Published in final edited form as:

Crystallogr Rev. 2016 ; 22(2): 128–140. doi:10.1080/0889311X.2015.1065489.

Application of the phase extension method in virus crystallography

Vijay S. Reddy*

Department of Integrative Structural and Computational Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA

Abstract

The procedure for phase extension (PX) involves gradually extending the initial phases from low resolution (e.g., $\sim 8\text{\AA}$) to the high-resolution limit of a diffraction data set. Structural redundancy present in the viral capsids that display icosahedral symmetry results in a high degree of non-crystallographic symmetry (NCS), which in turn translates into higher phasing power and is critical for improving and extending phases to higher resolution. Greater completeness of the diffraction data and determination of a molecular replacement solution, which entails accurately identifying the virus particle orientation(s) and position(s), are important for the smooth progression of the PX procedure. In addition, proper definition of a molecular mask (envelope) around the NCS-asymmetric unit has been found to be important for the success of density modification procedures, such as density averaging and solvent flattening. Regardless of the degree of NCS, the PX method appears to work well in all space groups, provided an accurate molecular mask is used along with reasonable initial phases. However, in the cases with space group P1, in addition to requiring a molecular mask, starting the phase extension at a higher resolution (e.g., $6\text{\AA}$) overcame the previously reported problems due to Babinet phases and phase flipping errors.

Introduction

Crystals of spherical virus capsids that display icosahedral symmetry results in a high degree of structural redundancy, at the minimum 5-fold non-crystallographic symmetry (NCS) and a maximum of 60-fold NCS or multiples of 60-fold NCS depending on the number of particles that occupy the crystallographic asymmetric unit. This redundancy is reflected in the diffraction intensities and in turn is imposed as a constraint on the structure factor

*Correspondence should be address to: Vijay S. Reddy, Department of Integrative Structural and Computational Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA, Ph.: (858) 784-8191; reddyv@scripps.edu.

Vijay Sai Reddy received a B.Sc., from the S.K.R. Govt. College, Gudur (A.P.), India and a M.Sc., in Chemistry from the University of Hyderabad, India. He was awarded a Ph.D. in 1991 for his work on protein-ligand docking studies under the guidance of Professor V.S.R. Rao from the Molecular Biophysics Unit, Indian Institute of Science, Bangalore, India. He then spent a long post-doctoral career with Professor Jack Johnson, an expert in virus crystallography, initially at Purdue University and then at The Scripps Research Institute. He was promoted to an Assistant Professor in 2000 in the Department of Molecular Biology, at The Scripps Research Institute. His significant research accomplishments are: 1) the determination of crystal structure of human adenovirus virus ($\sim 100\text{ nm}$ in diameter), the biggest biomolecule yet studied by the X-ray diffraction methods; and 2) the establishment of a spherical virus structure database, namely the V_Irus Particle Explorer database (VIPERdb), <http://viperdb.scripps.edu>. He is currently an Associate Professor in the Department of Integrative Structural and Computational Biology at The Scripps Research Institute, La Jolla, California.

amplitudes and phase angles [1]. A reciprocal space procedure to derive NCS phase relationships was proposed during the 1960s [2, 3, 4]. At that time, though feasible, this procedure turned out to be computationally expensive given the limited resources available. However, a similar procedure in real space [5, 6] became popular to use and lead to the structure determination of the first virus structure at high resolution [7, 8]. The redundancy in the NCS can be used to improve the quality of the electron density (e.d.) maps by averaging the densities of the equivalent subunits within the molecular boundary (envelope) and flattening the density outside of the envelope [6, 9, 10, 11, 12]. Back transforming the modified e.d. map results in improved phases that are more accurate than the original phases that were used to calculate the initial map. This procedure can be repeated 3 to 5 times or until there is no significant change ($< 5^\circ$) in the phases. The phases derived from the (converged) density-modified map are superior and contain slightly higher resolution information than the original unaveraged map at a given resolution. Hence, phases to higher resolution (increased by 1–2 reciprocal lattice points) can be obtained from the averaged map and combined with the associated observed structure factor amplitudes (*Fobs*) to calculate an e.d. map at correspondingly higher resolution. The resultant higher-resolution map is further improved by density averaging and solvent flattening. The first reported application of the phase extension (PX) procedure was during the structure determination of *panulirus interruptus* Haemocyanin by Wim Hol and colleagues [13, 14], where they successfully extended the phases from 4.0 to 3.2 Å resolution employing 6-fold NCS. This success was immediately followed by the structure determinations of Human rhinovirus [15, 16] and Poliovirus [17], where the authors extended the phases from 6.0 to 3.0 Å resolution. Many of the virus structures determined after that utilized the PX procedure in some form or other.

Methodology

Briefly, the steps involved in the phase extension (PX) procedure are as follows.

1. Obtain a molecular replacement solution that entails accurate determination of virus particle orientation(s) and position(s), from which the non-crystallographic symmetry (NCS) operators are derived. [Programs: *GLRF* [18], *PHASER* [19], *AMORE* [20].]
2. Employ rigid body refinement (RBR) of the molecular replacement (MR) model against the diffraction data at moderate resolution (e.g., 20–5 Å) to adjust for any variations in the size of the virus particles and calculate the initial set of phases from the RBR model. [Programs: *X-PLOR* [21], *CNS* [22]]
3. Generate a molecular mask (envelope) surrounding the RBR model that represents the NCS asymmetric unit. Alternatively, a generic wedge-shaped mask can be generated based on atoms placed on a 2–4 Å three-dimensional (3D) grid within the volume that represents a trapazoidal icosahedral (NCS) asymmetric unit connecting the points on adjacent 5-fold, 2-fold, 3-fold, 2-fold axes (or 5-fold, 3-fold, 2-fold and 3-fold axes) and with chosen inner and outer radii. The latter is particularly useful to

accommodate larger surface loops if the structure in question is significantly different from the MR model. However, having a well-defined molecular mask may be critical for the success of the PX procedure (see the discussion below). [*Programs: RAVE: MAMA* [12, 23], *CCP4: NCSMASK* [24]]

4. Density averaging and solvent flattening: electron density (obtained from initial phases) at each grid point within the molecular envelope is set to an average value of the initial densities at the corresponding NCS related grid points. Electron densities at the non-integral (NCS-related) grid points are obtained by interpolating the density values from the nearest integral grid points. The density at grid points outside the envelope is set to zero or to an average solvent density. The averaged e.d. in the molecular envelope is then expanded to generate the crystallographic asymmetric unit or the complete unit cell as necessary. [*Programs: RAVE: AVE, MAVE* [12], *CCP4: MAPROT* [24], *CCP4: DM* [25], *CNS* [22]]
5. This density is then back-transformed to obtain a set of amplitudes (F_{calcs}) and phases (α_{calcs}) corresponding to the observed structure factor amplitudes (F_{obs}). R-factors and correlation coefficients (C.C.) are calculated based on the F_{obs} and F_{calcs} for each cycle as a way to monitor the progress of the procedure. After each cycle of density averaging, at a given resolution range, improvements in the R-factors and C.C.s are expected. However, the improvements in the R-factors and C.C.s are necessary but not sufficient to assess the proper progress of the PX-procedure. Hence, it is strongly recommended that the e.d. maps be inspected to evaluate the success of the PX procedure. [*Programs: CCP4: SFALL, RSTATS* [24]]
6. The F_{obs} are then combined with newly obtained phases (α_{calcs}) using appropriate weights to generate a new e.d. map. This map is subjected to another round of density averaging by repeating the steps #4–#6, until there is no significant change in the phases between two consecutive cycles. Usually 3 to 4 rounds of the steps #4–#6 are needed for convergence. [*Programs: CCP4: SIGMAA, FFT* [24]]
7. The averaged map after convergence is back transformed (step 5) to slightly higher resolution (increased by 1–2 reciprocal lattice points usually along the longest unit cell axis) and these phases are combined with corresponding F_{obs} to calculate the e.d. map to slightly higher resolution. The map(s) at higher resolution is subjected to multiple rounds of steps #4–#7 until the high-resolution limit of the data set is reached. [*Programs: CCP4: SFALL, FFT* [24]].

A C-shell script file that executes the steps #4–#7 above in a sequence was created to perform the PX procedure between the chosen resolution limits. There exists a few self-contained programs: DM [25]; PARROT [26]; CNS [22]; and Phenix [27] that can be used for phase extension. However, the programs CNS and Phenix might require geometrically

well defined models, i.e., previously known structures or proper homology models to do the phase extension.

A review case study

The structure determination of Seneca valley virus

The above procedure was employed on a diffraction data set of Seneca valley virus 001 (SVV), whose structure we published a few years ago [28]. However, the details of the phase extension (PX) procedure employed were not reported. This article provided the opportunity to discuss the details and analysis of the aforementioned PX method. The data were processed using HKL2000 [29]. The space group of the SVV data set is H3 with two virus particles in the crystallographic asymmetric unit (6 particles in the unit cell) resulting in 40-fold NCS, 20-fold redundancy coming from each of the two unique particles. Table 1 shows the data statistics obtained from a single flash-frozen crystal. The molecular replacement program Phaser [19] was used to determine the particle orientations and positions using the polyaniline model of 1/3rd of Mengo Encephalomyocarditis virus particle containing 20 copies of the icosahedral asymmetric unit (IAU) [30, 31] (PDB-ID: 2MEV; sequence identity of 35%). According to the Phaser solution, the orientations (Euler notation) and positions (fractional coordinates) of the two particles in the crystallographic asymmetric unit are $\alpha=211.610^\circ$, $\beta=159.087^\circ$, $\gamma=180.001^\circ$, $X=0.333$, $Y=-0.333$, $Z=0.0$ and $\alpha=66.191^\circ$, $\beta=125.215^\circ$, $\gamma=45.002^\circ$, $X=-0.333$, $Y=-0.666$, $Z=0.166$, respectively. One IAU of the 2MEV polyaniline search model was placed in the SVV unit cell according to the phaser solution and subjected to rigid body refinement (RBR) using the data between 20–5Å resolution, enforcing strict 40-fold NCS constraints, employing the program X-PLOR [21]. The root mean square deviation between the starting and ending models was 1.5Å, while the model R-factor dropped by ~3% from 0.495 to 0.467.

The PX procedure was started with the data between 20.5Å and 12.26Å with starting phases calculated from the above RBR polyaniline model of Mengo Encephalomyocarditis virus placed in the SVV unit cell. A wedge-shaped mask with inner and outer radii of 90 and 165Å, respectively, was generated using the program MAMA [23]. A C-shell script file that performs the series of steps in the protocol, described in the methods section, was used to do the PX procedure by extending the phases to the next higher resolution step increased by one reciprocal lattice point. At each step, 4 cycles of density averaging and solvent flattening, back transformation and new map generation was performed prior to extending the phases to the next higher resolution step. Remarkably, the procedure was successful in extending the phases all the way to 2.1Å without any intervention. This required 106 steps of the phase extension and 3 different grid sizes to accommodate the maps beyond various resolution limits (12Å, 6Å and 3Å). The progress of the procedure was monitored using the correlation coefficients and R-factors, calculated between the F_{obs} and F_{calcs} (derived from the back-transformed e.d. maps), as a function of resolution (Figure 1). Representative electron density from the final map is shown in Figure 2.

To further test the requirements of the phasing model for the success of the PX procedure, two other models were chosen: 1) atomic model of Southern bean mosaic virus (SBMV) [32, 33] and 2) a continuous e.d. model, obtained by placing atoms (water molecules) on a

2Å 3D-grid filling the wedge-shaped volume that corresponds to the icosahedral asymmetric unit with inner and outer radii of 90 and 165Å, respectively. SBMV is a plant virus with a single coat protein subunit that displays T=3 icosahedral symmetry and has sequence similarity of 11%, 6% and 8% with VP1, VP2 and VP3 subunits of SVV. Initial phases from both these models were calculated by orienting them according to the MR solution previously obtained with the 2MEV model.

To our surprise, the PX procedure using the continuous density model (#2) proceeded smoothly resulting in an interpretable map for the region that is unique to the SVV structure (Figures 3 & 4B,C). But, the trial with the SBMV model did not fare that well, even though certain densities somewhat resemble that of the SVV structure (Figure 4D). Interestingly, the skeletonized model derived from the hard to interpret map at 3.5Å using the program BUCCANEER [34] and restarting the PX procedure did yield interpretable maps (Figure 4F). But, restarting directly with the “refined phases” from the first round of the phase extension did not do as well as restarting with the phases calculated from the skeletonized atoms, which were derived from the hard to interpret map (Figure 4E,F). This suggests that the phases from the skeletonized atoms provided satisfactory starting phases, without the phase flipping and Babinet opposite errors, analogous to using the phases from the 2MEV model. These results suggest that phases from a “random” model along with a generic wedge-shaped mask can be used to determine the structure of SVV, employing the PX procedure. However, a word of caution: such generic phasing models and masks may not work in all cases.

The requirements for successful application of the phase extension method

In addition to the reported cases with space group P1 [35, 36], there are examples of other space groups [10, 37, 38, 39], where the phase extension did not work smoothly. To evaluate the requirements for the success of the PX method, a comprehensive exercise was undertaken by revisiting the application of the PX procedure on a set of representative examples listed in Table 2. The deposited structure factor amplitudes and the coordinates were taken from the PDB [40] and VIPERdb [41], respectively. In each case, the molecular replacement (MR) solution was obtained using the program Phaser [19], rigid body refinement (RBR) was done using X-PLOR [21] and the initial phases from the corresponding RBR polyaniline models were obtained using the CCP4 suite of programs [24]. Unless otherwise explicitly mentioned (e.g., SVV), all the MR analysis was performed using the polyaniline MR probes containing multiple copies of icosahedral asymmetric unit (IAU) equivalent to the degrees of NCS as listed in the 3rd column of Table 2. The rigid body refinement was done using the program X-PLOR, enforcing strict NCS constraints according to the degrees of NCS (Table 2; 3rd column) on one copy of the IAU, oriented in the crystal cell according to the MR solution. The same PX procedure described above was applied to all the examples listed in Table 2. In addition to obtaining a molecular replacement solution and the availability of satisfactory starting phases, both of which are true for these examples, a few other parameters are found to influence the success of the PX procedure: 1) molecular mask - accurate molecular masks enhance the efficiency of density-averaging and solvent-flattening procedures; 2) the NCS redundancy - higher NCS redundancy amounts to greater phasing power [11]; and 3) the PX starting resolution – a

higher starting resolution (e.g., 6Å) for the phase extension appears to fare better than starting at low resolution (e.g., 12Å).

When the molecular mask was correctly defined, irrespective of the degree of NCS or space group symmetry, with the exception of P1, one could extend the phases from low resolution (~12Å) to highest resolution of the data set smoothly, resulting in easily interpretable maps. Interestingly, the presence of high degree of NCS (≥ 40) appears to eliminate the requirement of having a tight molecular mask (e.g., SVV (H3), CCMV (P212121), Pariacoto virus (P21)). A higher degree of NCS is possible when the particles do not occupy special positions and are asymmetrically packed in the unit cell. In these cases, the PX method can be applied successfully, again with the exception of cases in space group P1. In the example of Nudarelia capensis ω virus (space group P1), even with a proper molecular mask and a higher degree of NCS (60-fold), the PX procedure was only partially successful when extending the phases from low resolution (~12Å). Only the backbone could be traced in the resulting maps along with identification of a few side chains. However, the successful phase extension was achieved by extending the phases from 6Å resolution, instead of from 12Å (Figure 5). It appears that all three aforementioned parameters have to be fully met for the PX method to be successful in the space group P1. The errors associated with the PX procedure that manifest in the form of phase flipping and Babinet opposite solutions, which appear to propagate/accumulate at each PX step and may not get corrected in the space group P1. Perhaps there may be more of these errors in the space group P1 relative to its higher-symmetry counterparts because of the greater number of reflections. One of the ways to limit such errors is by reducing the number of steps associated with the PX procedure. Thus, starting the phase extension at a higher resolution would be helpful, particularly in the cases involving the space group P1.

Notwithstanding the differences in the degree of NCS (5-fold to 60-fold), all the cases that failed to yield a successful outcome of a readily interpretable map using a generic “wedge-shaped mask” and starting at 12Å resolution, have the particles centered at the origin (Table 2). Since the origin is considered a special position, these findings are consistent with the previous observation that difficult phase extensions are known to occur in the cases where particles occupy special positions [36]. However, using an atomic mask overcame the problems of starting the PX procedure from low resolution (12Å) in these cases, irrespective of the degree of NCS (Table 2).

Interestingly, in the cases where satisfactory molecular replacement solution and initial phases are available, one could directly apply phase refinement procedure instead of the phase extension. The phase refinement involves multiple cycles (8–12) of employing the steps #4–#6 (see methodology) by including all the reflections in the data set with the initial phases coming from the MR solution. This will likely avoid all the errors related to phase flipping and Babinet opposite solutions. These suggestions are borne out of the observation that interpretable electron density is obtained in the cases where the PX method had failed. These include the cases in Table 2: 1) Nudarelia capensis ω virus (NwV) with wedge-shaped mask, 60-fold NCS, and phases calculated from the polyaniline model of NwV; 2) Cowpea mosaic Virus (CPMV) with wedge-shaped mask, 5-fold NCS, and phases calculated from the polyaniline model of CPMV; and 3) SVV with the molecular replacement solution and

phases derived from the polyaniline model of 2MEV (Mengovirus). In all these cases, direct application of the phase refinement, using all the data with the phases derived from the MR model oriented according to the Phaser (MR) solution, yielded an interpretable electron density map. Even the phases calculated from only C-alpha atoms (one per residue) are found to be sufficient for the successful outcome. Furthermore, a proper atomic mask is not essential for the successful outcome of the phase refinement procedure, a generic wedge-shaped mask was found to be adequate. Obviously, this approach does not work if the initial phases are poor (e.g., coming from an envelope filled with random atoms) and/or there is not a satisfactory MR solution.

Conclusions

First and foremost, the success of the PX method depends on accurate determination of particle position(s) and orientation(s) as well as having satisfactory initial phases. The other requirements for the successful PX procedure are 1) a well-defined molecular mask, 2) high degree of non-crystallographic symmetry (NCS), and 3) extending phases from the highest resolution possible, assuming that one has satisfactory phases till $\sim 6\text{\AA}$ resolution. It is usually preferable to extend the phases in small steps, one reciprocal lattice point at a time along the longest unit cell axis. When starting from low resolution (e.g., $12\text{\AA}$), the PX procedure is successful when a proper molecular mask is used, even with 5-fold NCS, with the exception of space group P1.

Remarkably, the presence of a higher degree of NCS (≥ 40) eliminated the need for a well-defined mask in the case of R3, P212121 and P21 space groups. Of note, the PX procedure yielded a readily interpretable SVV map, even starting with the phases from a continuous electron density model derived by placing water molecules on a $2\text{\AA}$ 3D-grid filling the wedge-shaped volume. This suggests that the higher degree of NCS can also overcome the need for having the satisfactory starting phases. However, such a phasing model may not work in all the cases and most certainly will not be sufficient to obtain a molecular replacement solution. With the availability of over 400 virus structures in the public domain [40, 41], one should be able to find a suitable model for the molecular replacement (MR) analysis. Choosing a MR model from the same virus family or genus is a good start.

Higher degree of NCS results from the asymmetric packing of molecules within the unit cell. In the case of P1 space group, in addition to natively present 60-fold (or multiples of 60-fold) NCS and using a proper molecular mask, a higher resolution ($\leq 6\text{\AA}$) start for the PX procedure is needed for the successful outcome. The errors associated with the PX procedure, which manifest in the form of phase flipping and Babinet opposite solutions can be minimized by reducing the number of extension steps by starting the PX procedure at a higher resolution ($\leq 6\text{\AA}$), particularly it is recommended in the cases involving the space group P1. In the cases where the PX procedures did not work, skeletonizing the density at high resolution from the unsuccessful attempt and restarting the PX procedure with the phases derived from the skeletonized atoms resulted in a successful outcome.

Of note, it appears that in the case of icosahedral virus capsids, if one can obtain a molecular replacement solution using a suitable model, the resulting phases from this model are good

enough to apply the phase refinement procedure, using all the diffraction data, to obtain interpretable electron density. This success is mainly due to the presence high degree of NCS, which amounts to greater phasing power. From the CPMV example, a 5-fold NCS was sufficient to obtain an interpretable map, even though the background/solvent density was high due to the use of wedge-shaped mask. However, high degree of NCS (60-fold) resulted in a cleaner map in the example of NwV. In the cases, when a suitable MR model is not available, then *de novo* determination of molecular replacement solution is necessary using the self-rotation function analysis [18] and identifying the particle positions based on special positions and/or packing considerations. In these cases, where the initial phases are likely to be poor, the phase extension method using the density averaging and solvent-flattening procedures with a proper mask is required for the virus structure determination.

Acknowledgments

I would like to thank Drs. Jack Johnson and Michael Rossmann for the instruction and discussions on the workings of phase improvement methods over the period of a number of years, during which some of the scripts and programs that are described in the text were developed. I also would like to thank Drs. Gerard Kleywegt and Kevin Cowtan for generously sharing their programs/scripts and answering my questions regarding various features of their programs. I acknowledge the support from National Institutes of Health (NIH) grants AI070771 and AI103692 that were awarded to me in determining the crystal structures of adenovirus capsids and their complexes.

References

1. Rossmann MG, Blow DM. The detection of sub-units within the crystallographic asymmetric unit. *Acta Crystallogr.* 1962; 15:24–31.
2. Rossmann MG, Blow DM. Determination of phases by the conditions of non-crystallographic symmetry. *Acta Crystallogr.* 1963; 16:39–45.
3. Main P. Phase determination using non-crystallographic symmetry. *Acta Crystallogr.* 1967; 23:50–54.
4. Crowther RA. The use of non-crystallographic symmetry for phase determination. *Acta Crystallogr.* 1969; B25:2571–2580.
5. Bricogne G. Geometric sources of redundancy in intensity data and their use for phase determination. *Acta Crystallogr.* 1974; A30:395–405.
6. Bricogne G. Methods and programs for direct-space exploitation of geometric redundancies. *Acta Crystallogr.* 1976; A32:832–847.
7. Winkler FK, Schutt CE, Harrison SC, Bricogne G. Tomato bushy stunt virus at 5.5-Å resolution. *Nature.* 1977; 265:509–513. Epub 1977/02/10. [PubMed: 834302]
8. Harrison SC, Olson AJ, Schutt CE, Winkler FK, Bricogne G. Tomato bushy stunt virus at 2.9 Å resolution. *Nature.* 1978; 276:368–373. Epub 1978/11/23. [PubMed: 19711552]
9. Wang BC. Resolution of phase ambiguity in macromolecular crystallography. *Methods Enzymol.* 1985; 115:90–112. Epub 1985/01/01. [PubMed: 4079800]
10. Fry E, Acharya R, Stuart D. Methods used in the structure determination of foot-and-mouth disease virus. *Acta crystallographica.* 1993; 49(Pt 1):45–55. Epub 1993/01/01. [PubMed: 8382928]
11. Rossmann, MG. Determination of virus structures by the use of molecular replacement, density averaging. Warrington: Daresbury Laboratory; 1989.
12. Jones, TA. a. yaap. asap, @#*? A set of Averaging Programs. In: Dodson, EJ.; Gover, S.; Wolf, W., editors. *Molecular Replacement*. Warrington: SERC Daresbury Laboratory; 1992.
13. Gaykema WP, Hol WG, Vereijken JM, Soeter NM, Bak HJ, Beintema JJ. 3.2 Å structure of the copper-containing, oxygen-carrying protein Panulirus interruptus haemocyanin. *Nature.* 1984; 309:23–29. Epub.

14. Gaykema WP, Volbeda A, Hol WG. Structure determination of Panulirus interruptus haemocyanin at 3.2 Å resolution. Successful phase extension by sixfold density averaging. *Journal of molecular biology*. 1986; 187:255–275. Epub 1986/01/20. [PubMed: 3701867]
15. Rossmann MG, Arnold E, Erickson JW, Frankenberger EA, Griffith JP, Hecht HJ, Johnson JE, Kamer G, Luo M, Mosser AG, et al. Structure of a human common cold virus and functional relationship to other picornaviruses. *Nature*. 1985; 317:145–153. Epub 1985/09/12. [PubMed: 2993920]
16. Arnold E, Luo M, Vriend G, Rossmann MG, Palmenberg AC, Parks GD, Nicklin MJ, Wimmer E. Implications of the picornavirus capsid structure for polyprotein processing. *Proceedings of the National Academy of Sciences of the United States of America*. 1987; 84:21–25. Epub 1987/01/01. [PubMed: 3467351]
17. Hogle JM, Chow M, Filman DJ. Three-dimensional structure of poliovirus at 2.9 Å resolution. *Science*. 1985; 229:1358–1365. Epub 1985/09/27. [PubMed: 2994218]
18. Tong LA, Rossmann MG. The locked rotation function. *Acta Crystallographica A*. 1990; 46(Pt 10): 783–792.
19. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *J Appl Crystallogr*. 2007; 40:658–674. Epub 2007/08/01. [PubMed: 19461840]
20. Navaza J. AMoRe: an automated package for molecular replacement. *Acta Crystallogr*. 1994; A50:157–163. Epub.
21. Brunger, AT. X-PLOR, version 3.1: A system for X-ray crystallography and NMR. New Haven, Connecticut: Yale Univ. Press; 1992.
22. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr*. 1998; 54:905–921. Epub 1998/10/03. [PubMed: 9757107]
23. Kleywegt GJ, Jones TA. Software for handling macromolecular envelopes. *Acta Crystallogr D Biol Crystallogr*. 1999; 55:941–944. Epub 1999/03/25. [PubMed: 10089342]
24. CCP4. The CCP4 suite: programs for protein crystallography. *Acta Crystallographica D*. 1994; 50:760–763.
25. Cowtan K. dm: An automated procedure for phase improvement by density modification. *Joint CCP4 & ESF-EACBM Newsletter on Protein Crystallography*. 1994; 31:34–38. Epub.
26. Zhang KY, Cowtan K, Main P. Combining constraints for electron-density modification. *Methods Enzymol*. 1997; 277:53–64. Epub 1997/01/01. [PubMed: 18488305]
27. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr*. 2010; 66:213–221. Epub 2010/02/04. [PubMed: 20124702]
28. Venkataraman S, Reddy SP, Loo J, Idamakanti N, Hallenbeck PL, Reddy VS. Structure of Seneca Valley Virus-001: an oncolytic picornavirus representing a new genus. *Structure*. 2008; 16:1555–1561. Epub 2008/10/23. [PubMed: 18940610]
29. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol*. 1997; 276:307–326. [PubMed: 27754618]
30. Luo M, Vriend G, Kamer G, Minor I, Arnold E, Rossmann MG, Boege U, Scraba DG, Duke GM, Palmenberg AC. The atomic structure of Mengo virus at 3.0 Å resolution. *Science*. 1987; 235:182–191. Epub 1987/01/09. [PubMed: 3026048]
31. Krishnaswamy S, Rossmann MG. Structural refinement and analysis of Mengo virus. *Journal of molecular biology*. 1990; 211:803–844. Epub 1990/02/20. [PubMed: 2156078]
32. Abad-Zapatero C, Abdel-Meguid SS, Johnson JE, Leslie AG, Rayment I, Rossmann MG, Suck D, Tsukihara T. Structure of southern bean mosaic virus at 2.8 Å resolution. *Nature*. 1980; 286:33–39. Epub 1980/07/03. [PubMed: 19711553]
33. Silva AM, Rossmann MG. Refined structure of southern bean mosaic virus at 2.9 Å resolution. *Journal of molecular biology*. 1987; 197:69–87. Epub 1987/09/05. [PubMed: 3681993]

34. Cowtan K. Fitting molecular fragments into electron density. *Acta Crystallogr D Biol Crystallogr*. 2008; 64:83–89. Epub 2007/12/21. [PubMed: 18094471]
35. Plevka P, Kaufmann B, Rossmann MG. Analysis of phases in the structure determination of an icosahedral virus. *Acta Crystallogr D Biol Crystallogr*. 2011; 67:568–577. Epub 2011/06/04. [PubMed: 21636897]
36. Munshi S, Liljas L, Johnson JE. Structure determination of Nudaurelia capensis omega virus. *Acta Crystallogr D Biol Crystallogr*. 1998; 54:1295–1305. [PubMed: 10089506]
37. Valegard K, Liljas L, Fridborg K, Unge T. Structure determination of the bacteriophage MS2. *Acta crystallographica Section B, Structural science*. 1991; 47(Pt 6):949–960. Epub 1991/12/01. [PubMed: 1772631]
38. Larson SB, Koszelak S, Day J, Greenwood A, Dodds JA, McPherson A. Three-dimensional structure of satellite tobacco mosaic virus at 2.9 Å resolution. *Journal of molecular biology*. 1993; 231:375–391. Epub 1993/05/20. [PubMed: 8510153]
39. McKenna R, Xia D, Willingmann P, Ilag LL, Rossmann MG. Structure determination of the bacteriophage phiX174. *Acta crystallographica Section B, Structural science*. 1992; 48(Pt 4):499–511. Epub 1992/08/01. [PubMed: 1418820]
40. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. *Nucleic acids research*. 2000; 28:235–242. Epub 1999/12/11. [PubMed: 10592235]
41. Carrillo-Tripp M, Shepherd CM, Borelli IA, Venkataraman S, Lander G, Natarajan P, Johnson JE, Brooks CL 3rd, Reddy VS. VIPERdb2: an enhanced and web API enabled relational database for structural virology. *Nucleic acids research*. 2009; 37:D436–D442. Epub 2008/11/05. [PubMed: 18981051]
42. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry*. 2004; 25:1605–1612. Epub 2004/07/21. [PubMed: 15264254]
43. Lin T, Chen Z, Usha R, Stauffacher CV, Dai JB, Schmidt T, Johnson JE. The refined crystal structure of cowpea mosaic virus at 2.8 Å resolution. *Virology*. 1999; 265:20–34. Epub 1999/12/22. [PubMed: 10603314]
44. van den Worm SH, Stonehouse NJ, Valegard K, Murray JB, Walton C, Fridborg K, Stockley PG, Liljas L. Crystal structures of MS2 coat protein mutants in complex with wild-type RNA operator fragments. *Nucleic acids research*. 1998; 26:1345–1351. Epub 1998/04/04. [PubMed: 9469847]
45. Speir JA, Bothner B, Qu C, Willits DA, Young MJ, Johnson JE. Enhanced local symmetry interactions globally stabilize a mutant virus capsid that maintains infectivity and capsid dynamics. *Journal of virology*. 2006; 80:3582–3591. Epub 2006/03/16. [PubMed: 16537626]
46. Chandrasekar V, Johnson JE. The structure of tobacco ringspot virus: a link in the evolution of icosahedral capsids in the picornavirus superfamily. *Structure*. 1998; 6:157–171. Epub 1998/03/31. [PubMed: 9519407]
47. Tang L, Johnson KN, Ball LA, Lin T, Yeager M, Johnson JE. The structure of pariacoto virus reveals a dodecahedral cage of duplex RNA. *Nature structural biology*. 2001; 8:77–83. Epub 2001/01/03. [PubMed: 11135676]

Figure 1.

Plot showing the progress of the phase extension procedure using the starting phases from the 2MEV polyaniline model and a wedge-shaped mask having inner and outer radii of 90Å and 165Å, respectively. The overall R-factors (gray circles) and correlation coefficients (black squares) between F_{obs} and F_{calcs} calculated by back-transformation of density-averaged maps at the end of each extension step are plotted as a function of resolution (Å).

Figure 2.

A) Superposition of structures of Seneca valley virus (PDB-ID: 3CJI, in green) and Mengo Encephalomyocarditis virus (PDB-ID: 2MEV; in yellow). The loop containing residues 173–177 of VP2 that are unique to SVV are shown in red. B) Diagram showing the electron density at 2.1 Å for the unique residues 173–177 of SVV-VP2. The density was obtained at the end of the phase extension procedure starting with the phases from the polyaniline model of the Mengo virus structure (2MEV). The figure was generated using the program Chimera [42].

Figure 3.

Plot showing the progress of the phase extension procedures starting with different phasing models. Correlation coefficients (CC) between *Fobs* and *Fcalcs* calculated by back-transformation of density-averaged maps at the end of each extension step are plotted as a function of resolution (Å), up to 3.5Å. Different models included in this exercise are: 1) polyaniline model of Mengovirus (PDB-ID: 2MEV) (black line) that gave the best final statistics (Rfactor (Rf)=0.211, CC=0.865) and interpretable density; 2) polyaniline model of SBMV (PDB-ID: 4SBV) (orange line) with poor final statistics (Rf=0.229, CC=0.838) and poor density; 3) skeletonized model of the latter 4SBV derived density (gray line) with good final statistics (Rf=0.213, CC=0.862) and interpretable density; 4) the continuous electron density model derived by placing water molecules on a 3D-grid with 2Å spacing, placed within the wedge-shaped volume (dashed red line) with good final statistics (Rf=0.214, CC=0.861) and interpretable density; and 5) skeletonized atoms derived from the latter density (dashed blue line) that gave the best statistics (Rf=0.211, CC=0.865), comparable to that of using the 2MEV model and interpretable density.

Figure 4.

Diagrams showing the electron density at 3.5 Å resolution for the distinct residues 173–177 found in the VP2 subunit of SVV obtained using the phase extension procedure with starting phases from different models. A) The best (prototype) density obtained by using the starting phases from the polyalanine model of Mengo virus (2MEV). B) Density obtained by using starting phases from the continuous density model generated by placing water molecules on a (2 Å spaced) three-dimensional grid within the wedge-shaped volume. C) Density obtained by using the starting phases from skeletonized atoms of the map in panel B. D) Density obtained by using the starting phases from the Southern bean mosaic virus (SBMV) model.

E) Density obtained by restarting the phase extension with the refined phases of the map shown in panel D. F) Density obtained by using the starting phases from the skeletonized atoms of the density in panel D. Phases from the skeletonized atoms of the density fared better than using the refined phases that yielded the density. The figure was generated using the program Chimera [42].

Figure 5.

Diagrams showing the electron density (e.d) at 3.0Å resolution for the residues 59–65 of the B-subunit of Nudarelia capensis ω virus (NwV), obtained by the PX procedures using different masks or starting resolutions. A) E.d. at 3Å resolution, obtained by employing the PX procedure starting at 12Å resolution with a generic wedge-shaped mask having the inner and outer radii of 90Å and 220Å, respectively. Initial phases were calculated from a polyalanine model based on the refined coordinates of NwV (PDB-ID: 1OHF). B) E.d. at 3Å resolution, achieved by employing the PX procedure starting at 6Å resolution and with the

generic wedge-shaped mask as in the panel A. C) E.d. at 3Å resolution, obtained by employing the PX procedure starting at 12Å resolution with an atomic mask derived from the NwV coordinates (PDB-ID: 1OHF) and the polyalanine phases as described in the panel A. D) The best e.d. at 3Å resolution was achieved by employing the PX procedure starting at 6Å resolution employing the atomic mask as in panel C. The figure was generated using the program Chimera [42].

Table 1

Data statistics

Space group	H3; #146
Unit cell dimensions	a=b=311.42 Å, c=1526.49 Å, $\alpha=\beta=90^\circ$, $\gamma=120^\circ$
Resolution range (Å)	20.5 – 2.11 (2.23–2.11) [§]
Rmerge	0.166 (0.391) [§]
Completeness (%)	69.0 (29.2) [§]
$\langle I \rangle / \langle \sigma(I) \rangle$	5.9 (1.1) [§]

[§]Statistics for the highest resolution bin are shown in parenthesis

Table 2

Parameters and the results for the success of phase extension

Space group	Virus name	Degrees of NCS	Details of the phase extension			
			Start Res (Å)	End Res (Å)	Starting phases [‡]	Successful using atomic mask
I23 [‡]	Cowpea mosaic Virus (1NY7)[43] [§]	5	12.0	3.0	Polyala	No [¶] ; 89.2° [¥]
H32 [‡]	Bacteriophage MS2 (1AQ3)[44] [§]	10	12.0	3.0	Polyala	Yes [§] ; 49.3° [¥]
H3 [‡]	Flock house virus (4FTE) [§]	20	12.0	3.5	Polyala	Yes [§] ; 48.6° [¥]
H3	Seneca valley virus (3CJ1)[28] [§]	40	12.0	3.5	Polyala	Yes [§] ; 54.9° [¥]
P 2 1 2 1	Cowpea Chlorotic Mottle Virus (1ZA7)[45] [§]	60	12.0	3.0	Polyala	Yes [§] ; 60.7° [¥]
C 1 2 1 [‡]	Tobacco Ringspot Virus (1A6C)[46] [§]	30	12.0	3.0	Polyala	Yes [§] ; 46.4° [¥]
P 1 2 1	Pariacoto Virus (1F8V)[47] [§]	60	12.0	3.0	Polyala	Yes [§] ; 56.4° [¥]
P1 [‡]	Nudarelia capensis ω virus (1OHE)[36] [§]	60	12.0	3.0	Polyala	Yes [§] ; 41.0° [¥]
		60	6.0	3.0	Polyala	Backbone can be traced [*] ; 62.8° [¥]
						No [¶] ; 84.5° [¥]
						Backbone may be traced ^{**} ; 84.5° [¥]

[§] Shown in parenthesis are the PDB-IDs of the respective virus. Deposited structure factors and coordinates were obtained from the RCSB-PDB [40] and VIPERdb [41], respectively.

[‡] Molecular replacement analysis and starting phases were calculated using the polyalanine model derived from the respective PDB coordinates using Phaser [19] and CCP4 suite of programs [24].

[§] All of the backbone can be traced and most of the side chains can be readily identified

[¶] Not even backbone can be traced.

^{*} Majority of the polypeptide backbone can be traced and a few large side chains identified.

^{**} Polypeptide backbone may be traced with some difficulty

[‡] Particles are centered at the origin.

[¥] Phase difference between the initial phases and the phases at the end of the phase extension (PX) for all the reflections to the highest resolution limit specified in the column 5. Please note that in the cases, where the PX method did not work have phase difference close to 90°. However, in the PX trials of SVV that yielded interpretable e.d. maps, where the initial phases were calculated from either the continuous electron density model or the SBMV-skeletonized model (2nd attempt) produced the phase differences also close to 90°.