J. Synchrotron Rad. (1999). **6**, 816–821

Synchrotron radiation as a tool for investigating virus structures

Michael G. Rossmann

Department of Biological Sciences, Purdue University, West Lafayette, IN 47907-1392, USA. E-mail: mgr@indiana.bio.purdue.edu

(*Received 8 February 1999; accepted 8 April 1999*)

Synchrotron radiation is critically important to the determination of virus structures. Methods of data reduction and phase determination are explained. Some examples are mentioned.

Keywords: viruses; X-ray diffraction; auto-indexing.

1. Introduction

The smallest spherical viruses have a diameter of at least 200 Å [*e.g.* satellite tobacco necrosis virus (Liljas *et al.*, 1982) and satellite tobacco mosaic virus (Larson *et al.*, 1993)]. Many spherical viruses, such as southern bean mosaic virus (Abad-Zapatero *et al.*, 1980) or rhinoviruses (Rossmann *et al.*, 1985), have diameters of ∼300 Å, and many other viruses have much larger diameters [*e.g.* bluetongue virus (Grimes *et al.*, 1998)], with iridiviruses (Van Etten *et al.*, 1991) having diameters of as much as 2000 Å. Because of their spherical shape, these viruses are usually fairly easy to crystallize but, because of their size, the resolution, indexing and intensity measurements of the individual Bragg reflections create severe experimental problems. Fortunately, the use of synchrotron radiation (Figs. 1 and 2) has made it possible to solve these problems in most cases.

2. Data collection

Synchrotron radiation offers a number of advantages for the recording of diffraction patterns of viruses:

(i) The large unit cells mean that the number of Bragg reflections is very large, often requiring the recording of millions of reflections for a medium resolution (∼3 Å) data set. In turn, this implies that a larger number of exposures are required, each with only a small oscillation angle (0.15–0.50°, depending on the unit-cell size). Furthermore, the large unit-cell size means that each reflection is weaker than that of, for example, a protein crystal in proportion to the volume of the cell. The higher intensity of synchrotron radiation thus allows data collection of a complete data set in a period of days instead of years.

(ii) High-intensity synchrotron radiation allows data to be collected on unstable intermediates which alter the structure of the virus in a period of minutes. For instance, low pH destabilizes rhinoviruses but, nevertheless, exposures of less than 1 min allowed the analysis of the effect of pH on human rhinovirus 14 (HRV14) (Giranda *et al.*, 1992). Laue photography and suitable undulators can also be used effectively for reducing exposure times.

(iii) Radiation can be better controlled by high-intensity synchrotron radiation, both because less damaging shorter-wavelength radiation can be selected and because radiation damage is often time dependent. An excellent example is that of the structure determination of Mengo virus (Luo *et al.*, 1987), where a single short exposure produced an excellent diffraction pattern, leaving the crystal completely dead for any subsequent exposure. Although frozen crystals are far more stable, radiation damage is still a problem. In general, intense synchrotron radiation helps to keep the damage to a minimum because of the shorter exposure, but the radiation also heats the crystal, which may be a problem.

(iv) The quality of the data can be much improved using synchrotron radiation compared with the results using much weaker laboratory X-ray sources. This is not only due to the reduced radiation damage but also to reduced background. The latter is due to the shorter exposure times, the cleaner monochromated radiation, the tighter focal spots, and the possibility of using reduced oscillation angles. The synchrotron radiation, therefore, facilitates large unit-cell data collection.

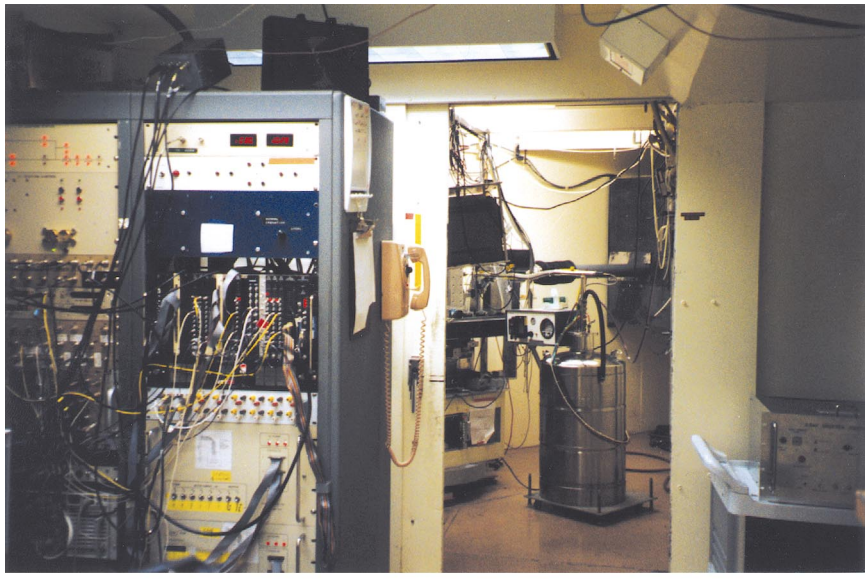
(v) As a consequence of the improved quality of diffraction patterns recorded with synchrotron radiation, the limit of resolution is often greatly enhanced. Carefully selected 'Fankuchen'-cut monochromating crystals greatly improve the energy resolution of the selected X-ray beam, an essential requirement for the large unit cells of virus crystals.

(vi) The ability to select a specific monochromated radiation from the white synchrotron radiation to collect multiple-wavelength anomalous dispersion (MAD) data is another very important advantage of synchrotron radiation, although its use for virus structure determination is not so relevant.

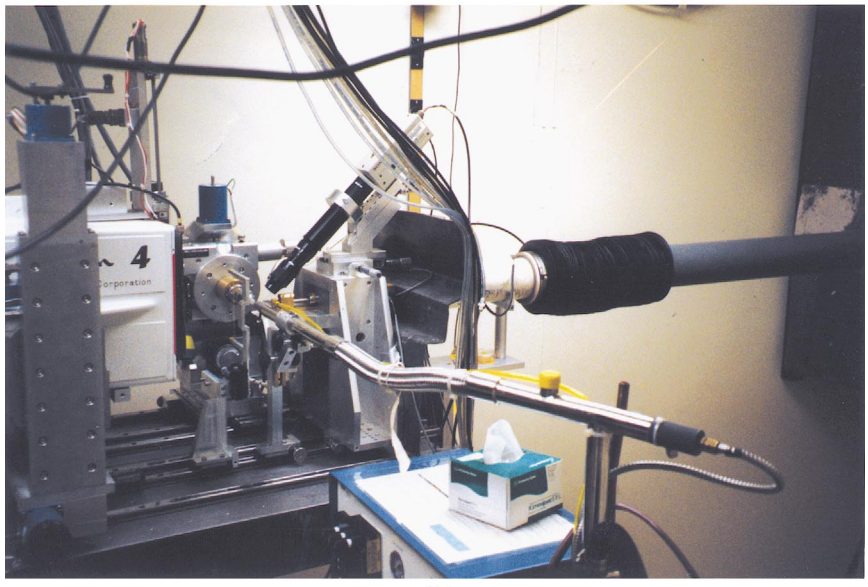
3. Data reduction

The high throughput of recorded diffraction data requires considerable automation of data processing and reduction to structure amplitudes. There is clearly an associated demand on computer storage and performance. Data reduction proceeds in a number of stages:

(i) *Auto-indexing.* The shortage of available synchrotron beam time, the need for reducing radiation damage, and the difficulty of optically orienting frozen crystals implies that, in general, crystals are mounted in a random orientation. The indexing, therefore, must be determined directly from the diffraction pattern, a process of 'shooting first and thinking later' dubbed the 'American method' (Rossmann & Erickson, 1983). A variety of early techniques were developed for this purpose (Vriend & Rossmann, 1987; Kabsch, 1988; Kim, 1989) but none were entirely satisfactory. The program *DENZO* (Otwinowski & Minor, 1997) revolutionized the procedure in providing a highly stable, reliable and automatic procedure, capable of suggesting likely Bravais lattices. Unfortunately, the nature of the algorithm has never been revealed, and the computer source code is not made available. However, a procedure at least as reliable has now been established (Steller *et al.*,



(*a*)



(*b*)

Figure 1
View of the new F1 beamline containment facility at the Cornell High Energy Synchrotron Source (CHESS). (*a*) General view through the hutch door. Note the incoming beamline at the back of the hutch and the liquid-nitrogen dewar. (*b*) More detailed view of the camera (Thiel *et al.*, 1998) with the Quad4 ADSC detector on the left, the black TV camera for viewing the crystal at a 45° angle, and the cold nitrogen liquid stream delivery tube on the right. (These photographs were taken by Dan Thiel at CHESS.)

1997) and is available as part of the Data Processing Suite (DPS) of programs and as part of *MOSFILM* (Collaborative Computational Project, Number 4, 1994).

(ii) *Pre-refinement.* Auto-indexing is dependent upon the measurement of the position of recorded reflections, as well as many camera parameters, such as the crystal-to-detector distance, pixel size, assumption of the X-ray beam being normal to the detector surface, and other difficult-to-measure parameters. It is, therefore, necessary to refine these parameters using not only the reflection positions but also their intensities (Rossmann, 1979). This procedure establishes crystal orientation defined by the orientation matrix $[A]$, where

$$\mathbf{x} = [A]\mathbf{h}$$

and $\mathbf{x}$ is a position of the Bragg reflection with indices $\mathbf{h}$ in reciprocal space defined with respect to the camera axes and X-ray beam. In turn, $\mathbf{X}$ is defined with respect to the camera axes by the matrix $[Q]$, where

$$\mathbf{X} = [Q]\mathbf{R}$$

and $\mathbf{R}$ are the coordinates of the reflection with respect to the scan directions on the detector.

(iii) *Integration.* With the $[A]$ and $[Q]$ matrices, it is then possible to predict the position of every full and partial reflection, as well as the degree of partiality. This then allows the measurement of intensity of each recorded reflection, either by integration or profile fitting (Rossmann, 1979). The average profile can be learned by taking a weighted average of the reflections surrounding reflections after background subtraction. This procedure should now be improved for partial reflections that are far more numerous for crystals with large mosaic spread, as is the case for frozen crystals. Furthermore, frozen crystals would now allow the determination of three-dimensional profiles.

(iv) *Post-refinement, scaling and averaging of data.* The process of scaling data together from differently recorded diffraction images has been discussed by Hamilton *et al.* (1965) and Fox & Holmes (1966). However, these techni-

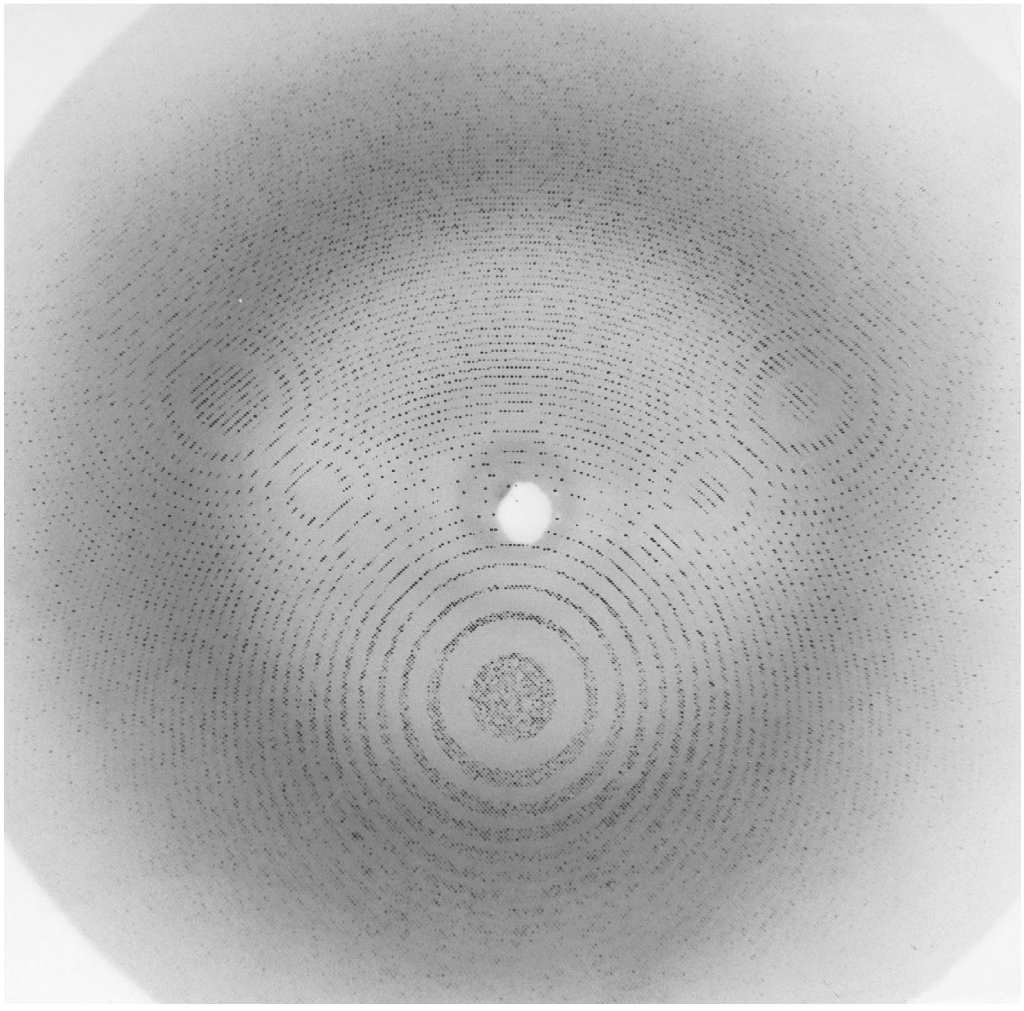


Figure 2
0.3° oscillation photograph of an HRV14 crystal taken on the old A1 beamline at CHESS. The highest-resolution reflections are 1.5 Å at the corners, and the resolution cut-off at the edge, along the spindle axis, is 3.0 Å. [Reprinted with permission from Arnold *et al.* (1987). Copyright by the International Union of Crystallography.]

ques do not permit the use of partial reflections. Because these dominate when frozen crystals are used, the Hamilton, Rollett and Sparks procedure needed modification (Bolotovsky *et al.*, 1998). Once the scale factors have been established, it is possible to perform 'post-refinement' of cell dimensions, crystal orientation and mosaic spread, a procedure introduced by Winkler *et al.* (1979). This depends only upon intensities and is completely independent of camera parameters. It is, therefore, a highly accurate procedure and one which is essential for successful electron density averaging, for, if the cell parameters are inaccurate, so will be the superposition of densities thought to be equivalent. Post-refinement does require knowledge of the relative scale factors between images (Rossmann *et al.*, 1979); thus, it has been usual to perform alternate cycles of scale factor and crystal parameter refinement. However, the procedure of Bolotovsky *et al.* (1998) does not explicitly differentiate between these two procedures. When scaling and post-refinement have converged, the data available for each independent reflection need to be analyzed for outliers and then averaged. Error estimates of intensities can be determined, both from the quality of the profile fitting and the divergence of the independent measurements of any reflections. Methods based on statistical estimates or error from the counts recorded for reflections (apparently used by *DENZO* and *SCALEPACK*) are not at all reliable and omit the obvious possibility of comparing observations.

4. Phase determination

Spherical viruses invariably have icosahedral (532) symmetry for reasons first recognized by Crick & Watson (1956, 1957). As fivefold symmetry cannot be incorporated into a periodic crystal lattice, there will always occur at least a fivefold redundancy of information in the diffraction pattern. However, quite often the whole virus particle [or even two particles (Muckelbauer *et al.*, 1995)] forms the crystallographic asymmetric unit. Thus, there is always ample non-crystallographic symmetry (NCS) redundancy in crystals of icosahedral viruses. Rossmann & Blow (1962) first pointed out the importance of this feature for the determination of the structure of viruses.

The NCS can be used to find the orientation of each virus particle in the crystal unit cell by means of a rotation function (Rossmann & Blow, 1962; Tong & Rossmann, 1990). Finding the position of a virus particle within the cell can be more difficult. If a virus twofold axis happens to be parallel to a crystallographic evenfold axis, then a Patterson function will readily determine the particle position. Otherwise, a search with a homologous virus of known structure or the sites of icosahedrally distributed heavy atoms may be necessary.

NCS is particularly powerful for improving initial phase determination by electron-density averaging. It can also be used for gradual phase extension from low resolution to the limit of the available data. Initial phasing could start from isomorphous replacement information (Arnold *et al.*, 1987), a homologous viral structure (Luo *et al.*, 1989), a low-resolution electron microscope structure (Rayment *et al.*, 1983; Ban *et al.*, 1998), an atomic model based on the structure of viral capsid proteins (Grimes *et al.*, 1998), or simply a hollow shell model (Tsao *et al.*, 1992).

Assumption of the NCS can help in determining the position of heavy atoms (Arnold *et al.*, 1987). The use of heavy metal clusters is also a powerful phasing start, as was performed to help in the determination of a ribosomal subunit (Ban *et al.*, 1998).

Although anomalous dispersion has become a very powerful tool in the determination of many protein structures, this has not yet been useful for the structure determination of viruses. That is because the impact of an anomalous scatterer on the structure amplitudes is likely to be very small. Nevertheless, such anomalous scatterers will be distributed in an icosahedral fashion and, hence, the impact need only be considered in proportion to the molecular weight of the non-crystallographic asymmetric unit, perhaps only 30000 Da in a favorable case. Thus, it seems possible that, in some future problems, Se might be used in the form of SeMet amino acids (Hendrickson, 1991; Smith, 1997) as an initial MAD phasing start in conjunction with NCS averaging.

It seems unlikely that direct methods might be used for phasing virus structures. The NCS is a very powerful tool which can extend accurate phasing from very low resolution (*e.g.* 20 Å) to the limit of available data. Thus, the importance of alternative phasing techniques is in obtaining a low-resolution phasing start. However, direct methods become useful only at resolutions where atoms can be resolved. Nevertheless, direct methods might be very useful, in conjunction with NCS imposition, to find the site of heavy atoms in isomorphous derivatives or the site of anomalous scatterers (Doublié *et al.*, 1998).

5. Results

Probably the first virus structure determination that depended heavily upon the use of synchrotron radiation was that of HRV14 (Rossmann *et al.*, 1985). However, that was quickly followed by many other similar determinations, including those of Mengo virus (Luo *et al.*, 1987), cowpea mosaic virus (Stauffacher *et al.*, 1987), canine parvovirus (Tsao *et al.*, 1991), simian virus 40 (Liddington *et al.*, 1991), MS2 (Valegård *et al.*, 1990) and many others. The most outstanding examples are those of the bluetongue virus core particles with an external diameter of 800 Å (Grimes *et al.*, 1998) and a complex of HRV14 with an Fab fragment of a neutralizing antibody (Smith *et al.*, 1996).

The biological significance of these results has justified the technical developments. There have been extensive studies, particularly by Smith (Smith *et al.*, 1996), on the mechanisms by which viruses are neutralized by antibodies. The interaction of viruses with their cellular receptors has also been extensively studied (Olson *et al.*, 1993; Rossmann,

1994; Bella *et al.*, 1998), leading to the canyon hypothesis. While the latter has been successful in predicting the site of interaction of receptor with virus in rhino- and polioviruses, its general applicability remains a topic of considerable discussion and experimentation. The structure of viruses has also shown that, surprisingly, many viruses must have had a common evolutionary precursor. That is because a great many viruses (including ssRNA, dsRNA, ssDNA and dsDNA viruses with animal, plant and bacteria as hosts) have an eight-stranded antiparallel β-barrel as a structural motif for forming their capsids. It is unlikely that this structure has evolved independently on numerous different occasions.

The single traditional and highly successful method of combating viral diseases has been through the development of vaccines. Not only have virus structures provided some understanding on the nature of vaccines, but it has been possible to use the available knowledge of virus structures to develop new vaccines, for instance against human immunodeficiency virus (Arnold *et al.*, 1996). The newer alternative to vaccines is the development of antiviral agents that target specific stages of the viral life cycle. Knowledge of virus capsid structures has provided a basis for the design of anti-rhinoviral and anti-enteroviral drugs that bind to the viral capsid. These drugs are currently in a final stage of testing (phase 3) against common cold infections, as well as more serious enterovirus.

6. ATPase

This issue of the *Journal of Synchrotron Radiation* is dedicated to John Walker and his Nobel Prize for the investigation of ATPase. The structure determination of this remarkable molecule (Abrahams *et al.*, 1994) strongly depended upon the lessons learned in the structural analyses of viruses, briefly summarized here.

The contributions made at Purdue University in the last two decades to the development of techniques for using synchrotron radiation in the study of virus structure determinations would not have been possible without the outstanding support of many individuals at various synchrotron laboratories. In particular, I would like to thank the staff of the Cornell High Energy Synchrotron Source, especially Keith Moffat, Don Bilderback, Wilfried Schildkamp, Steve Ealick and Dan Thiel; John Helliwell at the Daresbury synchrotron; Keith Wilson at the DESY synchrotron in Hamburg; Bob Sweet at the National Light Source at the Brookhaven National Laboratory; and Paul Phizackerley at the Stanford Synchrotron Radiation Laboratory. The work has been continuously supported by grants from the National Institutes of Health and the National Science Foundation, as well as a grant from the former Sterling Winthrop Company with matching funds from Purdue University. Finally, I would like to thank Sharon Wilder for her outstanding assistance over nearly all the time that I have been at Purdue University.

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.
Abrahams, J. P., Leslie, A. G. W., Lutter, R. & Walker, J. E. (1994). *Nature (London)*, **370**, 621–628.
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.
Arnold, G. F., Resnick, D. A., Smith, A. D., Geisler, S. C., Holmes, A. K. & Arnold, E. (1996). *Intervirology*, **39**, 72–78.
Ban, N., Freeborn, B., Nissen, P., Penczek, P., Grassucci, R. A., Sweet, R., Frank, J., Moore, P. B. & Steitz, T. A. (1998). *Cell*, **93**, 1105–1116.
Bella, J., Kolatkar, P. R., Marlor, C. W., Greve, J. M. & Rossmann, M. G. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 4140–4145.
Bolotovsky, R., Steller, I. & Rossmann, M. G. (1998). *J. Appl. Cryst.* **31**, 708–717.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Crick, F. H. C. & Watson, J. D. (1956). *Nature (London)*, **177**, 473–475.
Crick, F. H. C. & Watson, J. D. (1957). Ciba Foundation Symposium on *The Nature of Viruses*, edited by G. E. W. Wolstenholme & E. C. P. Millar, pp. 5–13. Boston: Little Brown & Co.
Doublié, S., Tabor, S., Long, A. M., Richardson, C. C. & Ellenberger, T. (1998). *Nature (London)*, **391**, 251–258.
Fox, G. C. & Holmes, K. C. (1966). *Acta Cryst.* **20**, 886–891.
Giranda, V. L., Heinz, B. A., Oliveira, M. A., Minor, I., Kim, K. H., Kolatkar, P. R., Rossmann, M. G. & Rueckert, R. R. (1992). *Proc. Natl Acad. Sci. USA*, **89**, 10213–10217.
Grimes, J. M., Burroughs, J. N., Gouet, P., Diprose, J. M., Malby, R., Ziéntara, S., Mertens, P. P. C. & Stuart, D. I. (1998). *Nature (London)*, **395**, 470–478.
Hamilton, W. C., Rollett, J. S. & Sparks, R. A. (1965). *Acta Cryst.* **18**, 129–130.
Hendrickson, W. A. (1991). *Science*, **254**, 51–58.
Kabsch, W. (1988). *J. Appl. Cryst.* **21**, 67–71.
Kim, S. (1989). *J. Appl. Cryst.* **22**, 53–60.
Larson, S. B., Koszelak, S., Day, J., Greenwood, A., Dodds, J. A. & McPherson, A. (1993). *J. Mol. Biol.* **231**, 375–391.
Liddington, R. C., Yan, Y., Moulai, J., Sahli, R., Benjamin, T. L. & Harrison, S. C. (1991). *Nature (London)*, **354**, 278–284.
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.
Luo, M., Vriend, G., Kamer, G. & Rossmann, M. G. (1989). *Acta Cryst.* B**45**, 85–92.
Muckelbauer, J. K., Kremer, M., Minor, I., Tong, L., Zlotnick, A., Johnson, J. E. & Rossmann, M. G. (1995). *Acta Cryst.* D**51**, 871–887.
Olson, N. H., Kolatkar, P. R., Oliveira, M. A., Cheng, R. H., Greve, J. M., McClelland, A., Baker, T. S. & Rossmann, M. G. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 507–511.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Rayment, I., Baker, T. S. & Caspar, D. L. D. (1983). *Acta Cryst.* B**39**, 505–516.
Rossmann, M. G. (1979). *J. Appl. Cryst.* **12**, 225–238.
Rossmann, M. G. (1994). *Prot. Sci.* **3**, 1712–1725.

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.
Rossmann, M. G., Leslie, A. G. W., Abdel-Meguid, S. S. & Tsukihara, T. (1979). *J. Appl. Cryst.* **12**, 570–581.
Smith, J. L. (1997). *Recent Advances in Phasing*, edited by K. S. Wilson, G. Davies, A. W. Ashton & S. Bailey, pp. 25–39. Warrington: Daresbury Laboratory.
Smith, T. J., Chase, E. S., Schmidt, T. J., Olson, N. H. & Baker, T. S. (1996). *Nature (London)*, **383**, 350–354.
Stauffacher, C. V., Usha, R., Harrington, M., Schmidt, T., Hosur, M. V. & Johnson, J. E. (1987). *Crystallography in Molecular Biology*, edited by D. Moras, J. Drenth, B. Strandberg, D. Suck & K. Wilson, pp. 293–308. New York/London: Plenum Press.
Steller, I., Bolotovsky, R. & Rossmann, M. G. (1997). *J. Appl. Cryst.* **30**, 1036–1040.
Thiel, D. J., Blank, B. & Laluppa, J. (1998). *J. Synchrotron Rad.* **5**, 914–916.
Tong, L. & Rossmann, M. G. (1990). *Acta Cryst.* A**46**, 783–792.
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.
Van Etten, J. L., Lane, L. C. & Meints, R. H. (1991). *Microbiol. Rev.* **55**, 586–620.
Vriend, G. & Rossmann, M. G. (1987). *J. Appl. Cryst.* **20**, 338–343.
Winkler, F. K., Schutt, C. E. & Harrison, S. C. (1979). *Acta Cryst.* A**35**, 901–911.