Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary analysis of a dsDNA bacteriophage capsid intermediate: Prohead II of HK97

William R. Wikoff,[a,b] Zhiwei Che,[a] Robert L. Duda,[c] Roger W. Hendrix[c] and John E. Johnson[a]*

[a]Department of Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA, [b]Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St Louis, MO 63110, USA, and [c]Department of Biological Sciences, University of Pittsburgh, Pittsburgh, PA 15260, USA

Correspondence e-mail: jackj@scripps.edu

Received 1 September 2003
Accepted 7 October 2003

HK97 Prohead II is an early intermediate in the maturation of HK97, a $T = 7$ dsDNA-tailed bacteriophage related to bacteriophage λ. Previously, selected capsid-protein genes of HK97 were expressed in *Escherichia coli* and spontaneously assembled to form an icosahedral capsid that followed a maturation pathway closely similar to the authentic virion. The crystal structure of the mature HK97 capsid (Head II) made in this way was reported at 3.5 Å resolution. Additional high-resolution structures of intermediates are needed to understand the maturation mechanism. The crystal structure of expressed Prohead II will elucidate the early steps of HK97 assembly. Crystals of the Prohead II mutant W336F were grown in 0.1 *M* HEPES pH 7.5, 0.2 *M* $CaCl_2$ and 2–3% PEG 4000 at a Prohead II concentration of 16.5 mg ml^{-1}. It was not possible to grow high-quality crystals of wild-type Prohead II. Diffraction was observed to 5 Å resolution from these crystals on beamline 14BM-C at the Advanced Photon Source and data were collected to 5.5 Å with a completeness of 77%. The space group was $P2_13$, with unit-cell parameter $a = 707.0$ Å and four particles in the unit cell. The particles are on the body diagonals of the cubic cell, with icosahedral threefold axes coincident with crystallographic threefold axes. Self-rotation function and locked-rotation function analysis determined the particle orientation and a one-dimensional *R*-factor search along the body diagonal indicated that the particle centers were close to (1/4, 1/4, 1/4) and symmetry-related positions. Molecular-replacement averaging and phase extension are under way.

1. Introduction

The double-stranded DNA bacteriophage HK97 has a 40 kbp genome and a long non-contractile tail and is similar to bacteriophage λ. Previously, selected capsid-protein genes of HK97 were expressed in *Escherichia coli* and spontaneously assembled to form an icosahedral capsid that followed a maturation pathway closely similar to the authentic virion (Duda, Hempel *et al.*, 1995; Duda, Martincic *et al.*, 1995). The crystallography described in this report was performed with these expressed particles, as were the cryoEM studies that defined the structural transitions associated with HK97 maturation. The expressed particle has a $T = 7$ icosahedral capsid, with 420 copies of the capsid protein (gp5) organized as pentamers and hexamers. The authentic virion has a unique pentamer replaced by the dodecameric portal protein, to which proteins associated with DNA packaging initially attach and, later, the assembled tail.

There are multiple intermediates in the assembly of HK97 before the mature capsid form (Head II) is reached (Conway

et al., 1995; Lata *et al.*, 2000). Cryo-electron microscopy (cryoEM) image reconstructions of each assembly form were determined (Conway *et al.*, 1995), outlining a pathway for the maturation. This process starts with the assembly of 420 copies of the intact capsid protein (gp5) into a $T = 7$ icosahedron called Prohead I. About 50 copies of the protease (gp4) are also packaged into the shell. Prohead I is converted to Prohead II by cleavage of 102 amino-acid residues from the amino-terminus of each molecule of gp5; the resultant fragments and the auto-digested protease (gp4) exit the capsid. The external morphologies of Prohead I and II are indistinguishable, both having a diameter of approximately 500 Å. *In vivo*, DNA genome packaging follows and the capsid goes through an expansion to Head I with a significantly larger size (~650 Å in diameter) and a more angular shape. The expressed Prohead II can be expanded in a similar fashion with a variety of partially denaturing solvents *in vitro* (Duda, Hempel *et al.*, 1995). Finally, Head I matures to Head II by completing the covalent cross-linking between lysine and asparagine residues through the formation of an isopeptide bond. The cross-linking, an unusual feature in virus maturation, stabilizes the entire capsid by linking the capsomers into concatenated protein rings (Wikoff *et al.*, 2000; Duda, 1998).

The empty mature capsid of HK97 (Head II) was crystallized (Wikoff *et al.*, 1998) and the structure was determined to 3.5 Å resolution (Wikoff *et al.*, 2000). The mature capsid structure showed the novel covalent cross-linking mechanism and viral chain-mail formation; however, structures of other intermediate forms are needed to reveal the entire viral maturation mechanism. We previously reported a 12 Å resolution structure of Prohead II determined by cryoEM and generated a pseudo-atomic model by fitting subunits from Head II into this density (Conway *et al.*, 2001). This provided a sense of the dramatic reorganization of the subunits that occurs during the maturation, but it was not possible to establish with precision the details of atom locations in Prohead II. The crystal structure of Prohead II will allow a detailed understanding of conformational changes in the maturation from Prohead II to Head II.

Here, we report the crystallization and crystallographic analysis of HK97 Prohead II W336F mutant. Conditions for the maturation of this mutant particle are considerably more vigorous than for the wild type, making it an ideal subject for crystallographic study. This is the first procapsid from a dsDNA virus to be crystallized. The crystals diffract to 5 Å and a structure determination to 5.5 Å resolution is under way. Although the current resolution is moderate, the electron density will be interpreted with the subunit atomic models derived from the Head II structure and we anticipate that a wealth of information on virus-particle dynamics and subunit trajectories in large-scale transitions will be generated from this study.

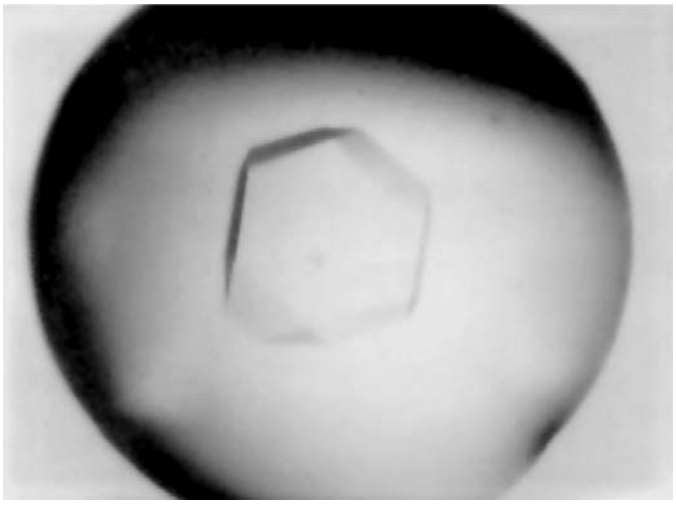

Figure 1
HK97 Prohead II W336F mutant crystal. The crystallization conditions were 0.1 *M* HEPES pH 7.5, 0.2 *M* $CaCl_2$, 2–3% PEG 4000 with a sample concentration of 26.5 mg ml^{-1}. The average crystal dimensions are 1.0 × 1.0 × 0.5 mm.

2. Materials and methods

2.1. Expression and purification

Site-directed mutagenesis to introduce the W336F mutation was followed by coexpression of the putative protease (gp4) and the capsid protein (gp5) in an *E. coli* T7 expression system. The Prohead II mutant was purified following the method for wild-type Prohead II, with modifications (Duda, Hempel *et al.*, 1995; Duda, Martincic *et al.*, 1995; Duda, 1998). The purified Prohead II sample was stored in 20 m*M* Tris–HCl pH 7.5 with 100 m*M* KCl and 1 m*M* 2-mercaptoethanol. Protein concentration was determined with the BCA assay (Pierce).

2.2. Crystallization, data collection and processing

Because HK97 procapsids spontaneously mature to capsids *in vitro* at low pH, crystallization conditions were limited to pH values above 6.0. Initial trials used hanging-drop vapor diffusion (McPherson, 1982) with a subset of conditions chosen from the Hampton Crystal screen in the pH range 6.0–8.5. The drop size was 3 µl, formed by mixing equal volumes of sample and mother liquor. The crystals were washed, crushed and checked using a Philips CM 100 electron microscope (Philips Electron Optics, Eindhoven, The Netherlands) after staining with uranyl acetate. The particle images on the micrographs were compared with those taken from the Prohead II W336F sample to ensure that the crystallization conditions did not induce the particles to mature to Head II.

Diffraction data were collected with a Quantum 4 CCD camera by oscillation photography at 277 K at the Advanced Photon Source (APS) BioCARS beamline 14BM-C (λ = 1.000 Å). The crystal-to-detector distance was 550 mm, with an oscillation angle of 0.3° and an exposure time of 10 s. 10–20 exposures were taken for each crystal from a total of 50 crystals.

Data were processed and scaled using a similar procedure to that used for HK97 Head II (Wikoff *et al.*, 2000). The diffraction patterns were viewed and indexed with the

programs *XDISPLAYF* and *DENZO* (Otwinowski & Minor, 1997). Each image was indexed and processed independently. The resolution of each image was determined by processing to 5.5 Å resolution and plotting the mean reflection $I/\sigma(I)$ as a function of resolution. The resolution at which the mean $I/\sigma(I)$ was less than 2 was used as the resolution limit to which the image was reprocessed. The procedure assures that only reflections of significant intensity are included at the highest resolution range. Reflections were scaled and post-refined with the program *SCALEPACK*, with the 'no merge, include partials' option (Otwinowski & Minor, 1997). Each image was aligned to a standard orientation by switching the crystallographic *a* and *b* axes when necessary. Mosaicity was refined independently for each crystal. Unmerged data from *SCALEPACK* were converted to *CCP*4 format with the program *ROTAPREP* and scaled by *AGROVATA*.

2.3. Particle orientation and position

Particle orientation was determined by computing the self-rotation function (Rossmann & Blow, 1962) and the locked self-rotation function in the resolution range 12–10 Å with the program *GLRF* (Tong & Rossmann, 1990). Subsequently, a one-dimensional *R*-factor search using a molecular-replacement model with data in the resolution range 50–15 Å was carried out from the origin to the center of the unit cell along the body diagonal using the programs *SFALL* and *RSTATS* from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994) to locate the center of the particle.

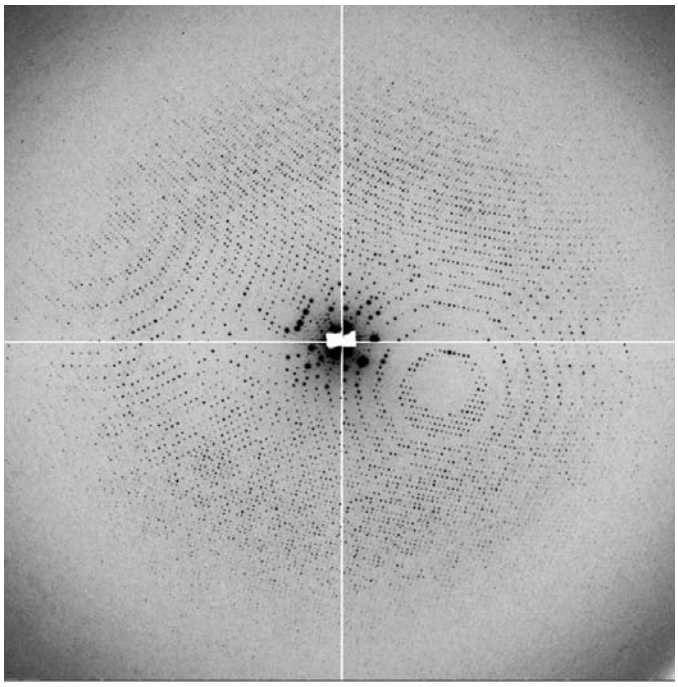

Figure 2
Diffraction pattern of HK97 Prohead II W336F mutant crystal. The data were collected at Advanced Photon Source (APS) BioCARS station 14BM-C with a Quanta 4 CCD camera (wavelength = 1 Å). The upper edge of the frame is at ∼5 Å resolution. The oscillation angle was 0.3° and the crystal-to-detector distance 550 mm with an exposure time of 10 s.

The search, initially coarse at 0.005 in fractional coordinates, was refined with the program *X-PLOR* (Brünger, 1992) at an interval of 0.001 in fractional coordinates. A Prohead II pseudo-atomic model was used for the search, built by fitting the coordinates of the HK97 Head II capsid monomer (Wikoff *et al.*, 2000) into the cryoEM map of Prohead II (Conway *et al.*, 2001; PDB code 1if0).

3. Results and discussion

3.1. Crystallization

Crystals were observed with various buffers in the pH range 6.0–8.5, all with either 0.2 *M* $CaCl_2$ or $MgCl_2$. The best crystals grew in 0.1 *M* HEPES pH 7.5, 0.2 *M* $CaCl_2$ and 2–3% PEG 4000, with the sample concentration of Prohead II W336F at 26.5 mg ml^{-1} (Fig. 1). The crystals typically appeared after 3–4 d and grew for an additional week to average dimensions of 1 × 1 × 0.5 mm.

Since Head II is significantly larger than the Prohead II particle (650 *versus* 500 Å), crystalline particles were checked by electron microscopy to ensure that the crystallization conditions did not induce maturation. A sample of crystals that had been washed, crushed and dissolved and a sample of crystal mother liquor were negatively stained with 1% uranyl acetate on carbon-film grids. Micrographs were recorded at 73 000× magnification and the particle size was estimated to be approximately 500 Å (data not shown) as expected for Prohead II.

3.2. Data collection, processing and scaling

All data were collected at APS BioCARS beamline 14BM-C. The best crystals diffracted to 5 Å resolution (Fig. 2). The unit-cell parameters were $a = b = c = 707.0$ Å, $\alpha = \beta = \gamma = 90^\circ$ using the indexing program *DENZO* (Otwinowski & Minor, 1997), suggesting that the space group may be cubic. The cubic space group $P2_13$ was confirmed by scaling in the resolution range 196–5.5 Å assuming Laue group 23 (initial $R_{\rm merge}$ = 16.1% from *SCALEPACK*) and identifying systematic absences of the odd reflections along the *h*00 axis. A total of 472 images from 26 crystals were included in the final data set, which extended from 196 to 5.5 Å resolution. The post-refined mosaicity from *SCALEPACK* ranged from 0.031 to 0.163 for individual crystals. Partial reflections were included to a minimum partiality of 0.7. The overall completeness was 77.4%, with a completeness of 19% at the highest resolution bin, for a total of 374 885 reflections in the data set (Fig. 3). The overall $R_{\rm merge}$ was 13.9% from *AGROVATA*.

3.3. Particle orientation and position

The self-rotation function was computed with the program *GLRF* (Tong & Rossmann, 1990) to determine the orientation of icosahedral non-crystallographic symmetry elements. The fivefold rotation function (Fig. 4) was consistent with four particles in the unit cell, each with an icosahedral threefold axis coincident with a lattice threefold axis. The fivefold search calculated in the outer resolution range of the data set (6.0–

5.5 Å resolution) confirmed that the quality of the data extended to this resolution. The threefold search shows four large crystallographic peaks in the directions of the four body-diagonal axes, each coinciding with an icosahedral threefold axis (Fig. 5).

To search for all non-crystallographic symmetry elements simultaneously, a locked-rotation function was computed (Fig. 6). By aligning one of the icosahedral threefold axes with the body diagonal ($\varphi = 45$, $\psi = 54.74^\circ$), the particle was rotated in a 360° range around the body diagonal to search for the angle between its orientation and the standard orientation.

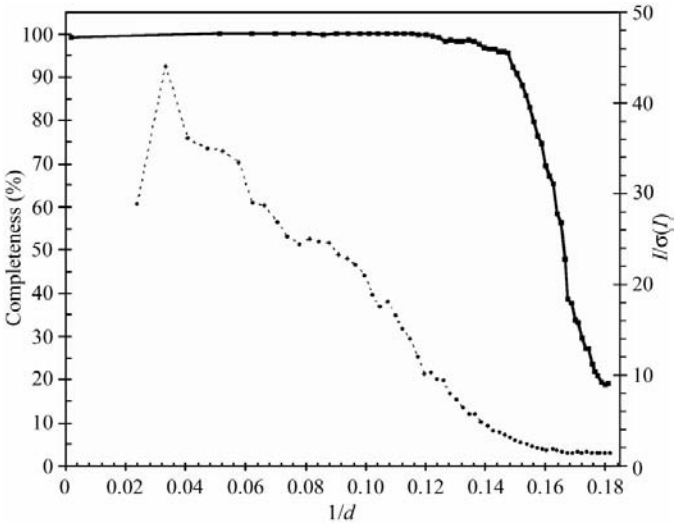

Figure 3
Plot of completeness (solid line) and $I/\sigma(I)$ (dashed line) *versus* resolution. The maximum resolution of the processed data is 5.5 Å.

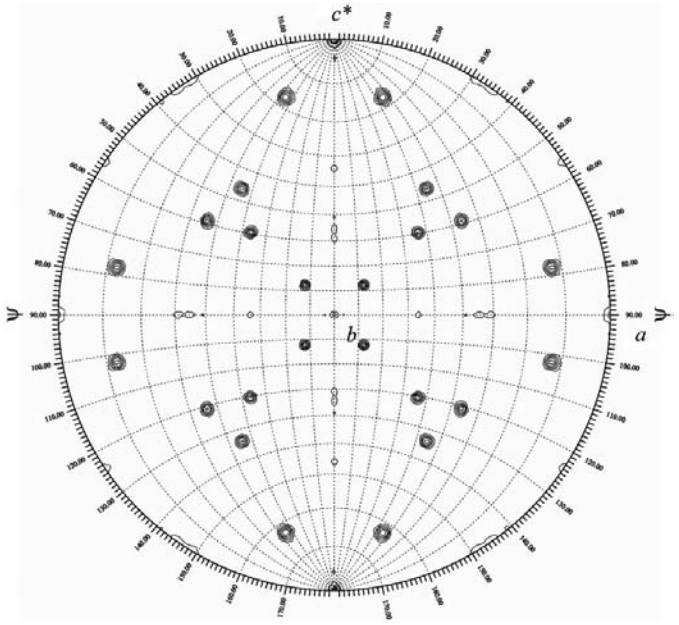

Figure 4
Fivefold self-rotation function peaks calculated in *GLRF* with data in the resolution range 12–10 Å. About 25 000 reflections were included and ~5000 large terms were used. The contour levels are between peak height 40 and 1000, with an increment of 20.

Three peaks were found related by the crystallographic threefold axis (120° apart) with a 97.7° (κ) angle for the first peak. At this orientation ($\varphi = 45$, $\psi = 54.74$, $\kappa = 97.7^\circ$), the skew matrix was determined to be

$$\begin{pmatrix} -0.174592200 & -0.707472900 & 0.684934910 \\ 0.968730900 & 0.000996000 & 0.248026900 \\ -0.176068100 & 0.706769600 & 0.685234830 \end{pmatrix}.$$

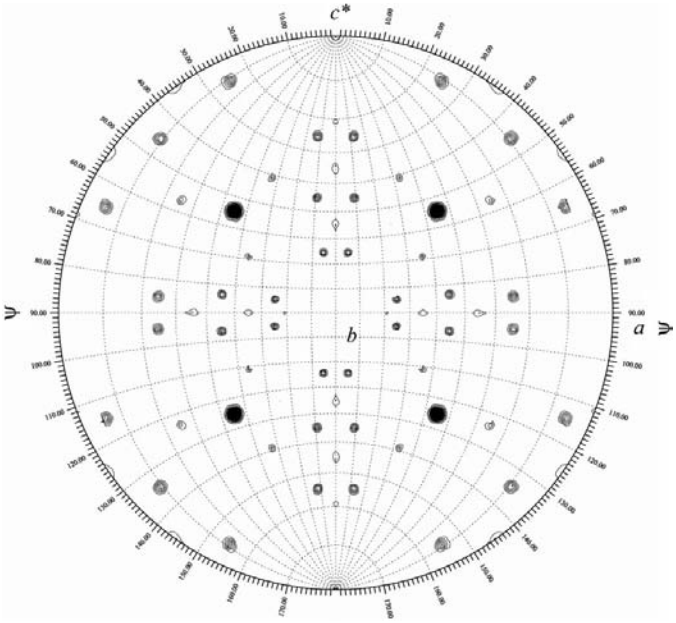

Figure 5
Threefold self-rotation function peaks calculated in *GLRF* with data in the resolution range 12–10 Å. The same conditions were used as for the fivefold self-rotation function. The four huge peaks are in the directions of four body diagonals of the cubic unit cell.

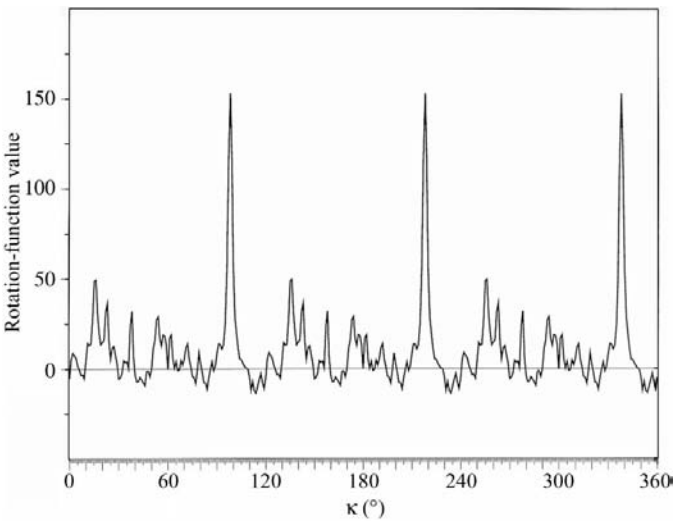

Figure 6
Locked self-rotation function peaks calculated in *GLRF* with data in the 12–10 Å resolution range by rotating the particle around the body diagonal of the unit cell. The first highest peak is at ($\varphi = 45$, $\psi = 54.74$, $\kappa = 97.7^\circ$). An obvious threefold relation is present in the three highest peaks.

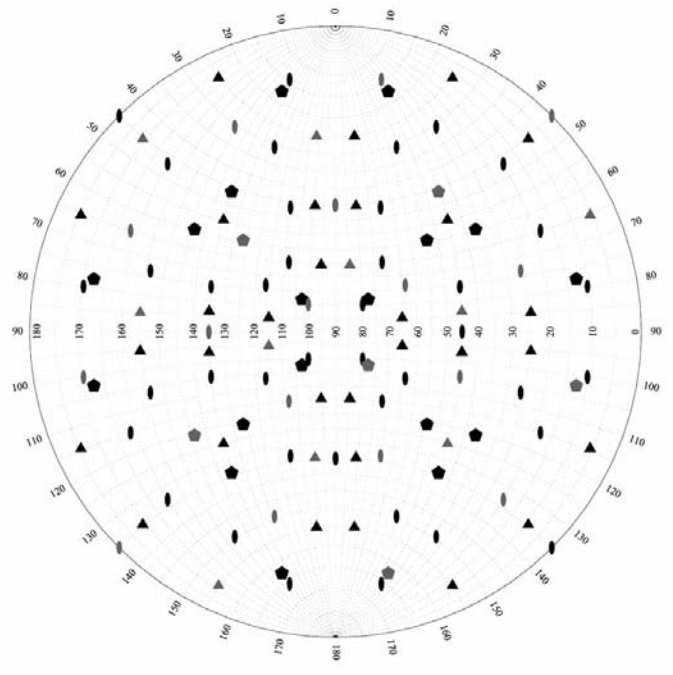

Figure 7
Constellations of icosahedral symmetry elements for the four particles in the unit cell viewed in the same orientation as the computed rotation functions. Symmetry elements for each of the four particles are shaded differently. Predicted positions agree well with self-rotation functions shown in Figs. 4 and 5.

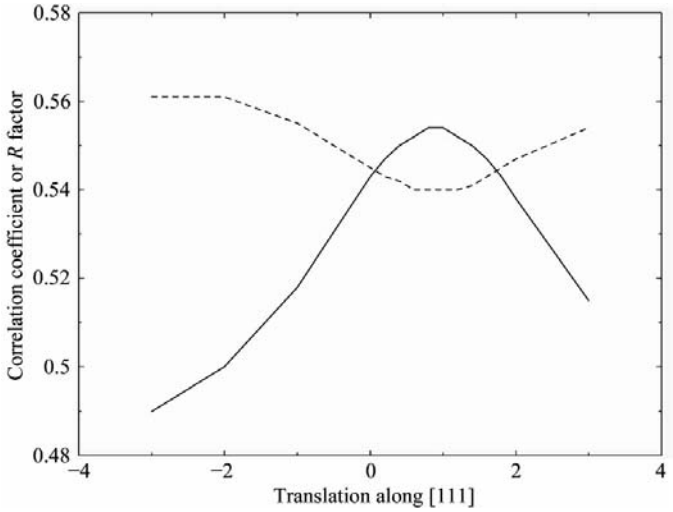

Figure 8
One-dimensional R factor (dashed line) and correlation coefficient (Wikoff *et al.*, 2000) (solid line) search along the body diagonal of the unit cell. Only the results around the (1/4, 1/4, 1/4) position are shown. The data used for the calculation were in the resolution range 50–15 Å. The final search grid was in 0.2 Å intervals (units of the abscissa are Å).

The particle orientations in the unit cell were confirmed by the locked-rotation function, with the icosahedral fivefold and threefold symmetry elements showing a perfect match with the peaks from the self-rotation function calculation (Fig. 7).

Using a pseudo-atomic model that was constructed by fitting the crystal structure of HK97 Head II subunit (Wikoff *et al.*, 2000) into the cryoEM image-reconstruction map of Prohead II (Conway *et al.*, 2001; PDB code 1if0), an R-factor search along the body diagonal was carried out with data in the resolution range 50–15 Å with a 3.5 Å search increment (Fig. 8). The results showed a minimum with the particle centered at (0.25, 0.25, 0.25). A finer grid (0.001) was used to search a small volume from (0.24, 0.24, 0.24) to (0.26, 0.26, 0.26), indicating an R-factor minimum with the particle centered at 0.251, 0.251, 0.251, which is 0.7 Å from the originally determined position.

Using the particle orientation and position determined above, the structure will be determined by molecular-replacement averaging (20-fold) and phase extension to 5.5 Å, using the starting pseudo-atomic Prohead II model (Conway *et al.*, 2001). Initial phasing attempts are under way.

We thank Drs Tianwei Lin and Vijay Reddy for helpful advice and discussion. We thank Dr Reinhard Pahl (APS) for assistance with data collection. Dr Jeffrey Speir is gratefully acknowledged for preparing Fig. 7 and for help in preparation of the manuscript. Use of the APS was supported by the US Department of Energy, Basic Energy Sciences, Contract No. W-31-109-Eng-38. Use of the BioCARS Sector 14 was supported by NIH grant RR07707. This work was supported by NIH grants R01-AI40101to JEJ and GM47795 to RWH.

References

Brünger, A. T. (1992). *X-PLOR Version 3.1 User's Guide.* Yale University, New Haven, CT, USA.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Conway, J. F., Duda, R. L., Cheng, N., Hendrix, R. W. & Steven, A. C. (1995). *J. Mol. Biol.* **253**, 86–99.
Conway, J. F., Wikoff, W. R., Cheng, N., Duda, R. L., Hendrix, R. W., Johnson, J. E. & Steven, A. C. (2001). *Science*, **292**, 744–748.
Duda, R. L. (1998). *Cell*, **94**, 55–60.
Duda, R. L., Hempel, J., Michel, H., Shabanowitz, J., Hunt, D. & Hendrix, R. W. (1995). *J. Mol. Biol.* **247**, 618–635.
Duda, R. L., Martincic, K. & Hendrix, R. W. (1995). *J. Mol. Biol.* **247**, 636–647.
Lata, R., Conway, J. F., Cheng, N., Duda, R. L., Hendrix, R. W., Wikoff, W. R., Johnson, J. E., Tsuruta, H. & Steven, A. C. (2000). *Cell*, **100**, 253–263.
McPherson, A. (1982). *Preparation and Analysis of Protein Crystals.* New York: John Wiley.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Rossmann, M. G. & Blow, D. M. (1962). *Acta Cryst.* **15**, 24–31.
Tong, L. & Rossmann, M. G. (1990). *Acta Cryst.* A**46**, 783–792.
Wikoff, W. R., Duda, R. L., Hendrix, R. W. & Johnson, J. E. (1998). *Virology*, **243**, 113–118.
Wikoff, W. R., Liljas, L., Duda, R. L., Tsuruta, H., Hendrix, R. W. & Johnson, J. E. (2000). *Science*, **289**, 2129–2133.