($CaIn_2$, $CeCu_2$) or anticuboctahedron (Ni_3Sn) to multispheric polyhedra gives the maximal total energy of the atomic interaction between the central atom and the first coordination sphere because of the maximal coordination number.

It is worth noting that many ZP are bispheric and presumably this property is essential for the maximal coordination number and for the gain in total energy of the atomic interactions.

The intermetallic structures have many examples of dual polyhedra in adjacent coordination spheres. This is one of the important properties of these structures in contrast to the structures of the ionic compounds.

References

ASLANOV, L. A. (1988*a*). *Acta Cryst.* B**44**, 449-458.
ASLANOV, L. A. (1988*b*). *Acta Cryst.* B**44**, 458-462.
ASLANOV, L. A. (1989*a*). *Acta Cryst.* A**45**, 671-678.
ASLANOV, L. A. (1989*b*). *Structures of Substances.* Moscow Univ. Press (in Russian).
ASLANOV, L. A. (1991). *Acta Cryst.* A**47**, 63-70.
ASLANOV, L. A. & MARKOV, V. T. (1989). *Acta Cryst.* A**45**, 661-671.
VILLARS, P. & CALVERT, L. D. (1985). *Pearson's Handbook of Crystal Data for Intermetallic Phases*, Vol. I. Metals Park, OH: American Society for Metals.
WELLS, A. F. (1984). *Structural Inorganic Chemistry.* Oxford: Clarendon Press.
ZALGALLER, V. A. (1969). *Convex Polyhedra with Regular Faces.* New York: Consultants Bureau.

Acta Cryst. (1992). A**48**, 293-301

Ab Initio Phase Determination for Viruses with High Symmetry: a Feasibility Study

BY JUN TSAO,* MICHAEL S. CHAPMAN† AND MICHAEL G. ROSSMANN

Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907, *USA*

(*Received* 22 *July* 1991; *accepted* 5 *November* 1991)

Abstract

Conditions that would permit the complete structure determination of spherical viruses that have high internal symmetry were examined starting only from an initial spherical shell model. Problems were considered that might arise due to the following. 1. Creation of centric phases due to the simple shell model and its position in the unit cell. The centric symmetry can generally be broken on averaging an initial electron density map based on observed structure amplitudes, provided that the internal molecular symmetry is sufficiently non-parallel to the crystallographic symmetry. 2. Choice of the average model shell radius. Some incorrect radii led to the Babinet opposite solution (electron density is negative instead of positive). Phases derived from other models with incorrect radii failed to converge to the correct solution. 3. Error in structure amplitude measurements. 4. Lack of a complete data set. 5. Error in positioning the initial spherical-shell model within the crystal unit cell. It was found that an error of 1.6 Å caused noticeable phasing error at a resolution greater than 20 Å.

Introduction

There is now abundant literature on the refinement and extension of phases to higher resolution in the presence of noncrystallographic symmetry (Rossmann, 1990; Bricogne, 1974). Use of the power of noncrystallographic symmetry is particularly applicable in the structure determination of icosahedral viruses because they possess ample symmetry, including fivefold axes which can only be local in nature. However, noncrystallographic symmetry is also used more and more frequently in the structure determination of oligomeric proteins or proteins that can be induced to crystallize in more than one crystal form. Here we explore the feasibility of extending phases derived from a very simple low-resolution model to high resolution. The effects of various errors on such phasing were examined theoretically by Arnold & Rossmann (1986). We examine experimentally the effects of various errors and data defects which may inhibit or retard such phasing. A similar investigation has previously been published by Rayment (1983). However, we emphasize, the requirements for phase extension to higher (3 Å) resolution as well as errors in molecular positioning that have not previously been considered.

Spherical-virus-structure determinations have depended on the use of noncrystallographic symmetry

* Current address: Department of Biochemistry, Center for Macromolecular Crystallography, 262BHS, THT 79, University of Alabama, University Station, Birmingham, Alabama 35294, USA.
† To whom correspondence should be addressed.

0108-7673/92/030293-09$03.00

averaging. The structure of tomato bushy stunt virus (Harrison, Olson, Schutt, Winkler & Bricogne, 1978), southern bean mosaic virus (SBMV; Abad-Zapatero, Abdel-Meguid, Johnson, Leslie, Rayment, Rossmann, Suck & Tsukihara, 1980) and satellite tobacco mosaic virus (Liljas, Unge, Jones, Fridborg, Lövgren, Skoglund & Strandberg, 1982) were all determined by the refinement of high-resolution multiple isomorphous replacement (MIR) phases. Human rhinovirus 14 (Rossmann, Arnold, Erickson, Frankenberger, Griffith, Hecht, Johnson, Kamer, Luo, Mosser, Rueckert, Sherry & Vriend, 1985) and poliovirus (Hogle, Chow & Filman, 1985) extended MIR phases to higher resolution using symmetry averaging, thereby building upon the technique of Gaykema, Hol, Vereijken, Soeter, Bak & Beintema (1984). In subsequent cases, there was always available [with the exception of the structure determination of MS2 (Valegård, Liljas, Fridborg & Unge, 1990)] a low -resolution phasing set from which phase extension could be initiated. The results we report here pertain to the structure determination of canine parvovirus (CPV; Tsao, Chapman, Agbandje, Keller, Smith, Wu, Luo, Smith, Rossmann, Compans & Parrish, 1991) where there was no readily available homologous atomic structure or isomorphous data for low-resolution phase determination. Instead, the initial phases were derived from a crude representation of the virus as a hollow shell with uniform density, a model which is probably only valid to about 20 Å resolution. The lessons learned here are applicable to the case where either a crude model can be constructed or where a model can be derived from electron microscopy images. Early examples for phase determination from a crude icosahedral model are that of southern bean mosaic virus (Johnson, Akimoto, Suck, Rayment & Rossmann, 1976) and polyoma virus (Rayment, Baker, Caspar & Murakami, 1982), although in neither case was phasing extended beyond 20 Å resolution.

An inherent difficulty in using an initial model such as a sphere or shell of uniform density is that it is centrosymmetric. Thus, if the centers of viruses within the unit cell are related by a center of symmetry, the phases based on this model will be centric also. However, this centricness will be broken during the first cycle of electron density averaging, provided that none of the noncrystallographic symmetry elements are parallel to crystallographic elements of the same order. Unfortunately, with the 60-fold symmetry of icosahedral viruses, there is a significant chance of crystallographic and noncrystallographic symmetry elements being nearly parallel. For example, the angle between neighboring twofold axes in an icosahedron is only 36°. Hence, there is a significant probability that one of the icosahedral twofold axes is roughly parallel to a crystallographic even-fold axis. This, then, creates difficulties in breaking the centricness of the initial phases at low resolution where a small difference in direction becomes negligible. That was the case for CPV where the positions had to be centrosymmetric and one of the icosahedral axes was inclined by only 2.5° to the monoclinic 2_1 axis. A more extreme case occurred in R32 SBMV (Johnson *et al.*, 1976) where the crystallographic symmetry coincided with much of the icosahedral symmetry of the virus particle. In this case, averaging alone could never break the center of symmetry introduced by the initial model. In general, it is necessary to consider whether the phase determination can escape from the centricness of the initial phasing model.

This work was initiated during attempts to extend phases from a 20 Å spherical model of the CPV structure. Earlier simulations with model structure factors had shown that it should be possible to refine CPV phases at 20 Å. However, unlike the model calculations, the correlation coefficients of CPV decreased as the resolution was increased (Tsao, Chapman, Wu, Agbandje, Keller & Rossmann, 1992). This work was motivated by the need to know how various kinds of error would influence phase extension.

The model structure

The atomic structure of mengo virus (Luo, Vriend, Kamer, Minor, Arnold, Rossmann, Boege, Scraba, Duke & Palmenberg, 1987) was used to calculate a set of 'observed' structure amplitudes (F_{atom}) which could be used for testing various starting conditions for redetermining the phases. As CPV crystallized in a monoclinic unit cell of space group $P2_1$ with one particle per crystallographic asymmetric unit, the Mengo virus structure was placed into a CPV-like $P2_1$ cell with the orientation and position of CPV (Tsao *et al.*, 1992). However, as the mean outer radius of Mengo virus was 1.2 times that of CPV, the unit-cell dimensions of the CPV-like unit cell were 1.2 times larger than the real CPV unit cell. The Mengo virus particle was placed at (0.2544, 0.2500, 0.2412) in a cell with dimensions $a = 319.1$, $b = 423.0$, $c = 323.8$ Å and $\beta = 90.82°$.

Phase extension starting from a spherical model

Phases from a series of hollow-shell models, with outer radii ranging from 145 to 155 Å in steps of 1 Å, were compared to the phases of the F_{atom} 'observed' data set (mengovirus has an approximate radius of 150 Å). The inner radius was also varied, but it was not very sensitive to the agreement of phases. The quality of fit was based on plotting the mean cosine difference between phases derived from the F_{atom} 'observed' data and the spherical-shell models (Fig. 1). The mean cosine phase difference should be 1.0 for perfect phase agreement, zero for random agree-

ment and −1.0 for phases that are 180° different (the 'Babinet' solution). The phases derived from the best spherical model, which had an outer radius of 149 Å and an inner radius of 105 Å, agreed well with the correct phases inside 40 Å resolution and gave a positive agreement beyond about 30 Å resolution (Fig. 1). However, there was a narrow resolution shell between 40 and 30 Å where a slight negative correlation was observed.

Starting with the hollow-shell model of 140 Å outer radius, phase refinement was initiated using the F_{atom} 'observed' data to 20 Å resolution with full weight. The viral envelope was defined by using a spherical shell with outer (165 Å) and inner (90 Å) radii of sufficient size to include all parts of the protein shell of the mengovirus structure. Tangent planes, placed midway between the centers, defined the boundaries between overlapping spheres. The electron density was set to zero inside and outside these radii and icosahedrally averaged (Johnson, 1978) within the viral envelope. After averaging and solvent flattening, the modified electron density map was Fourier back-transformed to give a set of calculated structure factors, F_{b-t}. The calculated phases were then combined with the suitably weighted F_{atom} 'observed' amplitudes to give a new electron density map for the next cycle of averaging. Weights were the product of Sim weights (Sim, 1959, 1960) and Rayment weights (Rayment, 1983) as used in previous structure determinations (Arnold, Vriend, Luo, Griffith, Kamer, Erickson, Johnson & Rossmann, 1987). These weights downweigh those terms whose phases are deemed to be poorly estimated as assessed from the agreement between F_{atom} and F_{b-t}.

After 15 cycles of averaging and Fourier back-transformation, agreement between the target 'observed' phases and the phases derived from the initial spherical model had greatly improved (Fig. 2). The refined phases matched the 'observed' phases almost perfectly beyond 25 Å resolution. The initial negatively correlated region, between 40 and 30 Å resolution, became positively correlated, although there remained a trough in the mean cosine phase-difference curve (Fig. 2). The distributions of phases are shown in Fig. 3 at different stages of refinement. Assuming the pseudo-center of symmetry has not shifted from its original position at (0, 0, 0), it is apparent that the initial centric distribution gradually disappears. As expected, the greatest degree of centricness occurs in the larger reflections at lower resolution.

Phases were then extended in 14 steps to 12 Å resolution, corresponding to an increment of one reciprocal-lattice unit at a time. Three to five cycles of phase improvement were performed at each resolution step. The final phases beyond 20 Å resolution had less than 10° mean phase difference from the correct solution (Fig. 2). However, the trough at 40 to 30 Å resolution still existed in both the mean cosine phase differences (Fig. 2) and the correlation coefficients (Fig. 4). Further phase extension to higher resolution was not deemed necessary as it had now been demonstrated that the correct phase solution could be regained from the F_{atom} 'observed' structure amplitudes and a hollow-shell spherical model. The major difficulty of phase extension from a hollow shell was in bridging the gap between 20 and 12 Å resolution. Furthermore, by this stage, the resolution would be higher than at the start of the successful phase extension for MS2 (Valegård *et al.*, 1990).

An electron density map was calculated using the F_{atom} 'observed' amplitudes to 12 Å resolution and the phases extended from the 20 Å spherical model.

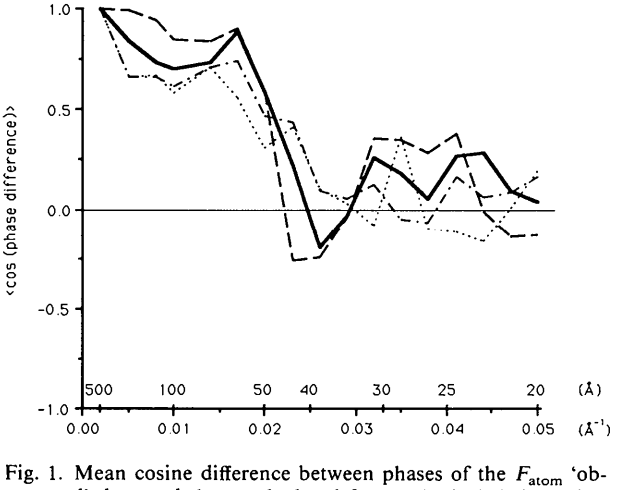


Fig. 1. Mean cosine difference between phases of the F_{atom} 'observed' data and those calculated from spherical-shell models with outer radii 152 Å (dashed) 149 Å (solid), 146 Å (dashed dotted) and 144 Å (dotted).

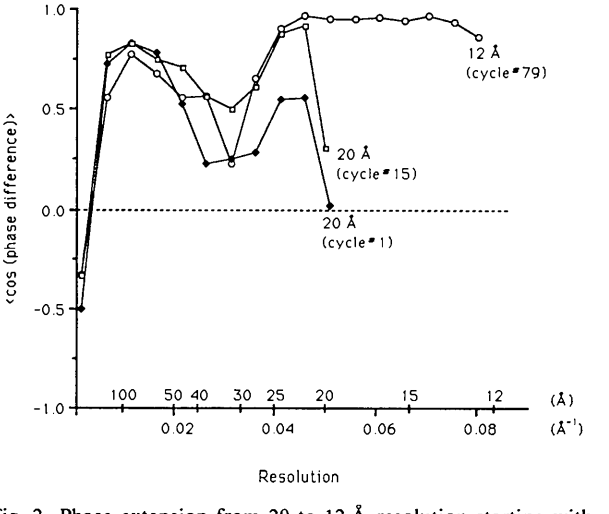
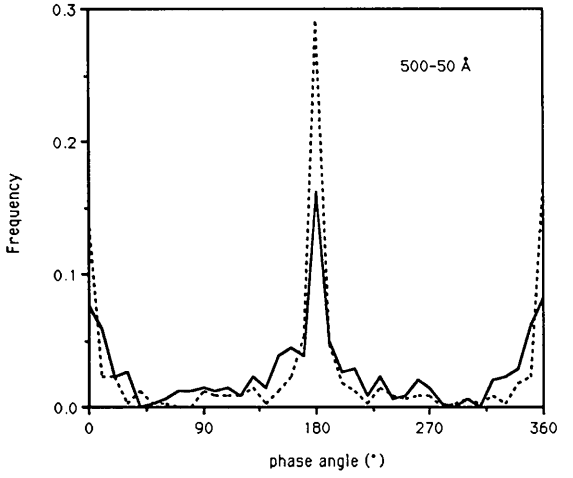
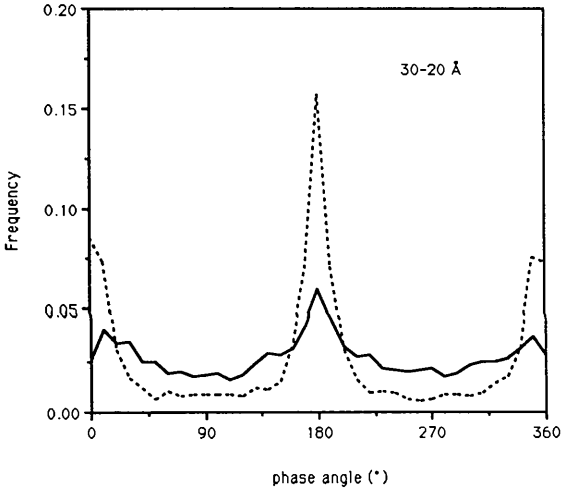


Fig. 2. Phase extension from 20 to 12 Å resolution starting with phases from a hollow-shell model with an outer radius of 149 Å and an inner radius of 105 Å.

It is compared with a map based on the F_{atom} 'observed' amplitudes and phases in Fig. 5 and shows all the correct structural features. These results show that phases from a spherical model converged to the correct solution at 12 Å resolution by using noncrystallographic symmetry averaging. The center of symmetry inherent in the starting phases had been overcome by selecting a hand corresponding to the icosahedral tilt of the particle at a selected center. Phases at resolution less than 25 Å refined more slowly, suggesting that the bias of the initial centric phases could only be removed beyond 25 Å resolution. The poorly matched phases between 40 and 30 Å resolution did not prevent the phases from converging beyond 25 Å resolution.



(*a*)



(*b*)

Fig. 3. Normalized frequency distributions showing the disappearance of the centric distribution of phases on phase refinement from a hollow shell. Dashed lines correspond to cycle 1, continuous lines correspond to phases after 15 cycles of refinement.

The effect of error in model radii

One of the major sources of difficulty in an *ab initio* structure determination is the choice of a suitable outer radius. At low resolution, the structure amplitudes of a typical icosahedral virus can correspond rather well to the transform of a uniform spherical shell (Fig. 6) (Johnson *et al.*, 1976). However, if the observable data are absent at very low resolution, then some confusion may arise as to choice of a suitable radius. For example, in Fig. 7, radii of 149 and 139 Å are equally probable for data at about 22 Å resolution. In this example, the choice of the wrong radius led to a solution with phases 180° out of phase with the correct solution, corresponding to the Babinet opposite solution, that is with a solution where positive protein density is negative. This situation may have occurred in the structure determinations of MS2 (Valegård *et al.*, 1990), CPV (Tsao *et al.*, 1991) and φX174 (McKenna, Xia, Willingmann, Ilag, Krishnaswamy, Rossmann, Olson, Baker & Incardona, 1992). A serious problem occurs if the radii are intermediate such that the phases are not able to converge to either the true solution or its Babinet opposite.

In light of the curves shown in Fig. 7, a number of different radii were chosen for starting models. These radii were 152 Å (slightly too large), 149 Å (the correct radius), 146 Å (slightly too small), 144 Å (halfway between the correct radius and the 'opposite' Babinet solution at 139 Å) and 139 Å (the 'opposite' solution). It was shown in the previous section that, after 15 cycles, a starting radius of 149 Å converged

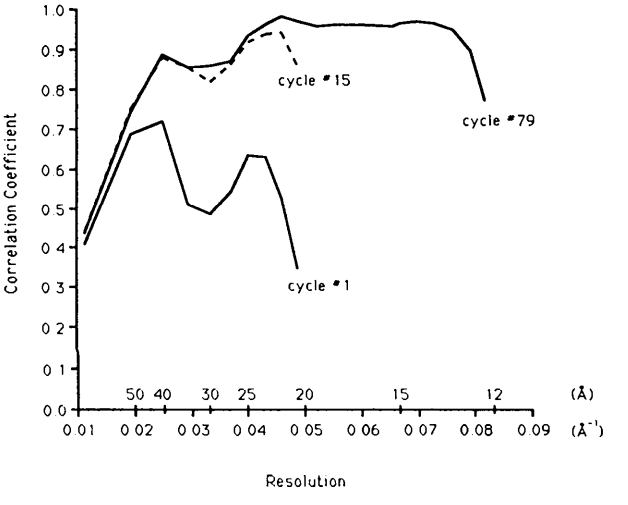
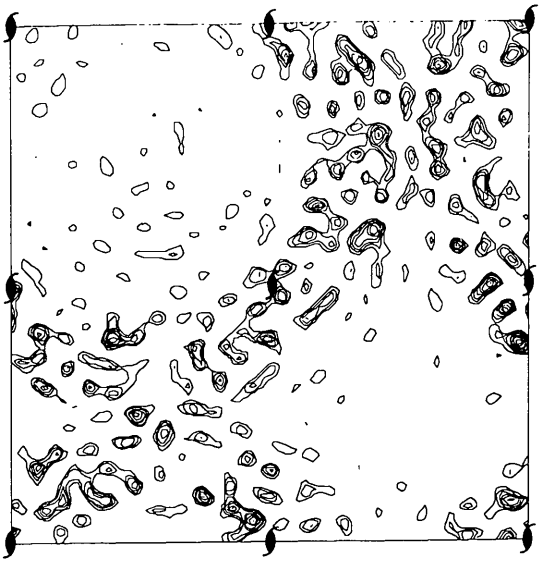


Fig. 4. Correlation coefficients, C, for phase extension from 20 to 12 Å starting with phases from a hollow-shell model with radii 149 and 105 Å.

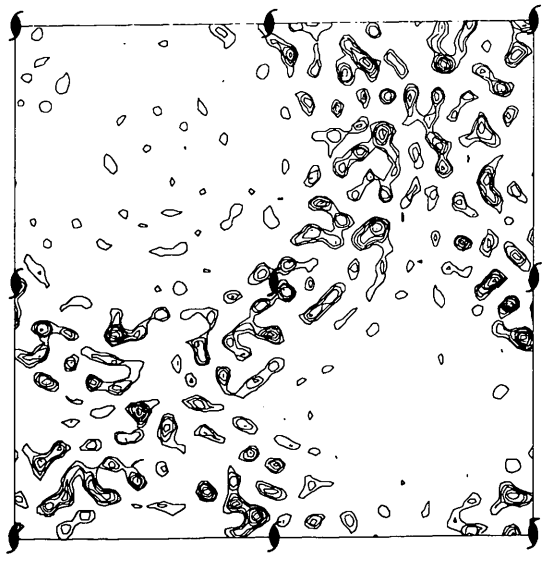
$$C = \sum (\langle F_{atom}\rangle - F_{atom})(\langle F_{b-t}\rangle - F_{b-t}) \times \left[\sum (\langle F_{atom}\rangle - F_{atom})^2 \sum (\langle F_{b-t}\rangle - F_{b-t})^2\right]^{-1}.$$

to essentially the 'observed' phasing set calculated from the atomic structure. Phases calculated with model radii of 152 and 146 Å also progressed to the correct phases but for the model radius of 146 Å convergence was impaired (Figs. 8*a*, *b*). However, a model radius of 144 Å completely failed to converge to any sensible solution (Fig. 8*c*).

The initial phases of the model with an outer radius of 139 Å correlated well with the 'observed' phases within 40 Å resolution. However, beyond 30 Å resolution the phases tended to be negatively correlated to the 'observed' phases (Fig. 9). Phase refinement and extension to 12 Å resolution enhanced this effect and eventually the agreement between derived and 'observed' phases beyond 25 Å resolution was as good as it was when the correct radius of 149 Å had been used (Fig. 9), except that they represented the opposite Babinet solution.



(*a*)



(*b*)

Fig. 5. Corresponding sections of electron density maps at 12 Å resolution (*a*) calculated from the atomic model and (*b*) calculated from F_{atom} amplitudes and phases extended from the hollow-shell model with outer radius 149 Å.

The effect of errors on structure amplitude measurements

Phase refinement and extension is largely dependent on an accurate knowledge of structure amplitudes. Previous tests were made in the absence of random error on structure amplitudes. However, in the real case, there will be error; the greater that error, the greater will be the uncertainty of the phase determination. A more realistic, F'_{atom} 'observed' structure-amplitude set was simulated by randomly introducing

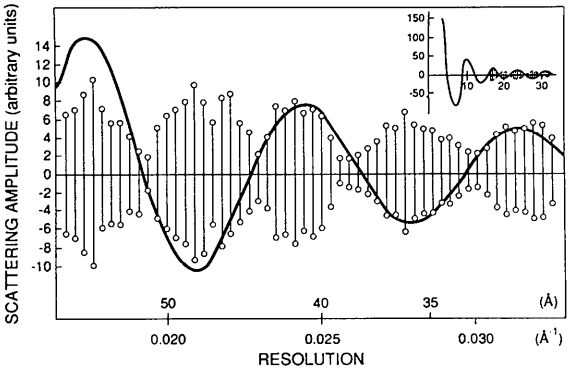


Fig. 6. Structure amplitudes of SBMV crystals averaged within shells of reciprocal space, shown in relation to the Fourier transform of a sphere of diameter 284 Å. Inset is shown the complete spherical transform from infinity to 30 Å resolution. [Reprinted with permission from Johnson *et al* (1976). Copyright by Academic Press. Inc.]

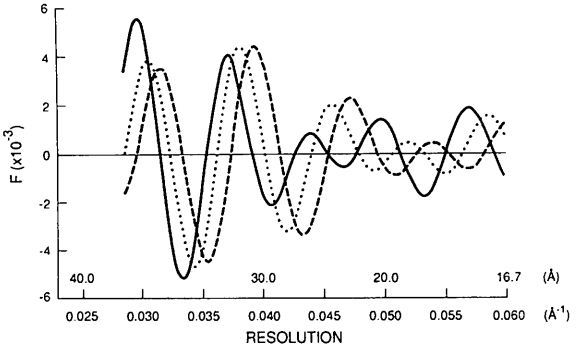


Fig. 7. The Fourier transform of hollow shells with inner and outer radii of 105 and 149 Å (solid), 105 and 144 Å (dotted) and 103 and 139 Å (dashed), respectively. Note that at 22 Å resolution the shells with outer radii of 149 and 139 Å are of opposite sign.

error on each structure factor with a random-number generator. The overall error distribution was assumed to be Gaussian, with no regard for structure amplitude or resolution. The errors were scaled to give an R factor of 13% [equivalent to the R_{merge} factor for CPV (Tsao *et al.*, 1992)] for data between 500 and 5.5 Å resolution. These F'_{atom} data were then used for phase refinement at 20 Å resolution starting from the 149 Å spherical-shell-model phases. As usual, 15 cycles of phase refinement were performed at 20 Å resolution, after which phases were extended to 16 Å resolution in the manner previously described. The final mean cosine phase differences (Fig. 10) were slightly worse than those which had been determined in the absence of error on the amplitudes.

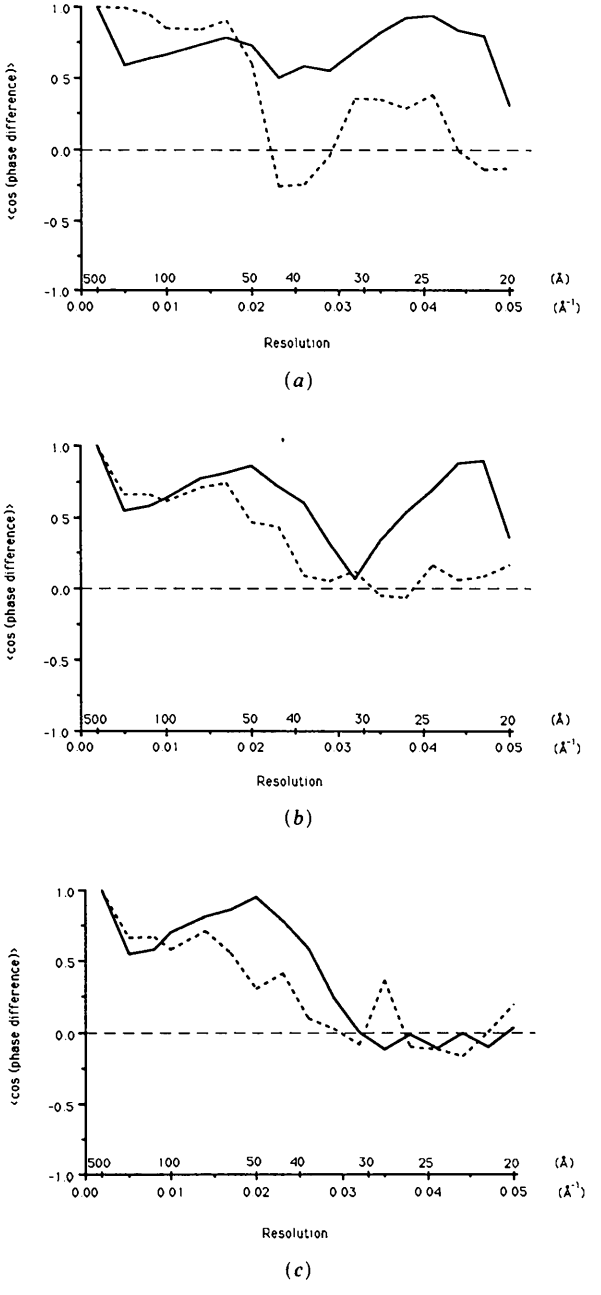


Fig. 8. Mean cosine phase difference after 15 cycles of phase refinement at 20 Å resolution starting from hollow-shell models with an inner radius of 105 Å and outer radii of (*a*) 152 (*b*) 146 and (*c*) 144 Å. The starting models with external radii of 152 and 146 Å, which were only slightly different from the correct 'observed' model, converged to the correct phase solution. However, the starting model with an outer radius of 144 Å was sufficiently in error to make it impossible to find a correct phase solution.

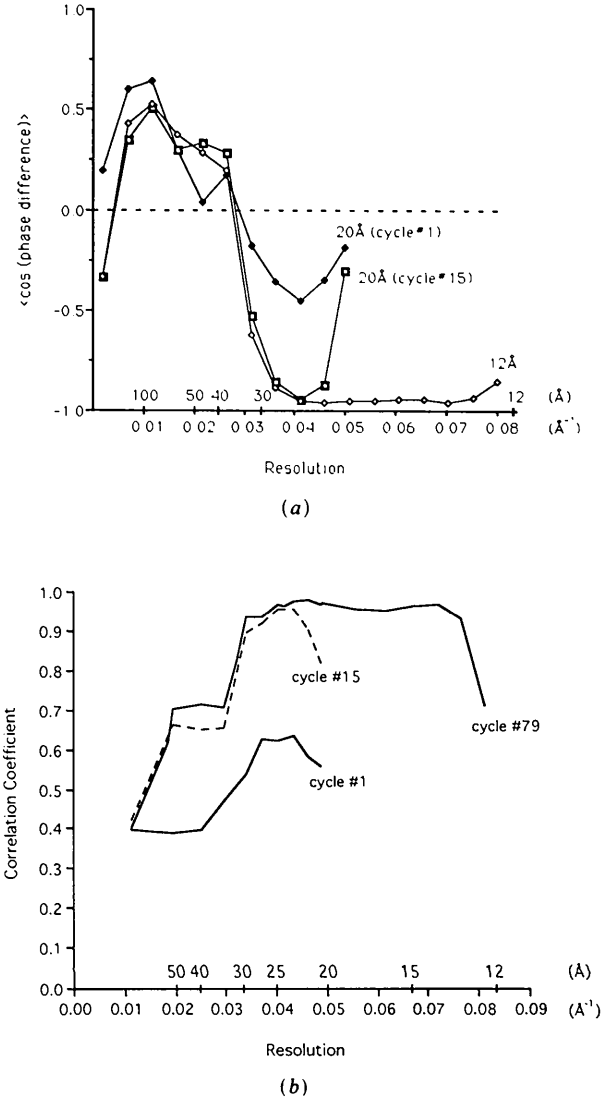


Fig. 9. (*a*) Mean cosine phase differences for phase extension from 20 to 12 Å resolution starting from a hollow-shell model with an inner radius of 102 Å and an outer radius of 139 Å. Note that the outer radius of the starting model causes a spherical transform to be exactly out of phase at about 25 Å resolution to the correct model (Fig. 7). The phases converged well, but to the opposite Babinet solution. (*b*) Correlation coefficients for the same phase extension.

The effect of using an incomplete data set

In a real structure determination it is unusual to have a complete data set to a given resolution either because some reflections are too weak for observation or (more importantly) some parts of the reciprocal space have not been explored. The latter is particularly true for the 'American method' of data collection (Rossmann & Erickson, 1983) where crystals are used in unknown orientation. A 50% complete data set was selected from the F_{atom} 'observed' data (without error) by random selection. The partial data set was used for refinement at 20 Å resolution with phases obtained from the spherical-shell model with an outer radius of 149 Å. After the first averaging, structure factors could be calculated for the missing data. These were then included in the calculation of a new electron density map. Rayment (1983) pointed out that use of calculated structure factors in place of unmeasured data provides a more accurate representation of the continuous Fourier transform of the structure than if they were omitted. Consequently, it is now general practice to include all such terms in phase refinement and extension. During the initial stages of phase refinement, the calculated terms will, of necessity, be small and, hence, they will be included with only small weights. Nevertheless, inclusion of these 'unobserved' terms in general increases the agreement of the observed and calculated structure amplitudes as measured by the correlation coefficients.

After an initial 15 cycles of molecular averaging at 20 Å resolution, the phases were extended to 12 Å resolution. The phases converged to essentially the correct solution (Fig. 11), but the mean phase difference between the derived and 'observed' phases was larger at the resolution limit than for the phases derived with the complete data set.

The effect of error on defining the molecular position

Another set of parameters important for phase extension is the orientation and the position of the virus particle in the unit cell. These parameters define the symmetry operations with which the density averaging is performed. In general, the particle orientation can be determined with sufficient accuracy by calculating an appropriate rotation function (Rossmann & Blow, 1962) to a sufficient resolution. Finding the particle position is often more problematic. Frequently, the best center can be found by searching over a restricted area for the position that minimizes the variation of electron density between noncrystallographically equivalent positions. This calculation can be repeated at successive steps in phase extension, but some initial error in particle position is likely because the map in which the density variation is minimized would have been calculated using phases which have a bias towards the previously assumed position. Phase extension at 9 Å resolution during the CPV structure determination (Tsao *et al.*, 1992) produced correlation coefficients that were rather lower than usual. The cause was suspected to be an incorrect particle center position. The following test was conducted to determine the effects of a plausible positional error.

A set of initial phases was calculated using a spherical shell model of 149 and 105 Å external and internal radii, respectively. However, the particle position was set at (0.2558, 0.2500, 0.2364), which

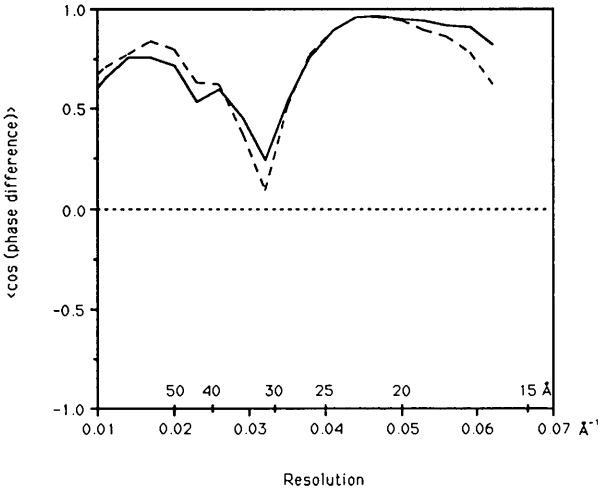


Fig. 10. Mean cosine phase differences for phase extension to 16 Å resolution using data which had a 13% R factor relative to the F_{atom} 'observed' set (dashed). The starting phasing model was a hollow shell with an outer radius of 149 Å and an inner radius of 105 Å. Similar results for error-free data are represented by the solid line.

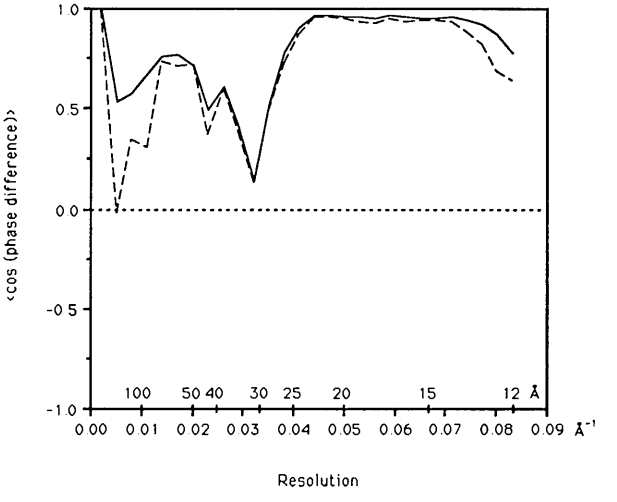


Fig. 11. Mean cosine phase differences for phase extension from 20 to 12 Å using only 50% of the available F_{atom} 'observed' data. The starting phasing model was a hollow shell with an outer radius of 149 Å and an inner radius of 105 Å. The solid line shows results for phase extension when there is no error in data.

was 1.6 Å from the correct position. This proved to be roughly the amount of displacement of the original CPV center from its final refined position. The phases were refined at 20 Å resolution assuming the displaced particle position. After the initial 15 cycles, the phases were extended to 16 Å resolution (Fig. 12), again assuming the displaced particle position. The phases converged to essentially the correct solution, but the mean phase difference between 'observed' and derived phases was 30° at 20 Å resolution. This is the largest phase difference of all the tests covered in this paper. Thus, it appears that the phase determination is particularly sensitive to particle position and increasingly so with higher resolution. These results are consistent with the experience of the CPV structure determination (Tsao *et al.*, 1992). Correlation coefficients that were declining with resolution forced the initiation of a search for isomorphous heavy-atom derivatives. Retrospective tests showed that had the particle center been correct, then the phases would have been sufficiently accurate. If the methods of positional refinement used at high resolution had been applied during extension from 20 Å resolution, then it is probable that the use of the single isomorphous replacement phases would not have been necessary.

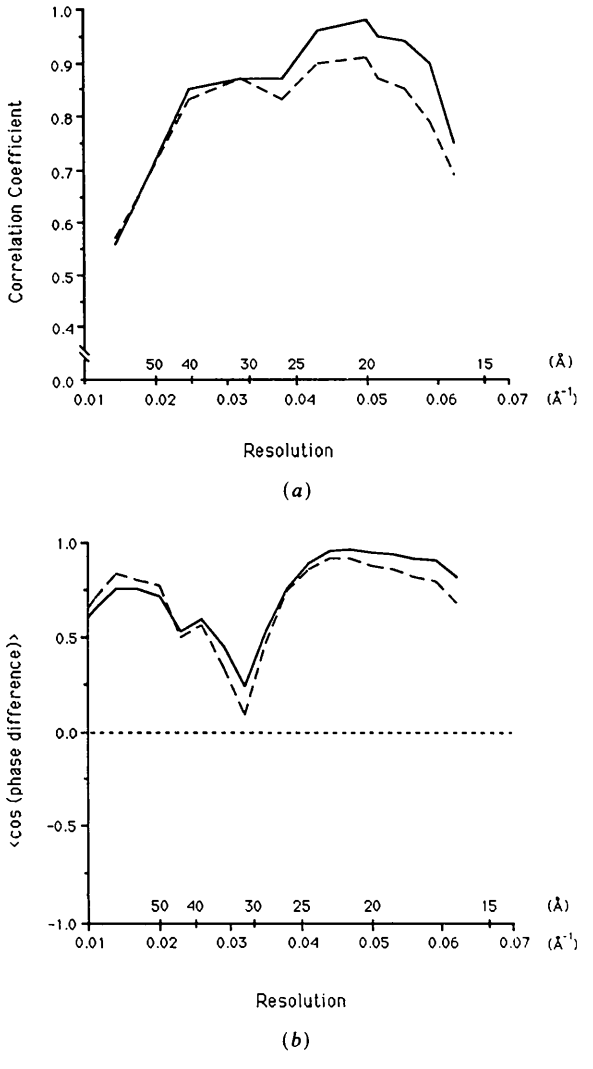


Fig. 12. The distribution (dashed) of (*a*) the correlation coefficient and (*b*) the mean cosine phase difference for phase extension from 20 to 16 Å resolution when the averaging was carried out with respect to a particle position which had a 1.6 Å error (dashed). The starting phasing model was a hollow shell with an outer radius of 149 Å and an inner radius of 105 Å. The solid line shows results for phase extension with no error in the data.

Concluding remarks

A variety of studies exist on the feasibility of phase extension from very-low-resolution data. Argos, Ford & Rossmann (1975) used experimentally observed structure amplitudes for the orthorhombic crystals of glyceraldehyde 3-phosphate dehydrogenase to determine whether the structure could have been solved by making an initial simple low-resolution assumption about the molecular structure. The molecule itself contained 222 symmetry, giving a noncrystallographic redundancy of four. It was found that reasonable phases could be determined given an accurate knowledge of the molecular orientation and position plus a rather detailed knowledge of the molecular envelope. Phasing was initiated by setting the density inside the envelope to a constant and to zero outside. In a real case, knowledge of the envelope would not be available. Here we have shown that with sufficient redundancy it is possible to obtain a useful phasing start with a simple model provided phasing is initiated at very low resolution.

Rayment (1983) examined the conditions required to determine the correct structure of polyoma virus at 22.5 Å resolution from various starting models. In this case, he wished to establish whether a variety of errors might have accounted for the unexpected appearance of pentameric building blocks in the viral capsid. His starting models contained greater detail than the spherical-shell models used here and were based on features that might be observed in electron microscopy studies. Because of this detail and because of the lack of a center of symmetry for the $T=7$ surface lattice in the polyoma virus capsid, Rayment did not have an initial centric distribution of phases. We extend his results to the use of a simpler starting model to obtain initial phases, the effect of error and quantity of observed data and the effect of error in particle position. Since Rayment's publication in 1983, it has been shown that in general, provided the correlation coefficients are better than about 0.9 at medium resolution, phase extension invariably leads to a unique, chemically correct, solution given sufficient noncrystallographic redundancy. Thus, the emphasis of the present paper is not to show that the

eventual solution is unique, but rather to explore the limits of deficiency in data and the initial phasing model to attain convergence to the correct solution.

Unlike the more complicated model used by Rayment (1983), a simple spherical-shell model was used for a virus. Similarly, a solid sphere might be used for a protein molecule with sufficiently high noncrystallographic symmetry. It was shown that the centricity of the phases could be eliminated automatically as resolution increased under the stringent conditions where the inclination between the molecular and crystallographic twofold axes was only 2.5°. It was also shown that the choice of a suitable outer radius is critically important. Although time-consuming, systematic trials using phase extension can be used to establish a suitable external radius; the internal radius is of less importance. A parallel study (Chapman, Tsao & Rossmann, 1992) shows how the radii can be determined directly from the diffraction data.

We are grateful to Jin-bi Dai, Jeffrey T. Bolin, Janet L. Smith, John E. Johnson and Craig Smith for useful discussions. We thank Sharon Wilder for typing the manuscript. The work was supported by grants from the National Institutes of Health and the National Science Foundation.

References

ABAD-ZAPATERO, C., ABDEL-MEGUID, S. S., JOHNSON, J. E., LESLIE, A. G. W., RAYMENT, I., ROSSMANN, M. G., SUCK, D. & TSUKIHARA, T. (1980). *Nature (London)*, **286**, 33-39.

ARGOS, P., FORD, G. C. & ROSSMANN, M. G. (1975). *Acta Cryst.* A**31**, 499-506.

ARNOLD, E. & ROSSMANN, M. G. (1986). *Proc. Natl Acad. Sci. USA*, **83**, 5489-5493.

ARNOLD, E., VRIEND, G., LUO, M., GRIFFITH, J. P., KAMER, G., ERICKSON, J. W., JOHNSON, J. E. & ROSSMANN, M. G. (1987). *Acta Cryst.* A**43**, 346-361.

BRICOGNE, G. (1974). *Acta Cryst.* A**30**, 395-405.

CHAPMAN, M. S., TSAO, J. & ROSSMANN, M. G. (1992). *Acta Cryst.* A**48**, 301-312.

GAYKEMA, W. P. J., HOL, W. G. J., VEREIJKEN, J. M., SOETER, N. M., BAK, H. J. & BEINTEMA, J. J. (1984). *Nature (London)*, **309**, 23-29.

HARRISON, S. C., OLSON, A. J., SCHUTT, C. E., WINKLER, F. K. & BRICOGNE, G. (1978). *Nature (London)*, **276**, 368-373.

HOGLE, J. M., CHOW, M. & FILMAN, D. J. (1985). *Science*, **229**, 1358-1365.

JOHNSON, J. E. (1978). *Acta Cryst.* B**34**, 576-577.

JOHNSON, J. E., AKIMOTO, T., SUCK, D., RAYMENT, I. & ROSSMANN, M. G. (1976). *Virology*, **75**, 394-400.

LILJAS, L., UNGE, T., JONES, T. A., FRIDBORG, K., LÖVGREN, S., SKOGLUND, U. & STRANDBERG, B. (1982). *J. Mol. Biol.* **159**, 93-108.

LUO, M., VRIEND, G., KAMER, G., MINOR, I., ARNOLD, E., ROSSMANN, M. G., BOEGE, U., SCRABA, D. G., DUKE, G. M. & PALMENBERG, A. C. (1987). *Science*, **235**, 182-191.

MCKENNA, R., XIA, D., WILLINGMANN, P., ILAG, L. L., KRISHNASWAMY, S., ROSSMANN, M. G., OLSON, N. H., BAKER, T. S. & INCARDONA, N. L. (1992). *Nature (London)*, **355**, 137-143.

RAYMENT, I. (1983). *Acta Cryst.* A**39**, 102-116.

RAYMENT, I., BAKER, T. S., CASPAR, D. L. D. & MURAKAMI, W. T. (1982). *Nature (London)*, **295**, 110-115.

ROSSMANN, M. G. (1990). *Acta Cryst.* A**46**, 73-82.

ROSSMANN, M. G., ARNOLD, E., ERICKSON, J. W., FRANKENBERGER, E. A., GRIFFITH, J. P., HECHT, H. J., JOHNSON, J. E., KAMER, G., LUO, M., MOSSER, A. G., RUECKERT, R. R., SHERRY, B. & VRIEND, G. (1985). *Nature (London)*, **317**, 145-153.

ROSSMANN, M. G. & BLOW, D. M. (1962). *Acta Cryst.* **15**, 24-31.

ROSSMANN, M. G. & ERICKSON, J. W. (1983) *J. Appl. Cryst.* **16**, 629-636.

SIM, G. A. (1959). *Acta Cryst.* **12**, 813-815.

SIM, G. A. (1960). *Acta Cryst.* **13**, 511-512.

TSAO, J., CHAPMAN, M. S., AGBANDJE, M., KELLER, W., SMITH, K., WU, H., LUO, M., SMITH, T. J., ROSSMANN, M. G., COMPANS, R. W. & PARRISH, C. R. (1991). *Science*, **251**, 1456-1464.

TSAO, J., CHAPMAN, M. S., WU, H., AGBANDJE, M., KELLER, W. & ROSSMANN, M. G. (1992). *Acta Cryst.* B**48**, 75-88.

VALEGÅRD, K., LILJAS, L., FRIDBORG, K. & UNGE, T. (1990). *Nature (London)*, **345**, 36-41.

Acta Cryst. (1992). A**48**, 301-312

Ab Initio Phase Determination for Spherical Viruses: Parameter Determination for Spherical-Shell Models

BY MICHAEL S. CHAPMAN,* JUN TSAO† AND MICHAEL G. ROSSMANN

Purdue University, Department of Biological Sciences, West Lafayette, Indiana 47907, *USA*

(*Received* 22 *July* 1991; *accepted* 5 *November* 1991)

Abstract

The structure determination of canine parvovirus depended on the extension of phases calculated initially from a spherical-shell model [Tsao, Chapman, Wu, Agbandje, Keller & Rossmann (1992). *Acta Cryst.* B**48**, 75-88]. Such *ab initio* phasing holds the promise of obviating initial experimental phasing by isomorphous or molecular replacement, thereby expediting the structure determinations of spherical virus capsids. In this paper, it is shown how parameters such as radii, DNA density and particle

* To whom correspondence should be addressed.

† Current address: Department of Biochemistry, Center for Macromolecular Crystallography, 262BHS, THT 79, University of Alabama, University Station, Birmingham, Alabama 35294, USA.

0108-7673/92/030301-12$03.00