Acta Cryst. (1997). D53, 738–746

Resolution of Space-Group Ambiguity and the Structure Determination of Nodamura Virus to 3.3 Å Resolution from Pseudo-*R*32 (Monoclinic) Crystals

ADAM ZLOTNICK,† PADMAJA NATARAJAN, SANJEEV MUNSHI‡ AND JOHN E. JOHNSON*

Department of Molecular Biology, MB-13, The Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla, CA 92037, USA. E-mail: jackj@scripps.edu

(*Received 7 February 1997; accepted 16 May 1997*)

Abstract

Monoclinic crystals of nodamura virus (NOV) have two virus molecules per asymmetric unit. Packing analysis reveals a pseudo-rhombohedral (pseudo-*C*2 monoclinic) arrangement of particles in the actual $P2_1$ space group (a = 562.1, b = 354.1, c = 612.8 Å, β = 110.9°). The *R*32 symmetry is broken rotationally and translationally. The pseudo-symmetry of the unit cell results in three possible monoclinic origins and also restrains the four particles in the unit cell to similar orientations. NOV particles deviate by less than 3° from the ideal orientations, causing overlap of peaks in the rotation function and the generation of peaks that were not interpretable as particle symmetry elements. The space-group ambiguity was resolved by analysing the relationship between the particle orientations determined by high-resolution rotation functions and the attenuation of peak heights in native Patterson maps. Particles were centered less than 1 Å from the *R*32 special positions. Three different approaches were required to identify the correct particle center. Following the solutions of the rotation and translation problems, phases were computed using the coordinates of flock house virus (FHV), another member of this virus family. The phases were improved by real-space molecular averaging with a 120-fold non-crystallographic symmetry and by solvent flattening with a spherical mask. The final model for the NOV structure was built using the 3.3 Å averaged map. While the overall subunit structure was very similar to that of other nodaviruses, FHV and black beetle virus, NOV showed distinct structural features near particle threefold and quasi-threefold axes and at the protein–RNA interfaces that are consistent with phenotype differences among the related viruses.

1. Introduction

Nodamura virus (NOV) is the type member of the nodavirus family. Its natural hosts are insects, though it can also infect and kill neonatal mice (for reviews see: Kaesberg, 1987; Hendry, 1991). The capsid of NOV is constructed of 180 copies of the coat protein α in three different, but quasi-equivalent, environments corresponding to a T = 3 surface lattice (Caspar & Klug, 1962). *In vivo* the virus coat protein co-assembles with the bipartite RNA genome to form a provirion. The provirion matures to the more stable infectious form by autoproteolysis of α protein to β protein and γ peptide (Hosur *et al.*, 1987; Gallagher & Rueckert, 1988). A mechanism has been proposed for this cleavage (Zlotnick *et al.*, 1994) based on the structures of two closely related viruses: black beetle virus (BBV; Hosur *et al.*,1987; Wery, Reddy, Hosur & Johnson, 1994) and flock house virus (FHV; Fisher & Johnson, 1993).

Our preliminary results of crystallographic studies on NOV reported by Zlotnick *et al.* (1993) are summarized as follows. NOV crystallized in a primitive monoclinic form with four 8 MDa virus particles per unit cell (a = 562.1, b = 352.1, c = 612.8 Å, β = 110.9°). Pseudo-*C*2 symmetry was manifested at low resolution by systematically weak $h + k$ odd reflections. Pseudo-face centering was also seen in the *A* and *B* faces but was not as strong as the *C* centering. This suggested a rhombohedral (*R*32) packing arrangement with particles located at pseudo-special positions in a monoclinic cell

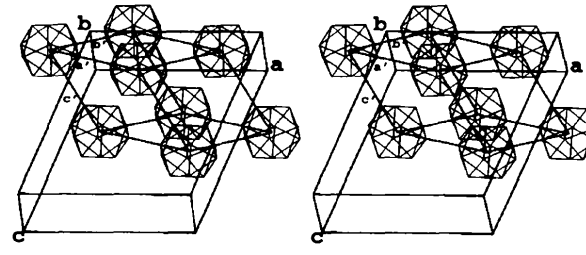


Fig. 1. Packing diagram showing the pseudo-*R*32 symmetry observed in the monoclinic NOV crystals. The virus particles are positioned in the monoclinic unit cell with axes a, b and c. The packing is pseudo-face-centered. An icosahedral twofold axis of each of the four virus particles is almost parallel to the monoclinic b axis. An icosahedral threefold axis is nearly parallel to the monoclinic 101 axis, which defines the body diagonal of the rhombohedral unit cell with axes a', b', c'. (For clarity, the icosahedra representing the virus particles in this diagram are shown as T = 1 icosahedral cages with reduced diameter compared with virus diameter.)

† Present address: Laboratory for Structural Biology Research, NIAMS-NIH, Bethesda, MD 20892, USA.

‡ Present address: WP44-B122, Merck Research Laboratory, Lansdale, PA 19486, USA.

Table 1. *Post-refinement results and data statistics*

Diffraction patterns were collected on 12 × 12 cm films and 25 × 21 cm Fuji image plates using oscillation photography at the Cornell High Energy Synchrotron Source (CHESS) F-1 station. An oscillation angle of 0.3° and a crystal-to-detector distance of 210 mm for films and 360 mm for image plates was used to minimize the number of overlapped reflections (the larger distance for image plates was required because of their reduced spatial resolution compared to film). The data were processed and refined with the Purdue University suite of programs (Rossmann, 1979; Rossmann, Leslie, Abdel-Meguid & Tsukihara, 1979).

(*a*) Post-refinement results

	Film data	Image-plate data	Combined data†
Number of films/image plates	206	82	206 films + 81 image plates
R factor [$I/\sigma(I)$ cut-off = 4.0] ‡	0.1400	0.1232	0.1410
Vertical mosaicity (°)	0.1376	0.1853	0.1253 (films), 0.1961 (image plates)
Horizontal mosaicity (°)	0.1971	0.1554	0.1961 (films), 0.1792 (image plates)
λ (Å) for data processing	0.910	0.913 §	
Post-refinement cell parameters (Å, °)		a = 562.1, b = 354.1, c = 612.8, β = 110.9	

(*b*) Data statistics for the combined data set

Resolution (Å) of data (for films and image plates)	3.3
No. of unique reflections	1714600
No. of independent observations	2030082
No. of possible reflections	3326074

Resolution bins (Å)	25.0–20.0	20.0–15.0	15.0–11.0	11.0–9.0	9.0–7.5	7.5–6.2	6.2–5.5	5.5–4.7	4.7–4.1	4.1–3.7	3.7–3.3
% Data	78	60	57	66	65	59	59	60	57	50	39

† The mosaicity parameters were refined separately for film and image-plate data in the final combined data set. ‡ $R = \{[\sum_h \sum_i (I_h - I_{hi})]/(\sum_h \sum_i I_{hi})\}$, where I_h is the mean of the I_{hi} observation for reflection h. § The wavelength 0.913 Å was determined by comparison of post-refinement unit-cell dimensions from image-plate data with those obtained previously from film data.

(Fig. 1). The $R32$ orientation for the particles was confirmed by low-resolution self- and locked-rotation functions. The corresponding packing arrangement was confirmed by the face-centered peaks observed in a native Patterson map. At that time, the rotation function could not be interpreted, the choice of three possible origins could not be resolved and, therefore, it was impossible to determine if there was any translational deviation from the pseudo-special positions. This paper describes how these issues were resolved as well as the phase determination that produced the NOV structure at 3.3 Å.

The structure determination of coxsackie virus B3 (CVB3) (Muckelbauer *et al.*, 1995) was complicated by a nearly identical form of pseudo-symmetry. The simultaneous analysis of the pseudo-symmetry in NOV and CVB3 was of mutual benefit to the two projects.

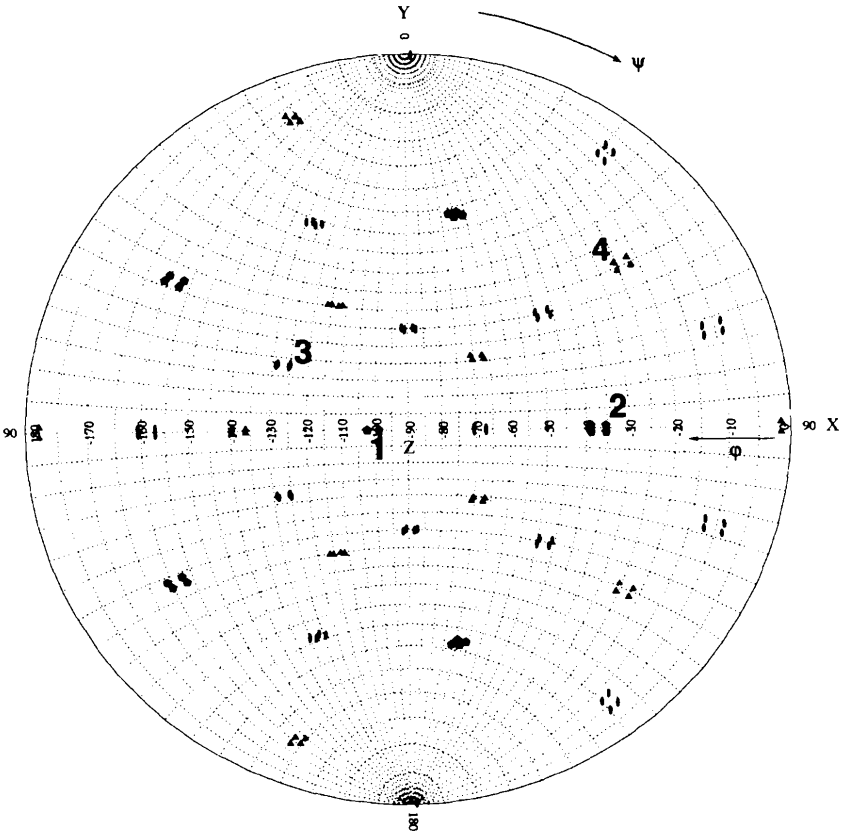
2. Data collection and processing

Preparation of material, crystallization and a preliminary analysis of a partial data set collected on film have been reported by Zlotnick *et al.* (1993). Table 1 gives the details of the later refinement and statistics for the final 3.3 Å data set. The post-refinement unit-cell dimensions were a = 562.1, b = 354.1, c = 612.8 Å, β = 110.9°. The final data set was compiled from 206 films (39 crystals) and 81 image plates (also 39 crystals) resulting in 1 714 600 unique reflections with $I/\sigma(I) > 4.0$, from 2 030 082 independent observations (800 812 whole reflections and 1 229 270 partial reflections with calculated partiality > 0.5).

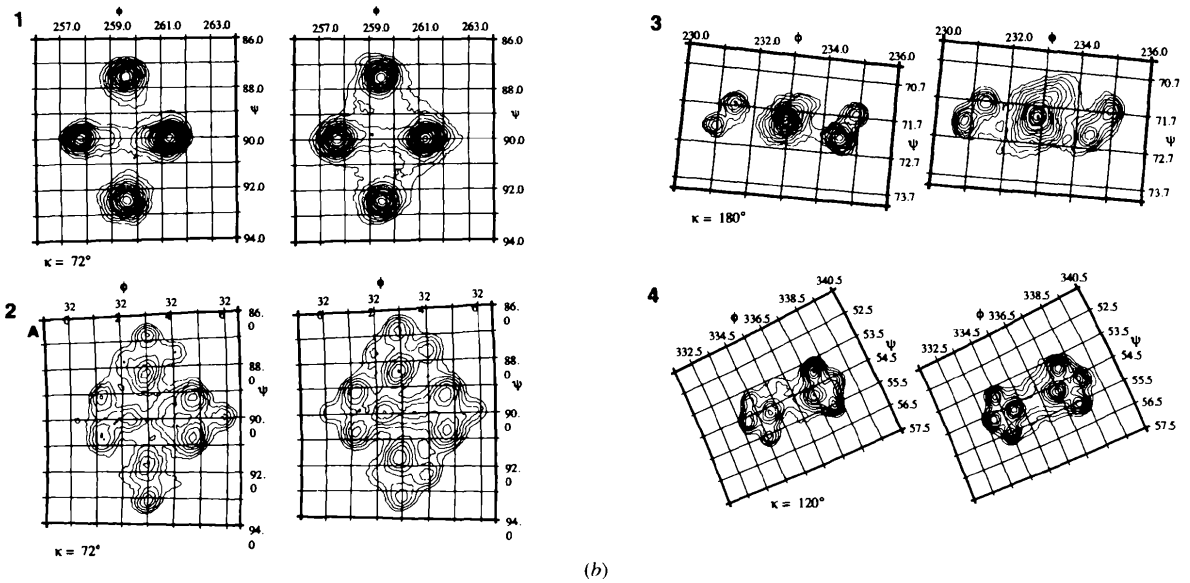
3. Rotation-function solution

Self-rotation functions, computed with reflections in the resolution range 3.9–3.7 Å from the partial data set reported by Zlotnick *et al.* (1993) yielded a constellation of peaks consistent with the orientation of a single icosahedron for the four particles in the unit cell. This showed that the particles were in nearly identical orientations with an icosahedral twofold axis of the independent particles in the asymmetric unit being nearly parallel to the crystallographic twofold axis (b axis). Since the pseudo-systematic absences in the data and the different Patterson peak heights observed for the A, B and C faces (described below) did not correspond to the higher symmetry space group, more than one icosahedral constellation of peaks in the rotation function was expected. The pseudo-symmetry was so close to perfect at this resolution, however, that the assignment of more accurate orientations for the individual particles was impossible.

The self-rotation functions with the more complete current data set showed splitting for each of the two-, three- and fivefold peaks (Fig. 2). Since each peak split into four or more peaks, it was clear that the independent particles in the asymmetric unit not only differed in orientation with respect to the crystallographic screw axis, but also were rotated with respect to one another.



(*a*)



(*b*)

Fig. 2. Stereographic projections of the self-rotation functions for NOV. (*a*) Distribution of icosahedral axial positions resulting from the four particles in the unit cell. Assignments of the axial positions here are based on the interpretation of the high-resolution (3.9–3.7 Å) self-rotation function peaks and the interpreted orientations of the two particles in the asymmetric unit are: particle 1, $\varphi_1 = 247.0$, $\psi_1 = 2.60$, $\kappa_1 = 23.0°$; particle 2, $\varphi_2 = 240.0$, $\psi_2 = 2.22$, $\kappa_2 = 19.4°$. The two-, three- and fivefold axial positions are shown. (*b*) Peak splitting seen in the high-resolution self-rotation functions. Peaks observed for five- ($\kappa = 72°$), two- ($\kappa = 180°$) and threefold ($\kappa = 120°$) symmetry operators marked 1, 2, 3 and 4 in (*a*) correspondingly decompose into several maxima shown here. Identification of the additional peaks, possibly arising from the special packing arrangement of the particles in the NOV unit cell, was critical in obtaining the particle orientations. Peak splitting for the observed data (on the left) match well with those for the model structure factors (on the right), confirming the particle orientations.

The rotation function should yield four maxima for each peak, corresponding to the individual orientations of the four particles in the unit cell. Yet, some peaks resolved into clusters of five or more (up to 16) discrete maxima (Fig. 2*b*). It was necessary, therefore, to identify the peaks arising from the independent particle orientations. We noticed that in every cluster there was a subset of flanking peaks that could always be divided into peaks that split horizontally and vertically about a central position expected for the *R*32 peak. An icosahedron in an *R*32 orientation would have a value of κ of 20.9°.† The horizontal peak splitting results from the deviations in κ from 20.9°, corresponding to the differences in orientations of the independent particles. Thus, the independent particles were assigned κ values of 23 and 19.4°, respectively. The φ and ψ angles were estimated based on the vertical spacing of maxima around the mirror plane at $\psi = 90°$ (*xz* plane in Fig. 2*a*). The initial estimates for φ and ψ angles were 350 and 1.5°, respectively, for both the independent particles.

Refinement (Rossmann & Argos, 1976) of the observed and idealized direction cosines resulted in the following orientations for the two particles (Fig. 2): particle 1, $\varphi_1 = 247.0$, $\psi_1 = 2.60$, $\kappa_1 = 23.0°$; particle 2, $\varphi_2 = 240.0$, $\psi_2 = 2.22$, $\kappa_2 = 19.4°$.

The final orientation of the particles shifts any point at a radial distance of 140 Å by ~7 Å from the locations dictated by *R*32 symmetry. The final r.m.s. deviations between the observed and calculated self-rotation peak positions were 0.19 and 0.28 Å at a radius of 140 Å, corresponding to a maximum deviation of less than 0.2°. The twofold-related particles have φ orientations of 67 and 60° due to the crystallographic symmetry.

Since the observed rotation-function peaks were resolved into more component peaks than would arise from individual symmetry elements of the icosahedra, an experiment was performed with calculated data to see if the additional peaks could be reproduced. The atomic coordinates of FHV (Fisher & Johnson, 1993) were oriented precisely as described for particles 1 and 2 above. Data were computed between 3.9 and 3.7 Å resolution and rotation functions calculated. This rotation function was compared with that obtained for the observed data and it was clear that the complex decomposition of the peaks was nearly identical for observed and calculated data (Fig. 2). Although the calculated data demonstrate identical distribution of peaks, the rationale for these peaks in terms of particle packing (Åkervall *et al.*, 1972) is not obvious.

† Particle orientations described here follow the spherical polar-coordinate convention defined by Rossmann & Blow (1962). The angles κ, ψ and φ describe the particle orientation with respect to a standard orientation. The three mutually perpendicular icosahedral twofold axes coincide with *x*, *y* and *z* of the reference coordinate system in the standard orientation ($\kappa = 0$, $\psi = 0$ and $\varphi = 0°$). ψ and φ define the direction of the rotation axis and κ is the rotation about this axis. ψ is measured as the tilt angle from the *x* axis to the rotation axis and φ is the angle made with the *x* axis by the projection vector of this rotation axis onto the *xz* plane.

4. Patterson analysis of pseudo-symmetry

The near-parallel directional coincidence of an icosahedral twofold axis of each of the independent particles in the asymmetric unit with the crystallographic screw axis should give the particle positions from the peak positions in the Harker sections of the Patterson function (Zlotnick *et al.*, 1993). Unfortunately, these peaks were coincident with the 'origin peaks' generated by the pseudo-face-centered symmetry present in the unit cell (Fig. 3*a*). However, the relative peak heights of the pseudo-origin peaks helped in choosing the correct particle arrangement. The observed peak height ratio of 1:3:6 for the *A*:*B*:*C* faces (Zlotnick *et al.*, 1993) deviated from the identical peak height that would result from an ideal *R*32 arrangement of particles. The ratio also indicated that the degree of similarity in orientation between the pairs of particles related by face centering increased from the pseudo-*A*-centered pairs to the pseudo-*B*-centered pairs, to pseudo-*C*-centered pairs.

Various particle arrangements were considered, accounting for the three unique crystallographic origins in the monoclinic system (Fig. 3*c*). Twelve unique arrangements are possible when no constraints are placed on the two independent particles. Six are possible when particles are placed at the *R*32 positions in the monoclinic setting. Since the similarity of orientations among the independent particles and symmetry-related particles in each of these arrangements dictate the relative heights of the Patterson peaks, the possible arrangements reduced to two primitive monoclinic space groups that were in agreement with the relatively weak peak for the *A* face: *P*2 and a nearly *C*-centered monoclinic $P2_1$ space group represented here as *C*-$P2_1$ (see 1 and 3 of Fig. 3*c*, respectively).

The space group $P2_1$ identified as *C*-$P2_1$ was consistent with the observed pseudo-systematic absences and was confirmed by the correlation between observed data and the data calculated with the FHV model (Table 2).

The orientations and idealized positions of the two independent particles in the asymmetric unit for *C*-$P2_1$ resulting from rotation-function and Patterson analyses were, particle 1: $\varphi_1 = 67.0$, $\psi_1 = 2.60$, $\kappa_1 = 23.0°$ and $x_1 = \frac{1}{4}$, $y_1 = 0$, $z_1 = 0$; particle 2: $\varphi_2 = 240.0°$, $\psi_2 = 2.22°$, $\kappa_2 = 19.4°$ and $x_2 = \frac{1}{4}$, $y_2 = \frac{1}{2}$, $z_2 = \frac{1}{2}$.

The 180° change of φ_1 was required for the *C*-$P2_1$ packing arrangement (Fig. 3*c*).

5. Refinement of particle positions from their initial *R*32 positions

The refinement of particle positions employed an approach similar to that used to solve the φX174

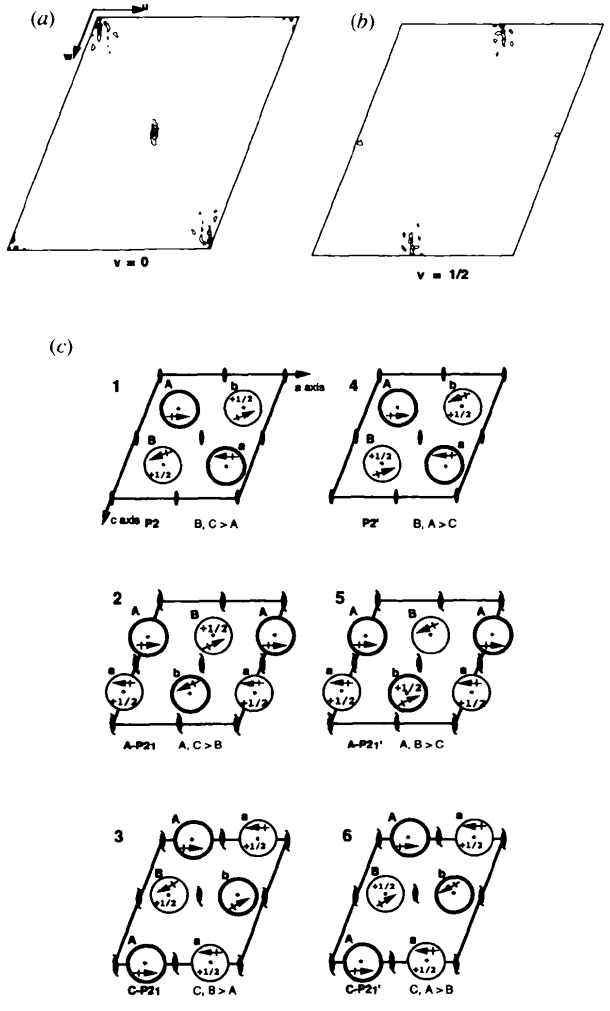


Fig. 3. Analysis of the native Patterson for identification of particle arrangement. (*a*), (*b*) The Harker planes in the 20–10 Å NOV native Patterson showed a relative peak height of 1:3:6 for the peaks on the *A*:*B*:*C* faces. The differential peak heights seen here are due to the small deviations in the particle orientations with respect to the crystallographic twofold axis. This is exploited in identifying the correct arrangement of particles among several possible choices in the monoclinic cell of NOV shown in (*c*): the six possible particle arrangements (numbered 1 to 6) in the monoclinic space-group setting for NOV by placing the particles at the *R*32 positions. As observed in the native Patterson map, both *P*2 and *C*-*P*2$_1$ arrangements (1 and 3) lead to a minimum relative height for the *A*-face peak because of maximum dissimilarity in orientations of particles on this face. Evidence for *C*-*P*2$_1$ being the correct arrangement comes from the better figures of merit (*R* factor and correlation coefficient) between the observed and calculated structure factors for *C*-*P*2$_1$ compared with those for *P*2. In addition, the systematic absences observed for the 0 k_{odd} 0 reflections confirmed the presence of a 2$_1$ screw axis along the *b* axis. [Particles are shown as circles. The arrows represent κ rotation of particles with respect to the twofold axis of the standard icosahedral particle oriented along the crystallographic *b* axis (in the monoclinic system) and the radial distance of the arrow from particle center is representative of the ψ-angle deviation of the particle from the crystallographic *b* axis ($\psi = 0.0°$). The prefixes *A*- and *C*- for the space group *P*2$_1$ represent centering symmetry on these faces.]

structure (McKenna, Xia, Willingmann, Ilag & Rossmann, 1992). Translation searches were performed with data from slightly different low-resolution ranges. The *R*32 special positions listed above were the starting points for all the searches. Three different procedures were used, as described below.

In the *TRANSF* procedure (Argos & Rossmann, 1980; Tong & Rossmann, 1993), calculated structure factors based on an atomic model or density are compared with all the observed structure factors above specified σ and amplitude thresholds. The program performs a systematic *R*-factor or correlation-coefficient search for the best possible position for a particle in the given translation space. For NOV, the initial *R*32 positions for the two independent particles of the asymmetric unit were refined one particle at a time using the procedure. In each search, the contribution to structure factors from one particle was fixed (immobile) while the other particle was allowed to move in the specified translation range until convergence was reached.

RIGID (Arnold & Rossmann, 1988) is a rigid-body least-squares minimization procedure requiring an atomic model as input. In the *RIGID* search, *R* factors are calculated using a small subset of the structure factors which exceed a very high amplitude threshold. The choice of observed structure factors used in the calculation is varied with each cycle to minimize the bias.

CLIMB (Muckelbauer *et al.*, 1995) is a real-space averaging procedure that minimizes $\sigma(\rho)$, the r.m.s. deviation from the mean density for the icosahedrally related points in the unaveraged density map. From the initial best NOV positions and orientations, a systematic search or 'climb' was performed in *x*, *y*, *z* (fractional coordinates) and θ_1, θ_2, θ_3 (Eulerian angles) to locate the positions and orientations that resulted in minimum $\sigma(\rho)$.

A consensus was developed from the very similar particle centers (0.2501, 0, −0.0001; 0.2499, 0.5002, 0.4998) obtained from the three search procedures (Table 2) and a map was calculated with data from 20 to 7.3 Å with phases based on an FHV polyalanine model. After eight rounds of molecular averaging with the program *ENVELOPE* (Rossmann *et al.*, 1992; Cornea-Hasegan *et al.*, 1995) with a spherical envelope (inner radius 80 Å, outer radius 175 Å, based loosely on the dimensions of FHV) the correlation coefficients and *R* factor converged (Table 2). The structure factors obtained from this averaged map were used for the subsequent round of translation searches with *TRANSF* and the unaveraged map computed with phases from the revised model at this stage was used in the *CLIMB* procedure.

The 7.3 Å map was displayed with the FHV model using the program *O* (Jones, Zou, Cowan & Kjeldgaard, 1991). The model fits the overall features of the map. Two specific differences between the model and the map gave confidence that the phase bias towards the model was broken by the 120-fold non-crystallographic averaging. First, the quasi-threefold protrusion of the map

Table 2. *Determination of the correct monoclinic space group*

Space group†	Resolution (Å)	Particle positions	Correlation coefficient/*R* factor‡ Model	Average (cycles)
C-$P2_1$	20–10.2	1/4, 0, 0 1/4, 1/2, 1/2	0.638/0.394	–
A-$P2_1$	20–10.2	0, 0, 1/4 1/2, 1/2, 1/4	0.185/0.608	–
C-$P2_1$	20–7.3	0.2501, 0, −0.0001 0.2500, 0.5002, 0.4998	0.572/0.400	0.854/0.217 (8)
C-$P2_1$	20–4.55	0.2503, 0, 0 0.2499, 0.5005, 0.4995	0.358/0.461	0.807/0.279 (9)
C-$P2_1$	20–4.2	0.2503, 0, 0 0.2499, 0.5005, 0.4995	0.372/0.443	0.755/0.265 (12)

† The prefix *A*- or *C*- to the monoclinic space-group representation corresponds to particles being nearly *A* or *C* centered (see also Fig. 3*c*). ‡ Figures of merit: the correlation coefficient (CC) and *R* factor are defined as follows: $CC = \sum(\langle F_o\rangle - F_o)(\langle F_c\rangle - F_c)/[\sum(\langle F_o\rangle - F_o)^2(\langle F_c\rangle - F_c)^2]^{1/2}$ and $R = [\sum_h|(|F_o| - k|F_c|)|]/\sum_h|F_o|$. Better figures of merit were observed for the correct particle arrangement in the monoclinic space group *C*-$P2_1$(see Fig. 3*c*). A polyalanine model of FHV for the first three cases and an NOV polyalanine model for the later two cases were used for calculating the structure factors.

was considerably larger than that of the model (Fig. 4*a*), consistent with the insertions in the NOV coat-protein sequence in this region with respect to the other nodaviruses (Kaesberg *et al.*, 1990). Second, no γ peptide for the *B* subunit was included in the phasing model since this density in FHV is very weak. In the NOV map, density corresponding to the *B*-subunit γ peptide was at least as strong as that of the other two γ peptides.

Another iterative series of translation functions using *TRANSF* were calculated using the averaged electron-density map, described above, as a search model. The results of the translation function from the correlation-coefficient and *R*-factor searches yielded the same solution for the particle centers: 0.2503, 0, 0 for particle 1 and 0.2499, 0.5005, 0.4995 for particle 2. Searches were run in narrow shells from 7.4 to 7.8 Å and from 8.3 to 9.0 Å, in both cases using ~23 000 large terms out of 30 000 reflections ($I > 0.3I_{\text{average}}$). *CLIMB* results based on a 7.3 Å map were identical with those obtained using *TRANSF*, reported above. We had greater confidence in these results because the new search model more closely resembled NOV.

Phase refinement at 4.55 Å was initiated with phases based on the FHV model, using the *TRANSF* positions (Table 2). The map was readily interpretable after nine cycles of molecular averaging and phase refinement. The FHV polyalanine model was extensively modified to correspond to the new electron density map using the program *O* (Jones *et al.*, 1991). Density was readily interpretable for only 17 large side chains; thus, most of the amino acids were modeled as alanine in the model based on a 4.55 Å resolution map.

This NOV model was used to calculate phases to 4.2 Å. Even before averaging, the improvement in the statistics was evident (Table 2). Previously, maps were averaged to the maximum extent of their resolution, without extending phases. In the case of the 4.2 Å map, six cycles of averaging were performed on a map calculated from 20 to 6 Å with initial phases computed from the atomic model. These phases were combined with model phases to 4.2 Å followed by six additional rounds of averaging. The final correlation coefficient and *R* factor are 0.755 and 0.265, respectively. The resulting map was superior to its predecessor in its detail and in the continuity of weak density. After re-modeling at this stage, new structure factors were calculated to 3.3 Å. This time, the 120-fold averaging and solvent flattening were performed to 3.3 Å in three phase-extension steps (4.2, 3.9 and 3.3 Å resolution), starting with 20–6.0 Å data and following the same procedure as explained above. Most of the residues (55–383 for *A*, 58–380 for *B*, and 13–29 and 58–380 for *C* subunits) were visible with their side-chain density in the 3.3 Å averaged map; the model was re-built using the graphics program *O* (Jones *et al.*, 1991).

6. Refinement of the NOV model using *X-PLOR*

The NOV atomic coordinates were refined using *X-PLOR* (Brünger, 1992) to obtain reasonable geometry and avoid unacceptable non-bonded atomic contacts. A total of 180 cycles of energy minimization were carried out with a fixed *B* value of 18.0 Å^2 using all reflections (1 377 412) in the resolution range 8.0–3.5 Å with no modeled water molecules; the *R* factor converged to 0.344. When all the reflections (1 708 847) in the resolution range 20.0–3.3 Å were considered, the *R* factor dropped from an initial value of 0.399 to 0.362. 94.0% of the amino acids (excluding glycines and prolines) were in the most favored or allowed regions of the Ramachandran plot (*PROCHECK*; Laskowski, MacArthur, Moss & Thornton, 1993). The new model was inspected for main-chain errors and for amino acids with disallowed φ and ψ angles, and corrected. A new map with phases calculated from the adjusted model was subjected to molecular

averaging followed by energy minimization and visual adjustments of the model using a graphics interface. This model was subjected to simulated annealing (T = 3000–300 K, time step = 0.5 fs, T-coupling temperature decrement of 50 K every 12.5 fs, Δf = 0.2 Å), using all reflections between 8.0 and 3.5 Å resolution and with no modeled water molecules. This was followed by 20 cycles of positional refinement (Δf = 0.0 Å) and 12 cycles of unrestrained individual B-factor refinement. The R factor for the model was 0.296 and r.m.s.

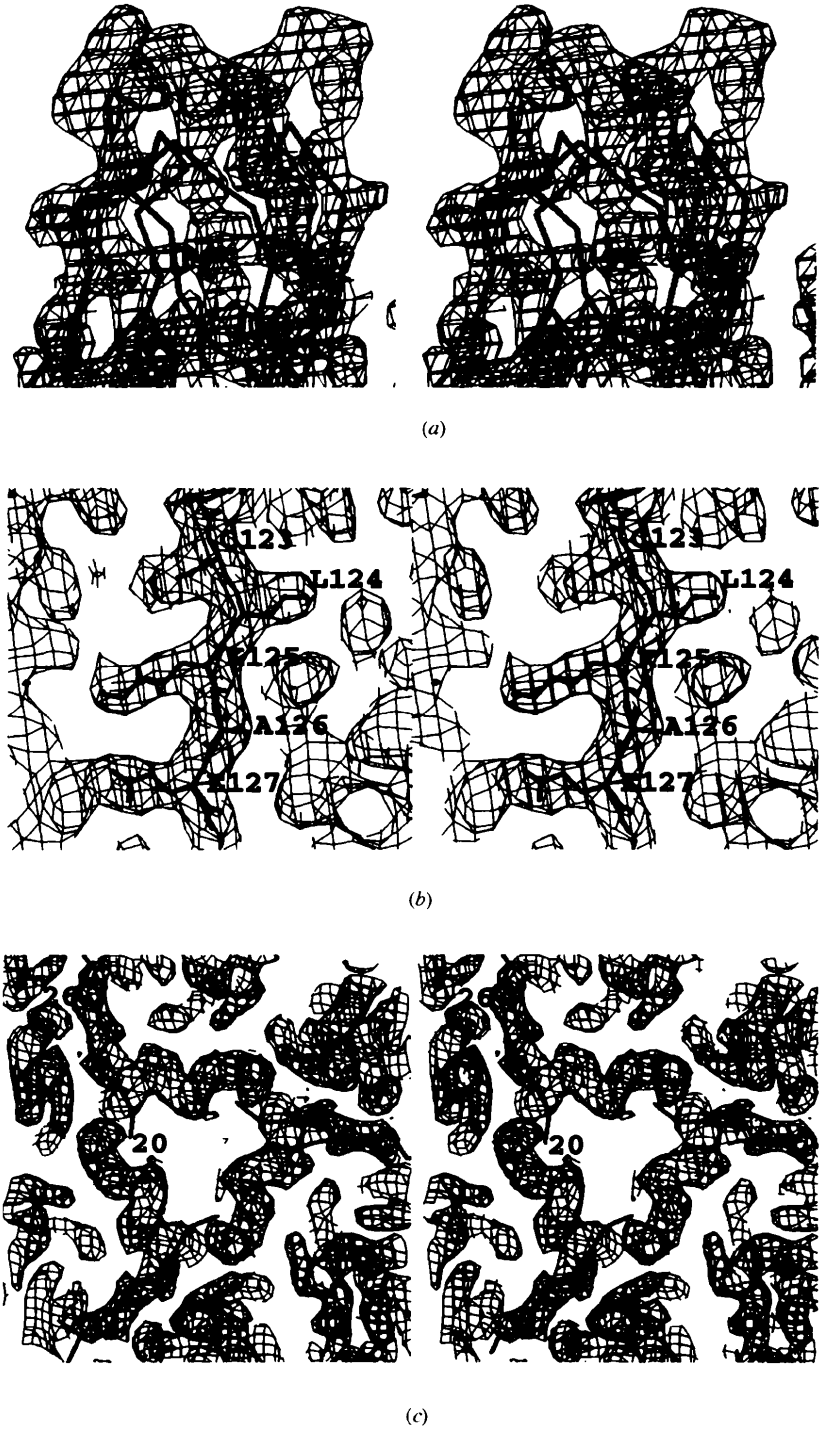


Fig. 4. Details of the 120-fold averaged electron-density map at 3.3 Å and the NOV model. (*a*) Electron density for the protrusions related by the quasi-threefold axis at the surface of NOV (in a view approximately perpendicular to the quasi-threefold axis). The protrusions are clearly much longer that the corresponding ones in the FHV model (Cα backbone shown here), demonstrating the absence of model bias in the averaged NOV map. (*b*) NOV model fit to the 3.3 Å electron-density map. The short stretch of residues 123–127 show here are from the *B* subunit of NOV (thick line). The FHV model, although structurally very close to NOV in this region (shown as Cα backbone; thin line) has a very different sequence (residues Y, W, S, A and S, respectively for the stretch 123–127). (*c*) Density for the peptide in the groove (residues 13–29 of N terminus of the *C* subunit in NOV). The peptide folds into a helix close to the icosahedral threefold axes. The unique helix is followed by the extended structure (residues 20–30) seen in corresponding peptides of FHV (solid lines) and BBV (not shown).

deviations in bond distances and bond angles were 0.009 Å and 1.646°, respectively.

The quality of the final electron-density map, computed with molecular replacement followed by phase refinement using real-space molecular averaging with 120-fold non-crystallographic symmetry, was very good (Fig. 4), despite the rather high crystallographic R factor. The high R factor can be attributed to the large unit cell and its effect on scaling statistics. At resolution beyond 6 Å there is significant overlap of reflections along the c^* direction. Deconvolution corrections were applied, but the scaling statistics suggest that they were only partially successful. A second factor affecting the overall R factor was the collection of data with two different detectors. While these data scaled reasonable well, the overall scaling R factor was high (14.1%). Despite these problems with the data, which was only 39% complete in the highest resolution shell (Table 1), the quality of the electron-density map was superb. This was undoubtedly due to the 120-fold redundancy used in the phase refinement.

The 3.3 Å map was used to build the final NOV structure. The overall fold of NOV was very close to those of FHV and BBV. There were several unique features that were present in NOV. The capsid protein cleaves into a 363-residue β protein and short γ peptide, about 44 residues long. The β protein which forms the contiguous capsid folds into an eight-stranded β-barrel, commonly found in non-enveloped icosahedral viruses (Rossmann & Johnson, 1989). The γ peptide located in the particle interior forms an amphipathic helix. The symmetry-related γ peptides at the icosahedral fivefold axes form a helical bundle. At particle twofold axes, γ peptides are in close proximity to a partially ordered RNA duplex (Fisher & Johnson, 1993).

Most of the unique features of NOV were seen in non-homologous regions, in loops connecting the strands in the β-barrel and in N and C terminal regions (Kaesberg *et al.*, 1990). The protrusions at quasi-threefold axes on the viral surface were longer in NOV than in FHV and BBV (Fig. 4*a*) as expected from insertions in the NOV sequence. In NOV, the peptide in the groove [residues 13–29 at the N terminus of the C subunit; Fig. 4(*c*)] folds into a helix close to the icosahedral threefold axes. The unique helix is followed by the extended structure seen in the corresponding peptides of BBV and FHV.

A detailed comparison is being made between structures of different members of the nodavirus family of viruses and will be related to their distinct biological and physico-chemical properties (Natarajan, Zlotnick & Johnson, 1997).

We are extremely thankful to Jodi Muckelbauer and Michael G. Rossmann of Purdue University for very useful discussions. We are grateful to our colleagues for help in data collection at CHESS and to the staff at CHESS. AZ was supported by a Merck graduate fellowship. The work was supported by a grant from NIH (GM34220) to JEJ.

References

Åkervall, K., Strandberg, B., Rossmann, M. G., Bengtsson, U., Fridborg, K., Johannisen, H., Kannan, K. K., Lovgren, S., Petef, G., Oberg, B., Eaker, D., Hjerten, S., Ryden, L. & Moking, I. (1972). *Cold Spring Harbor Symp. Quant. Biol.* **36**, 469–488.

Argos, P. & Rossmann, M. G. (1980). In *Theory & Practice of Direct Methods in Crystallography*, edited by M. F. C. Ladd & R. A. Palmer. New York: Plenum Press.

Arnold, E. & Rossmann, M. G. (1988). *Acta Cryst.* A**44**, 469–488.

Brünger, A. T. (1992). *X-PLOR.* Version 3.1. *A System for X-ray Crystallography and NMR.* New Haven, CT: Yale University Press.

Caspar, D. L. D. & Klug, A. (1962). *Cold Spring Harbor Symp. Quant. Biol.* **27**, 1–24.

Cornea-Hasegan, M. A., Zhang, Z., Lynch, R. E., Marinescu, D. C., Hadfield, A., Muckelbauer, J. K., Munshi, S., Tong, L. & Rossmann, M. G. (1995). *Acta Cryst.* D**51**, 749–759.

Fisher, A. J. & Johnson, J. E. (1993). *Nature (London)*, **361**, 176–179.

Gallagher, T. M. & Rueckert, R. R. (1988). *J. Virol.* **62**, 3399–3406.

Hendry, D. (1991). *Viruses of Invertibrates*, edited by E. Kurstak, pp. 227–276. New York: Marcel Dekker.

Hosur, M. V., Schmidt, T., Tucker, R. C., Johnson, J. E., Gallagher, T. M., Selling, B. H. & Rueckert, R. R. (1987). *Proteins Struct. Funct. Genet.* **2**, 167–176.

Kaesberg, P. (1987). *Molecular Biology of Positive Strand RNA Viruses*, edited by D. J. Rowlands, M. A. Mayo & B. W. J. Mahy, pp. 207–218. London: Academic Press.

Kaesberg, P., Dasgupta, R., Sgro, J.-Y., Wery, J.-P., Selling, B. H., Hosur, M. V. & Johnson, J. E. (1990). *J. Mol. Biol.* **214**, 423–435.

Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

McKenna, R., Xia, D., Willingmann, P., Ilag, L. L. & Rossmann, M. G. (1992). *Acta Cryst.* B**48**, 499–511.

Muckelbauer, J. K., Kremer, M., Minor, I., Tong, L., Zlotnick, A., Johnson, J. E. & Rossmann, M. G. (1995). *Acta Cryst.* D**51**, 871–887.

Natarajan, P., Zlotnick, A. & Johnson, J. E. (1997). In preparation.

Rossmann, M. G. (1979). *J. Appl. Cryst.* **12**, 225–238.

Rossmann, M. G. & Argos, P. (1976). *J. Mol. Biol.* **105**, 75–96.

Rossmann, M. G. & Blow, D. M. (1962). *Acta Cryst.* **15**, 24–31.

Rossmann, M. G. & Johnson, J. E. (1989). *Annu. Rev. Biochem.* **58**, 533–573.

Rossmann, M. G., Leslie, A. G. W., Abdel-Meguid, S. S. & Tsukihara, T. (1979). *J. Appl. Cryst.* **12**, 570–581.

Rossmann, M. G., McKenna, R., Tong, L., Xia, D., Dai, J., Wu, H., Choi, H. K. & Lynch, R. E. (1992). *J. Appl. Cryst.* **25**, 166–180.

Tong, T. & Rossmann, M. G. (1993). *J. Appl. Cryst.* **26**, 15–21.

Wery, J.-P., Reddy, V. S., Hosur, M. V. & Johnson, J. E. (1994). *J. Mol. Biol.* **235**, 565–586.

Zlotnick, A., McKinney, B. R., Munshi, S., Bibler, J., Rossmann, M. G. & Johnson, J. E. (1993). *Acta Cryst.* D**49**, 580–587.

Zlotnick, A., Reddy, V. S., Dasgupta, R., Schneemann, A., Ray, W. J. Jr, Rueckert, R. R. & Johnson, J. E. (1994). *J. Biol. Chem.* **29**, 13680–13684.