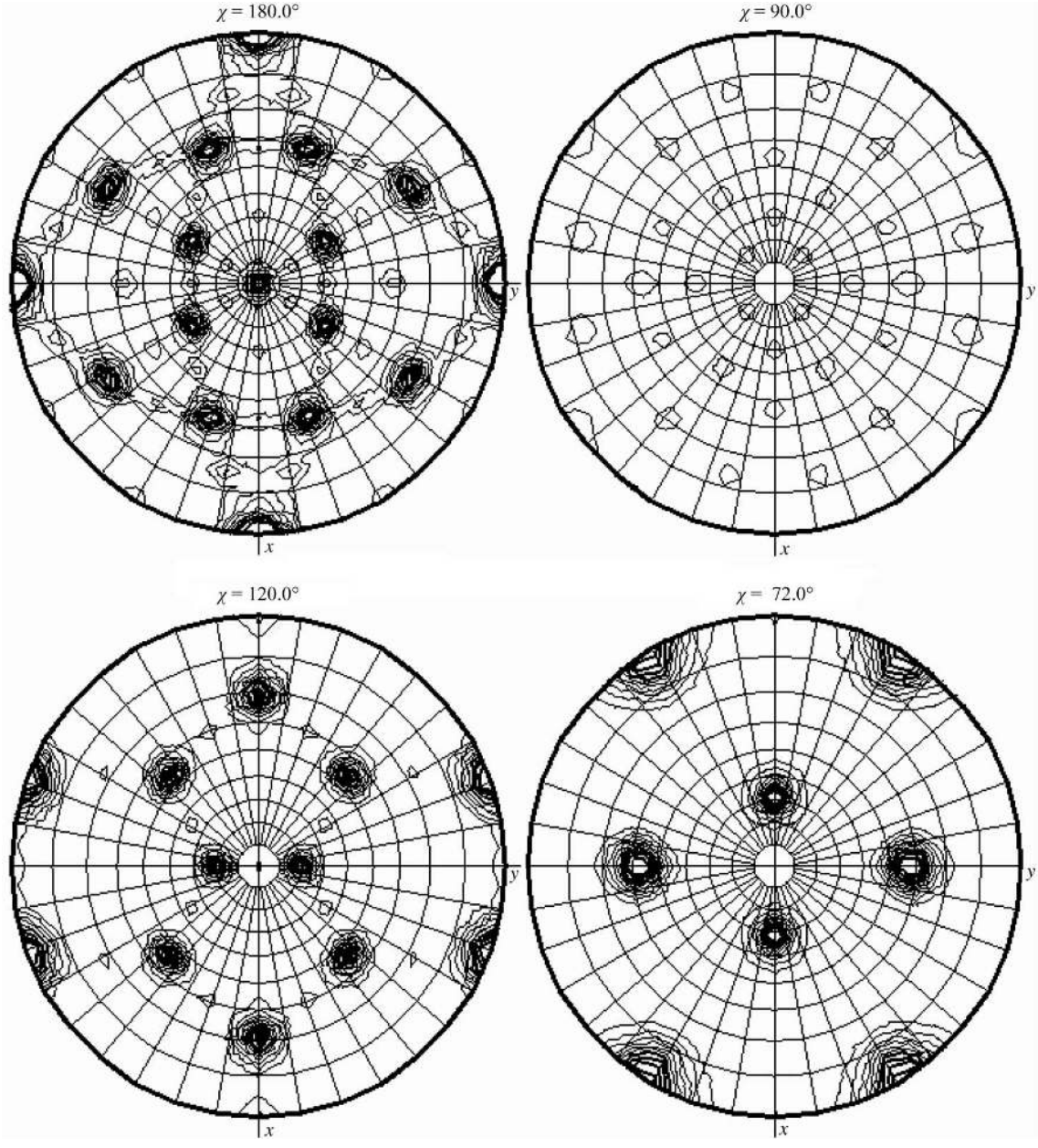


Figure 4
Self-rotation function *versus* (θ, φ) for χ = 180, 90, 120 and 72°. All data to 9 Å were used with a search radius of 140°. x = crystal a axis, y = crystal b axis, z (perpendicular to plane of page) = crystal c axis.

Table 2
Data-collection statistics.

Values in parentheses are for the highest resolution shell and values in square brackets are for the lowest resolution shell.

Resolution limits (Å)	60.41–8.9 (9.38–8.9) [60.41–28.14]
$R_{\rm merge}$†	21.9 (50.2) [5.4]
Total observations	322540 (45331) [8061]
Unique observations	81623 (11817) [2041]
$\langle I/\sigma(I)\rangle$	7.1 (2.8) [25.7]
Completeness (%)	99.6 (99.9) [87.7]
Multiplicity	4.0 (3.8) [3.3]

† $R_{\rm merge} = \sum_{\mathbf{h}}\sum_{l}|I_{\mathbf{h}l} - \langle I_{\mathbf{h}}\rangle|/\sum_{\mathbf{h}}\sum_{l}\langle I_{\mathbf{h}}\rangle$, where $I_{\mathbf{h}l}$ is the lth observation of reflection **h** and $\langle I_{\mathbf{h}}\rangle$ is the weighted average intensity for all observations l of reflection **h**.

the 82% solvent reported for the previously solved HBc particle crystal structure (Wynne *et al.*, 1999). A self-rotation map clearly shows twofold, threefold and fivefold axes consistent with icosahedral symmetry (Fig. 4), with the icosahedral twofold running parallel or coincident to the *c* axis. In $P2_12_12$ the twofold crystallographic axis may run through the icosahedral twofold and a molecular-replacement search was performed using a half a particle created from PDB entry 1qgt. The program *Phaser* was used to a resolution of 14 Å, which resulted in a top solution for $P2_12_12$ with a log-likelihood gain (LLG) value of 419 and a *Z* score of 14.8, whilst in $P2_12_12_1$ the top solution had an LLG value of 512 and a *Z* score of 23.6, clearly indicating the correct space group to be $P2_12_12_1$. The complete particle was subsequently positioned in the cell using the program *Phaser*, with the top solution giving a refined LLG of 9947 and a *Z* score of 36 for all data to 8.9 Å. This solution is consistent with the peaks found in the self-rotation map (Fig. 4), with the icosahedral twofold running parallel to the crystal *c* axis. Initial density modification was performed using the whole particle as a single unit (Cowtan, 1994) and strict 60-fold icosahedral noncrystallographic symmetry averaging was subsequently applied to the density modification. Further refinement awaits higher resolution data.

The packing within the crystal shows that all of the crystal contacts are made between the exterior spikes of adjacent particles. The NCS-refined electron density is clear and shows that the vast majority of the protein model is covered by density, with very little extraneous density (Fig. 5). Some disorder is present on some of the crystal contacting spikes and may indicate a degree of structural change in these regions. This change was not detected in the cryo-EM structures of HBV capsids that carry an extraneous peptide at the N-terminus of the monomer (Böttcher *et al.*, 1998; Conway *et al.*, 1998). Three of the amino-acid differences between HBcAgΔ and HBc-ΔCW are found on the capsid spikes and are the most likely cause of this change. In particular, the HBcAg protein contains a Thr at position 83 (Asn in HBc-ΔCW). This is located near the top of the spike and may result in changes in crystal packing.

Despite the low resolution of the crystal structure obtained in this study, extra density in the regions around the N-terminus of each monomer within the icosahedral asymmetric unit was clearly visualized. In these regions there is consistently unaccounted-for density that protrudes out from the surface of the particle. This suggests that the extra N-terminal extension protrudes out from the surface of the particle in a reasonably ordered fashion. Modelling of the extension reveals that all of the residues can be comfortably positioned within the density (Fig. 6). The N-terminal Met was not modelled as this has previously been shown to be absent (Tan *et al.*, 2003). This is in agreement with the previous electron cryomicroscopy results of Böttcher *et al.* (1998) and Conway *et al.* (1998). The position of the modelled extension closely resembles that observed by Conway *et al.* (1998).

It is now widely accepted that recombinant HBV capsid has a significant potential as a vaccine carrier (Murray & Shiau, 1999; Pumpens & Grens, 2001). The structures of HBcAg revealed by X-ray crystallography (Wynne *et al.*, 1999) and electron cryomicroscopy (Crowther *et al.*, 1994; Böttcher *et al.*, 1997; Conway *et al.*, 1998) offer an excellent guide for inserting foreign sequences, particularly at the N- and C-termini of the HBcAg monomer as well as at the tip of the spike formed by a dimer. Fusion of foreign peptides to the N-terminus of HBcAg has significant applications in vaccine development and gene delivery. For instance, fusion of a 23-residue peptide of the external domain of human influenza A M2 protein to the N-terminus of HBcAg conferred 100% protection against a lethal virus challenge in mice (Neirynck *et al.*, 1999). Recently, Brandenburg *et al.* (2005) demonstrated that a translocation motif (TLM)

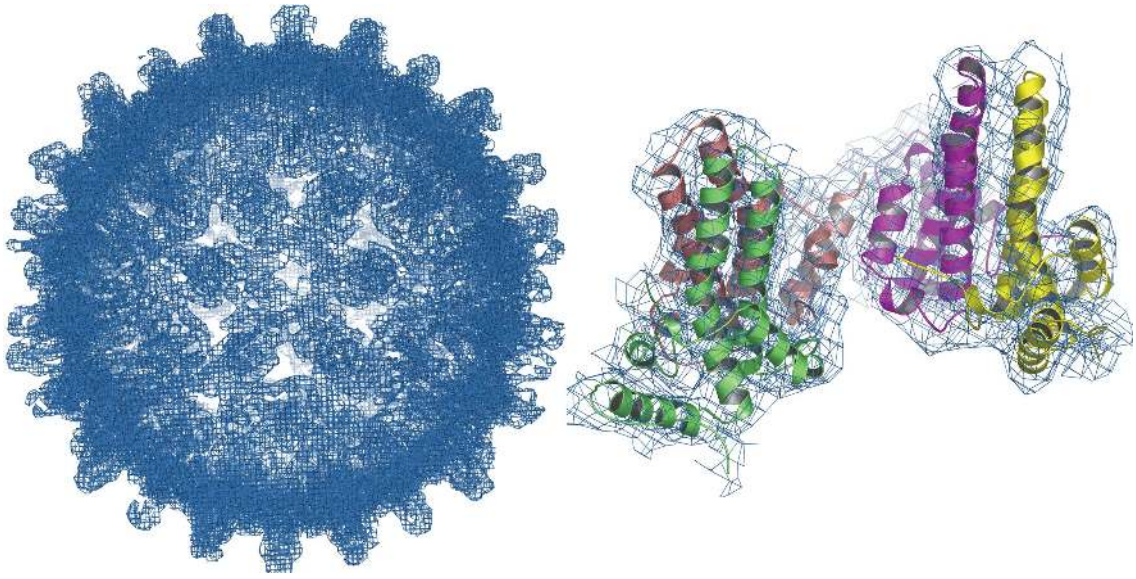


Figure 5
Electron density around a complete core particle approximately looking down the icosahedral twofold (pseudo-sixfold) and a close up of a tetramer; the protein is shown as cartoon ribbons in the close-up view. Density is contoured at 1σ.

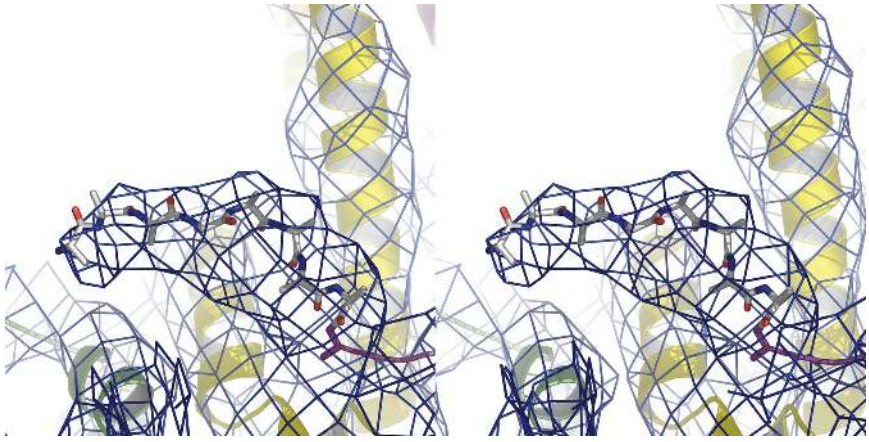


Figure 6
Stereo image showing extra electron density near the N-terminus. Density is shown for one monomer of the icosahedral asymmetric unit but is representative of density for the other three monomers. Modelled residues (shown as polyalanine) are shown to fit within the density. Density is contoured at 1σ.

comprising 12 amino acids fused to the N-terminus of HBcAg mediated efficient transfer of assembled chimeric particles and cargo into primary human hepatocytes. Our results support the view that the N-terminal fusion peptide is exposed on the exterior of the particle and this provides an explanation for the use of N-terminal fusion HBcAg as a vaccine carrier and a gene-delivery vehicle.

We thank Professor K. Murray for providing plasmid pR1-11E. X-ray data collection was supported by BM14UK/ESRF, Grenoble and SRS, Daresbury. WST was supported by a Die NorKen Stiftung Visiting Fellowship and KLH is the recipient of the Darwin Trust Scholarship.

References

Bachmann, B. J. (1972). *Bacteriol. Rev.* **36**, 525–557.
Böttcher, B., Dyson, M. R. & Crowther, R. A. (1998). *Proceedings of the 14th International Congress on Electron Microscopy*, edited by H. A. Calderon Benavides & M. J. Yacaman, Vol. 1, pp. 737–738. Bristol: Institute of Physics.
Böttcher, B., Wynne, S. A. & Crowther, R. A. (1997). *Nature (London)*, **386**, 88–91.
Brandenburg, B., Stock, L., Gutzeit, C., Roos, M., Lupberger, J., Schwartlander, R., Gelderblom, H., Sauer, I. M., Hofschneider, P. H. & Hildt, E. (2005). *Hepatology*, **42**, 1300–1309.
Burrell, C. J., MacKay, P., Greenaway, P. J., Hofschneider, P.-H. & Murray, K. (1979). *Nature (London)*, **279**, 43–47.
Clarke, B. E., Newton, S. E., Carroll, A. R., Francis, M. J., Appleyard, G., Syred, A. D., Highfield, P. E., Rowlands, D. J. & Brown, F. (1987). *Nature (London)*, **330**, 381–384.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Conway, J. F., Cheng, N., Zlotnick, A., Stahl, S. J., Wingfield, P. T. & Steven, A. C. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 14622–14627.
Cowtan, K. (1994). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **31**, 34–38.
Crowther, R. A., Kiselev, N. A., Böttcher, B., Berriman, J. A., Borisova, G. P., Ose, V. & Pumpens, P. (1994). *Cell*, **77**, 943–950.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Heermann, K. H., Goldmann, U., Schwartz, W., Seyffarth, T., Baumgarten, H. & Gerlich, W. H. (1984). *J. Virol.* **52**, 396–402.
Jung, M.-C. & Pape, G. R. (2002). *Lancet*, **2**, 43–50.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.
McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2005). *Acta Cryst.* D**61**, 458–464.
Murray, K. & Shiau, A. L. (1999). *Biol. Chem.* **380**, 277–283.
Neirynck, S., Deroo, T., Saelens, X., Vanlandschoot, P., Jou, W. M. & Fiers, W. (1999). *Nature Med.* **5**, 1157–1163.
Pasek, M., Goto, T., Gilbert, W., Zink, B., Schaller, H., MacKay, P., Leadbetter, G. & Murray, K. (1979). *Nature (London)*, **282**, 575–579.
Pumpens, P. & Grens, E. (2001). *Intervirology*, **44**, 98–114.
Stahl, S., MacKay, P., Magazin, M., Bruce, S. A. & Murray, K. (1982). *Proc. Natl Acad. Sci. USA*, **79**, 1606–1610.
Stahl, S. J. & Murray, K. (1989). *Proc. Natl Acad. Sci. USA*, **86**, 6283–6287.
Stewart, F. S. (1993). PhD thesis. University of Edinburgh, Scotland.
Tan, W. S., Dyson, M. R. & Murray, K. (2003). *Biol. Chem.* **384**, 363–371.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
Watts, N. R., Conway, J. F., Cheng, N., Stahl, S. J., Belnap, D. M., Steven, A. C. & Wingfield, P. T. (2002). *EMBO J.* **21**, 876–884.
Wynne, S. A., Crowther, R. A. & Leslie, A. G. (1999). *Mol. Cell*, **3**, 771–780.
Zlotnick, A., Cheng, N., Stahl, S. J., Conway, J. F., Steven, A. C. & Wingfield, P. T. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 9556–9561.
Zlotnick, A., Palmer, I., Kaufman, J. D., Stahl, S. J., Steven, A. C. & Wingfield, P. T. (1999). *Acta Cryst.* D**55**, 717–720.