

## Deconvolution of Fully Overlapped Reflections From Crystals of Foot-and-Mouth Disease Virus O<sub>1</sub> G67

BY SUSAN LEA

Laboratory of Molecular Biophysics, Rex Richards Building, University of Oxford, South Parks Road, Oxford OX1 3QU, England

AND DAVE STUART

Laboratory of Molecular Biophysics, Rex Richards Building, University of Oxford, South Parks Road, Oxford OX1 3QU, England, and Oxford Centre for Molecular Sciences, New Chemistry Laboratory, Oxford OX1 3QT, England

(Received 25 April 1994; accepted 25 July 1994)

### Abstract

Foot-and-mouth disease virus O<sub>1</sub> G67 forms crystals that appear similar to those of the closely related viruses O<sub>1</sub>K and O<sub>1</sub>BFS, both of which belong to space group *I*23. Statistical disorder in the O<sub>1</sub> G67 crystals means, however, that the measured diffraction data possess higher symmetry consistent with point group 432. It is shown that this is due to intimate twinning, with mosaic blocks randomly distributed between the two orientations. This results in a twofold loss of information due to the exact superimposition of non-identical reflections from the two orientations. A simple procedure has been devised to deconvolute these overlapped reflections by applying constraints in both real and reciprocal space. This procedure works well, providing interpretable electron-density maps for this virus. Other applications are discussed.

### Abbreviations

FMDV, foot-and-mouth disease virus.

$|F^{23}|$ , *I*23 structure-factor amplitude.

$|F_{\text{mutant}}^{23}|$ , estimate of *I*23 structure-factor amplitude for G67 produced by back-transformation of averaged/flattened map or by re-scaling of the estimates so derived.

$|F^{432}|$ , *I*432 structure-factor amplitude.

$|F_{\text{mutant}}^{\text{unfolded}}|$ , estimate of G67 space group *I*23 structure-factor amplitude produced by duplicating *I*432 amplitudes so that  $|F_{hkl}| = |F_{khl}|$ .

$|F_{\text{parent}}^{\text{unfolded}}|$ , as above but for the parental virus model.

$\alpha_{\text{parent}}$ , phase estimates for the parental virus.

### Introduction

Here we describe the procedures developed to make use of data collected from crystals of a foot-and-

mouth disease virus, G67. Foot-and-mouth disease viruses (FMDV's) are members of the picornavirus family. A number of picornavirus structures have been determined by X-ray crystallography (*e.g.* Hogle, Chow & Filman, 1985; Rossmann *et al.*, 1985; Luo *et al.*, 1987; Grant, Filman, Fujinami, Icenogle & Hogle, 1992; Luo *et al.*, 1992; Acharya *et al.*, 1989). They are revealed to be small (diameter = 300 Å) icosahedral viruses, the capsid consisting of 60 copies of a protomer made up of four proteins (VP1–4) arranged to form a protein coat with *T* = 1 symmetry (Caspar & Klug, 1962). The structures of two viruses closely related to G67 are known, namely virus O<sub>1</sub>BFS (Acharya *et al.*, 1989), which differs from G67 by 11 mutations in the four capsid proteins, and virus O<sub>1</sub>K (manuscript in preparation) which differs at only six positions. G67 is, in fact, derived from virus O<sub>1</sub>K by growing O<sub>1</sub>K *in vitro* in the presence of the selective pressures imposed by antibodies recognising the virus (Kitson, McCahon & Belsham, 1990; McCahon *et al.*, 1989).

### Nature of the G67 crystals

Crystals of G67 were grown using conditions similar to those used to grow crystals of O<sub>1</sub>K and O<sub>1</sub>BFS (Fox *et al.*, 1987; Curry *et al.*, 1992) *i.e.* crystals grew as rhombic dodecahedra from a solution of the virus at a concentration of approximately 10 mg ml<sup>-1</sup> (pH 7.6, 100 mM sodium phosphate buffer) by microdialysis against 21–23% saturated (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>. Visual examination of diffraction images (Fig. 1) indicated that the virus had crystallized in space group *I*23, with cell dimensions  $a = b = c = 345.0$  Å,  $\alpha = \beta = \gamma = 90^\circ$ , as had both O<sub>1</sub>BFS and O<sub>1</sub>K.

Analysis of diffraction data in *I*23 requires an arbitrary but consistent assignment of the orientation of the crystallographic axes. The ambiguity in the orientation of the axes corresponds to the two

different ways in which an icosahedron can be placed on a specific threefold axis. The two settings are related by a  $90^\circ$  rotation about a particle twofold (Fig. 2). Within any crystal the particles are all in the same relative orientation, but without consulting the distribution of intensities the choice will be random between crystals. We routinely divide data collected from such crystals into two streams by comparison with a reference data set: intensity data that have been indexed differently to the reference data set may then be brought to the standard setting by simply swapping a pair of reciprocal-space indices (e.g.  $h$  and  $k$ ).

We determine the orientation of data from any single crystal (the test data) by calculating a linear

correlation coefficient\* ( $C$ ) in reciprocal space between those data and the reference data set (which defines the standard indexing). With data collected from two closely related viruses  $C$  is normally of the order of 0.8 in the correct setting and less than 0.2 with the alternative indexing. However, when attempts were made to divide the G67 data into two streams the linear correlation coefficient in reciprocal space between the G67 data and the reference set (a set of data for the parent virus) was very low – for most crystals it was less than 0.5 in either orientation. The inability of the correlation coefficient to distinguish between the two indexing strategies suggested that extra, fourfold, symmetry might be present in these data. Fourfold symmetry is geometrically impossible since an icosahedral virus has no inherent fourfold symmetry and the unit cell is too small to allow association of viral particles around a fourfold symmetry axis, however, statistically the data did indeed appear to belong to point group 432. The data were processed in space group  $I432$ , yielding excellent merging  $R$  factors [ $R(I)^\dagger = 11.3\%$ ,

$$* C = \frac{[\sum_h ((I_h) - I_h) (I_{h,ref}) - I_{h,ref}) / \{[\sum_h ((I_h) - I_h)^2 \sum_h ((I_{h,ref}) - I_{h,ref})^2\}^{1/2}]}{}$$

$$^\dagger R(I) = \frac{[\sum_h \sum_i |I_h - I_{hi}|]}{(\sum_h \sum_i I_{hi})} \times 100.$$

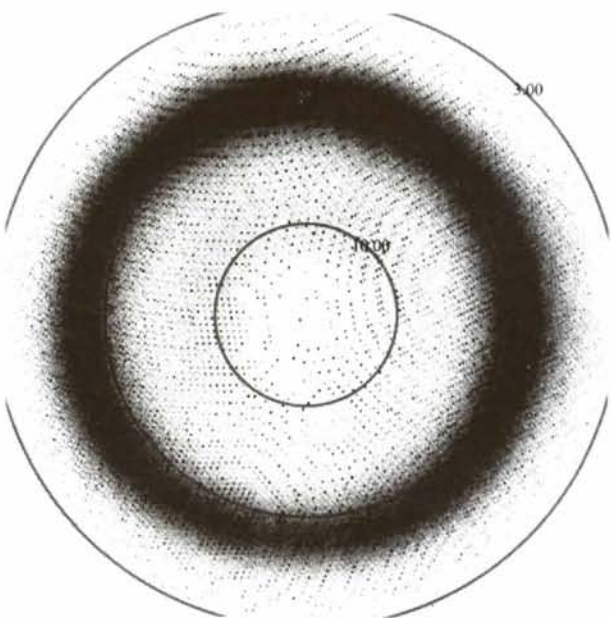


Fig. 1. G67 film image. Oscillation  $0.00\text{--}0.50^\circ$ , wavelength  $0.890 \text{ \AA}$ , distance  $190.0 \text{ mm}$ .

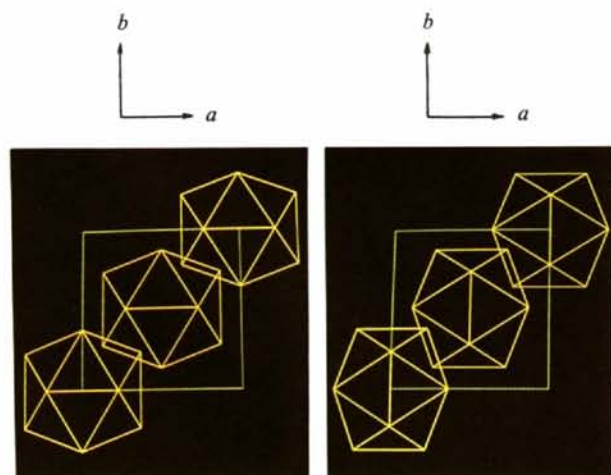


Fig. 2. Packing ambiguity in  $I23$ .

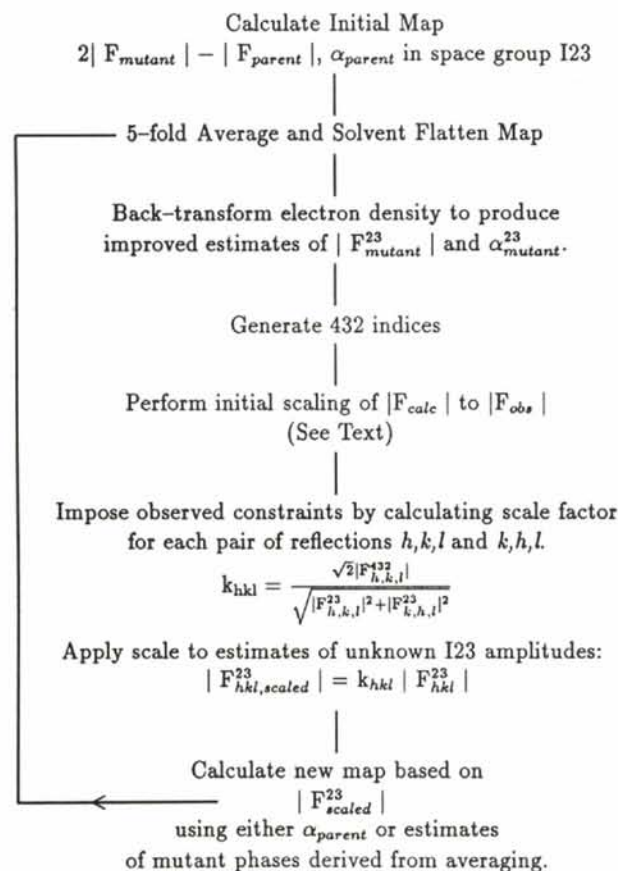


Fig. 3. Summary of the deconvolution procedure.

indistinguishable from those obtained in space group *I*23 and equivalent to the merging *R* factors for data collected from crystals of the related foot-and-mouth disease viruses processed in *I*23 (Table 1)]. Thus, no information was lost by processing the G67 data in this higher symmetry point group (we would emphasize that this was true for all the G67 crystals examined).

This 'statistical symmetry' may be explained by assuming a disordering of the crystal packing, such that both orientations are present within the same crystal. Since the two lattices are exactly superimposable it is impossible to identify this 'twinning' phenomenon simply by examination of the geometry of the diffraction pattern, furthermore, if the ratio of the two orientations is 1:1, the data will have apparent fourfold symmetry. In the current example the data obey 432 symmetry to within the error levels

Table 1. *Data quality*

	Reduced* G67	Oxidized* O,K
$\lambda$ (Å)	0.89	1.488
No. of crystals	26	35
No. of useful film packs	27	27
Resolution limits (Å)	3.0	2.6
<i>R</i> ( <i>I</i> )† (%)	11.3	11.3
Total No. observed reflections	165285	237674
No. of independent reflections	64004	138727
% Complete	92‡	67¶

\* See Logan *et al.* (1993).

†  $R(I) = [\sum_h \sum_l (I_h - I_{hl})] / (\sum_h \sum_l I_{hl}) \times 100$ .

‡ In point group 432.

¶ In point group 23.

expected from counting statistics, implying that the ratio of the two orientations is extremely close to 1:1. Such increases in the apparent symmetry of a system due to a twinning phenomenon are not uncommon (see for instance, Bragg & Bragg, 1933) and may be

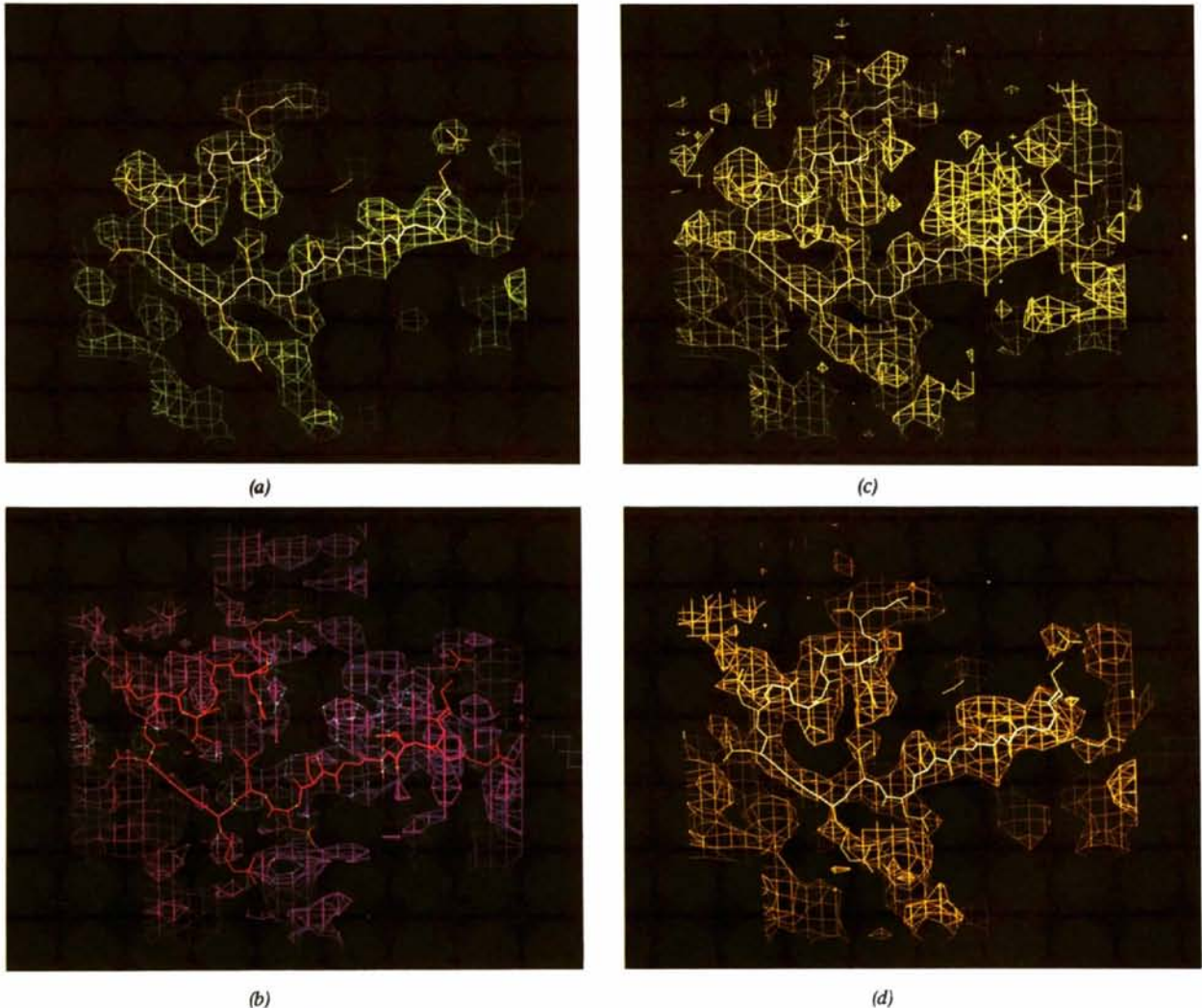


Fig. 4. Density produced at different stages of deconvolution of test data. (a) 'True' density,  $2|F_{\text{mutant}}^{23}| - |F_{\text{parent}}^{23}|$ ,  $\alpha_{\text{parent}}$ . (b) Unaveraged density,  $2|F_{\text{mutant}}^{\text{unfolded}}| - |F_{\text{parent}}^{\text{unfolded}}|$ ,  $\alpha_{\text{parent}}$ . (c) Cycle 1. (d) Cycle 7.

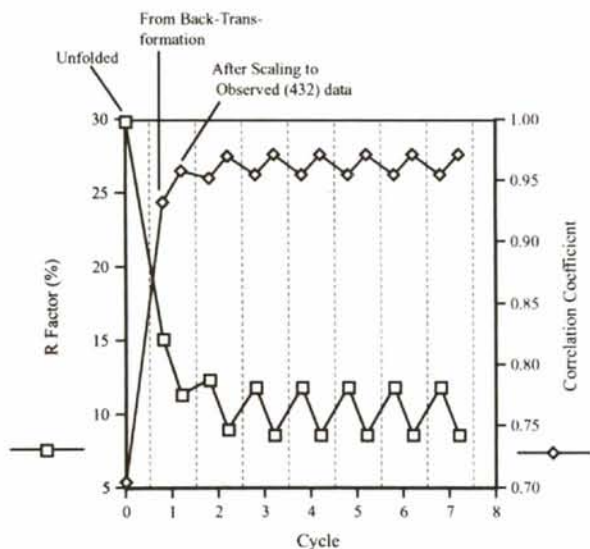


Fig. 5. Agreement between 'true' and estimated  $|F_{mutant}^{23}|$  during the test of the deconvolution protocol.

especially likely to occur in crystals of viruses due to their approximately spherical nature. This situation is a special case of merohedral twinning (Buerger, 1960) where the proportion of the two lattice orientations is, in general, unequal. Methods have been described to determine the relative proportions of the two lattices in the general case (see for instance, Rees, 1980; Fisher & Sweet, 1980; Britton, 1972). Given this information the methods described below are readily extended to any example of merohedral twinning.

**Models for the disorder**

To understand the disordering further we used data in point group  $I23$  from the closely related parental virus (O<sub>1</sub>K) to model the G67 data.

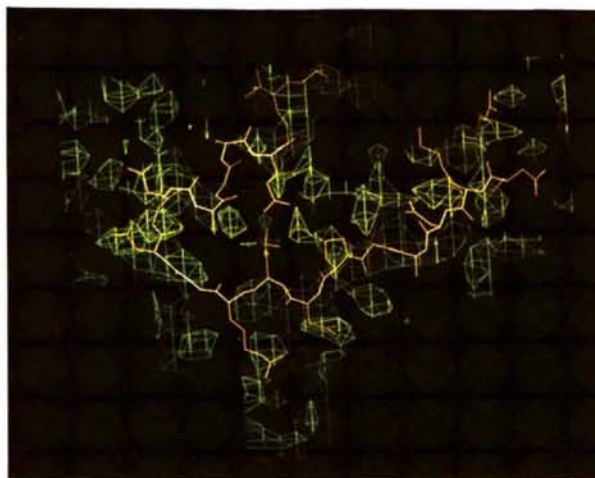
In point group 23 reflection  $h,k,l$  is independent of reflection  $k,h,l$ , i.e.  $F_{h,k,l} \neq F_{k,h,l}$  whereas, in point group 432 these reflections are equivalent:  $F_{h,k,l} = F_{k,h,l}$ . Two extreme models for the disordering may be used to reduce the point group 23 data to point group 432.

(1) If all virion twofolds are randomly arranged throughout the crystals then interference effects must be considered and a vector mean of the phased  $I23$  structure factors will model the  $I432$  data,

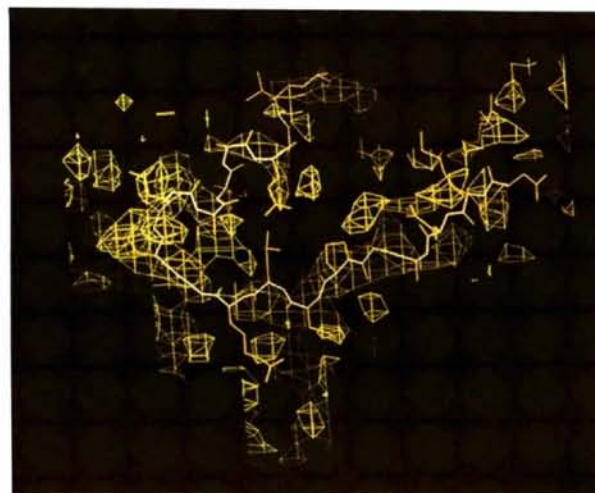
$$F_{h,k,l}^{432} = (F_{h,k,l}^{23} + F_{k,h,l}^{23})/2. \tag{1}$$

(2) If mosaic blocks of the crystal contain similarly orientated particles, but the different mosaic blocks are randomly orientated there will be no interference effects and the arithmetic mean of the intensities will model the  $I432$  data,

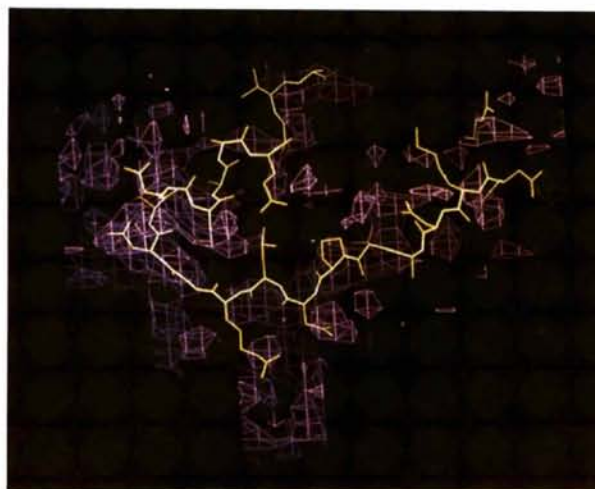
$$I_{h,k,l}^{432} = (I_{h,k,l}^{23} + I_{k,h,l}^{23})/2. \tag{2}$$



(a)



(b)



(c)

Fig. 6. Density produced at different stages of deconvolution of the G67-DIT data. (a) Unaveraged density,  $2|F_{G67-DIT}^{unfolding} - F_{parent}^{unfolding}|$ . (b) Cycle 1. (c) Cycle 3.

Table 2. *Scaling of model data based on O<sub>1</sub>K I23 data to 432 G67 data*

	R factor (%)	Correlation coefficient
Vector mean of structure factors	30	0.70
Arithmetic mean of amplitudes	14	0.89

Reduced data sets based on the O<sub>1</sub>K I23 data were constructed using both models. These model data sets were scaled in resolution shells to the observed G67 data and *R* factors and correlation coefficients calculated to judge the appropriateness of the model. The results are shown in Table 2. The arithmetic mean of the intensities provides an excellent model for the G67 data (to within the accuracy expected on the basis of the quality of the data) implying that disorder occurs between mosaic blocks within the crystals.

Thus, the G67 diffraction data are formed by the superimposition of two sets of non-equivalent lattice points leading to a loss of information. In such situations the usual procedure is to seek a more favourable crystal form. However, in this case the data inherently oversamples the molecular transform. This oversampling arises from two sources.

(1) The disordered solvent and RNA.

(2) The fivefold non-crystallographic symmetry (NCS) present in this system.

We have used two protocols which combine these sources of information with the observed X-ray intensity data and prior knowledge of the phases to produce interpretable electron-density maps.

(i) *The use of non-crystallographic symmetry to obtain interpretable difference maps*

A simple protocol has been developed to calculate difference maps between G67 and other related viruses which had crystallized in space group *I*23. Firstly the *I*23 data are reduced to point group 432 using equation (2). These data are scaled to the observed G67 data in resolution shells. The scaled data may then be 'unfolded' back to *I*23 by duplicating reflection records; *h,k,l* being written as *h,k,l* and *k,h,l* with the same structure-factor amplitude. By associating known, *I*23, phases with these structure-factor amplitudes difference Fourier maps may be calculated in *I*23. These maps are noisy, containing amplitude data for two orientations, but can be improved by one cycle of fivefold averaging using the NCS relationships appropriate to the *I*23 phases: the standard orientation is reinforced whilst information corresponding to the other orientation is destroyed. This single averaging step produced sufficient improvement in the maps to reveal the structural differences between the viruses and suggested the cause of the disordering (manuscript in preparation). The results of this procedure will not be discussed

further since the primary purpose of this paper is to present a method that does not depend on the use of difference Fourier methods.

(ii) *A cyclic procedure to deconvolute the overlapped data*

In this protocol we go beyond simple averaging of a difference Fourier to attempt to reconstruct the underlying structure-factor amplitudes for the true *I*23 cell. The procedure is iterative, for each cycle we calculate an electron-density map with the current best estimates for the *I*23 structure-factor amplitudes and phases and then use solvent flattening and NCS averaging in real space, followed by back-transformation in space group *I*23 to produce new estimates of the structure factors. We now use our observed data to provide a series of constraint equations in reciprocal space, one for each pair of reflections, *h,k,l* and *k,h,l*,

$$|F_{h,k,l}^{\text{OBS}-432}| = [(|F_{h,k,l}^{23}|^2 + |F_{k,h,l}^{23}|^2)/2]^{1/2}.$$

These constraints further improve the estimates of the unknown intensities in *I*23 by providing rescale factors for the pairs of reflections *h,k,l* and *k,h,l*. If the constraint information is sufficiently strong cycles of map calculation, real-space constraint imposition, structure-factor calculation and reciprocal-space constraint imposition should converge to provide reasonable estimates of the true electron density for the G67 virion.

**A test of procedure (ii)**

To enable us to test the deconvolution protocol synthetic data were generated from coordinates of DTT-soaked O<sub>1</sub>BFS (Logan *et al.*, 1993). Structure-factor amplitudes were generated in the range 30–3.5 Å for two slightly different models.

(1) To model the space group *I*23 data for the parental virus, structure factors were generated from the coordinate set with 30 residues deleted (130–160 of VP1, corresponding to a surface-exposed loop).

(2) To model the disordered G67 data the complete coordinate set (including the 130–160 loop) was used to generate structure-factor amplitudes in space group *I*23, which were then reduced to point group 432. The *I*23 amplitudes were then 'thrown away' for all further calculations, except where they provided a measure for the success of the deconvolution process.

The success of the deconvolution process was monitored in two ways. Firstly a subjective assessment based on the quality of the density for the truncated loop region – reconstructing this density is taken as the primary aim of the deconvolution. Secondly, for this test of the method, the agreement between the true *I*23 amplitudes and the estimates of

Table 3. Agreement between  $|F_{mutant}^{432}|$  and estimates of  $|F_{mutant}^{23}|$  at different stages of the deconvolution process

Cycle	$R_c$	$C$
'Unfolded'	0.0	1.00
1		
After averaging	26.0	0.73
After scaling to 432 data	24.0	0.77
2		
After averaging	28.7	0.70
After scaling to 432 data	27.3	0.73
3		
After averaging	29.7	0.70
After scaling to 432 data	28.4	0.72

Convergence achieved

these amplitudes obtained from the deconvolution process provides an objective and quantitative measure of success.

The process is summarized in Fig. 3. The programs used to perform the envelope determination, NCS symmetry averaging and solvent flattening are based on those of Bricogne (Bricogne, 1976), and further developed in our laboratory (Fry, Acharya & Stuart, 1993). We describe here only those procedures unique to the deconvolution process. The protein envelopes were generated using the parental  $I23$  data (modified to ensure that the loop to be reconstructed was not treated as solvent) – the parent and mutant structures are so similar that these should prove reasonably appropriate, and of better quality than envelopes calculated using the noisy 'unfolded' data. The areas of RNA and solvent defined by these envelopes were then independently 'flattened' and the protein areas averaged over the five NCS-related copies.

Back-transformation of the averaged and flattened electron-density map generates new estimates of the structure-factor amplitudes and phases, in  $I23$ , for the mutant virus. Initially, any phase information obtained from this back-transformation was discarded. This seems appropriate since we have an accurate set of phases for the parent virus obtained by 719 cycles of NCS symmetry averaging (Fry *et al.*, 1993) – the structural differences between the parent and the mutant are small, therefore the parental phases are likely to be closer to the true  $I23$  mutant phases than phase estimates produced by the deconvolution process.

The new estimates for the  $I23$  amplitudes produced by back-transformation are not on the same absolute scale as the observed data, therefore, before applying the scaling constraints arising from the reflection pairs, an overall scaling is performed. Scales are calculated in resolution shells based on the 'singlet',  $h = k$ , reflections for which there is no loss of information in the disordered crystals. These singlet scales are then applied to all the reflections

within each resolution shell, and the constraints derived from the observed data used to calculate and apply independent rescale factors for each reflection pair. These scaled amplitudes ( $|F_{scaled}|$ ) are then used with the parental phases ( $\alpha_{parent}$ ) to calculate a new map, and a further cycle of averaging, solvent flattening and deconvolution can begin.

The results of the test are shown in Fig. 4 (real space) and Fig. 5 (reciprocal space). Fig. 4(a) shows the target for the calculation where the map was calculated with true  $I23$  coefficients. Fig. 4(b) shows the start point where the density was calculated using amplitudes unfolded to space group  $I23$  (these 'unfolded' amplitudes scale to the 'true' amplitudes with  $R_c = 29.9\%$ ,\*  $C = 0.70$ , Fig. 5). Fig. 4(c) shows the map after one cycle of averaging and flattening (the estimates of the  $I23$  amplitudes now scale to the 'true' amplitudes with  $R_c = 15.1\%$ ,  $C = 0.93$ ). The first cycle of fivefold averaging and flattening is seen to have the greatest effect on the quality of the map. Application of the reciprocal-space constraints to these estimates of the  $I23$  amplitudes improved their agreement to the observed (432) data – from  $R_c = 25.0$  to  $24.0\%$  ( $C = 0.73$  to  $0.77$ ) as expected and, more importantly, increased the agreement with the 'true' amplitudes – from  $R_c = 15.1$  to  $11.3\%$  ( $C = 0.95$  to  $0.96$ ). The next two cycles of deconvolution further improve the agreement between the estimates and the 'true' amplitudes to  $R_c = 8.6\%$ ,  $C = 0.97$  (Fig. 5). The process has converged by the third cycle and no further improvement is seen, although seven cycles were performed (Fig. 4d shows the map after seven cycles). In achieving convergence to the true values the  $I23$  estimates diverge from the starting values derived from the observed data (Table 3).

We next investigated whether the protocol was sufficiently powerful to refine the phases as well as the amplitudes. The phase estimates produced by back-transformation of the averaged and flattened map at cycle 3 (the point at which convergence was achieved) were taken as the starting point for three further cycles of deconvolution where phase estimates were updated on each cycle to those produced by back-transformation of the averaged flattened map. This process also converged after three cycles (at which point the mean phase change between cycles was less than  $2^\circ$ ). Although the phases moved away from the parental phases (mean phase difference  $15.8^\circ$  at cycle 6) the phases also diverged from the 'true' mutant phases (mean phase difference  $14.6^\circ$  at cycle 3,  $16.0^\circ$  at cycle 6) and the agreement between the estimates of the mutant  $I23$  amplitudes and the true amplitudes was slightly worse than it had been before release of the phases with  $R_c = 12.0\%$  compared to  $11.9\%$  ( $C$  did not change).

\*  $R_c = [\sum_h (|F_{obs,h} - |F_{calc,h}|)] / [\sum_h |F_{obs,h}|] \times 100$ .

This test demonstrates that the deconvolution process can successfully reconstruct amplitudes, but that there is insufficient power (in this case) to reconstruct an accurate set of phases. It is likely, however, that in the presence of more real-space constraints (such as a more stringent definition of the protein envelope) phase refinement would be possible.

#### Deconvolution of the G67 Data

The deconvolution process was then applied to the experimental G67 data. Since we do not know the true values of the underlying G67 *I*<sub>23</sub> amplitudes the interpretability of the electron-density maps probably provides the best measure of success. As for the test case the RNA and solvent masks were calculated using the parental data. To ensure that genuine surface changes were not truncated the volume assigned to protein was overestimated. The maps produced are shown in Fig. 6; (a) shows the map calculated using 'unfolded' amplitudes, before averaging; (b) shows the map after one cycle of averaging; (c) shows the map after three cycles of deconvolution – at this point the process has converged and cycling was stopped. As for the test data the map quality is significantly improved by one averaging cycle, allowing unambiguous interpretation of the previously uninterpretable density for the loop. There is then a further, slight, improvement on cyclic deconvolution, thus, for instance the connectivity of the density around the 3<sub>10</sub> helix region of the loop (residues 148–155 of VP1) is improved. The electron density produced by this procedure was in accord with the results of the difference Fourier procedure. Since comparisons with 'true' G67 amplitudes cannot be made *R*<sub>c</sub> and *C* were calculated between the G67 estimates and the observed O<sub>1</sub>BFS amplitudes. After three cycles *R*<sub>c</sub> = 22.0% *C* = 0.85 whereas, the figures for the agreement between parent and 'mutant' amplitudes in the test case was *R*<sub>c</sub> = 12.9% *C* = 0.96. This reflects the larger structural differences between G67 and O<sub>1</sub>BFS than those modelled in the synthetic data and also the fact that real data are not error free.

In summary, the deconvolution protocol produces improved maps. The bulk of the improvement occurs during the first round of averaging and flattening and the process converges rapidly. As for the test data, optimization of the protein envelope would be expected to produce a further improvement.

#### Discussion

The method presented here is simple and works well, but is only one of a variety of approaches that might

have been taken to use these data. Other approaches could include use of difference Fourier coefficients combined with envelopes delineating the areas (in real space) within which structural alterations are presumed to occur. This would allow more stringent solvent-flattening constraints to be imposed, however this is a less general method and requires a prior estimate of the extent of the structural changes.

The approaches of Bricogne (Bricogne, 1991, 1993) using maximization of entropy and likelihood would allow a more rigorous treatment of such effects, indeed they are already proving effective in the case of powder diffraction data where there is even greater information loss. However the method we have developed is simple, quick and fits in easily with the phase-refinement procedures currently in use in many laboratories.

The power of this technique would be significantly greater in systems with higher NCS. The fivefold averaging possible in the true *I*<sub>23</sub> cell, is in effect, reduced to 2½-fold by the twofold ambiguity introduced by the disordering. Many virus systems crystallize with much higher NCS, e.g. Mengo (Luo, Vriend, Kamer & Rossmann, 1989) and BEV (Smyth *et al.*, 1993) with 60-fold NCS. In such cases, with greater phasing power to be gained from the averaging, we would expect that such a deconvolution process could successfully deconvolute entirely overlapped reflections even in the absence of an accurate set of starting phases.

The method, as described, is not specific to the deconvolution of completely overlapped reflections produced by crystal mispacking; it provides a simple and rather general method that may be directly applied in other cases, for instance to the deconvolution of the harmonic overlaps that occur using the wide wavelength range Laue method.

We thank the members of the molecular biology picornavirus group of the Institute for Animal Health, Pirbright Laboratory, for growth of the virus crystals; S. Curry, E. Fry, D. Goodridge, A. King, D. Logan and J. Newman for help in data collection; the staff of the Synchrotron Radiation Source, SERC Daresbury Laboratory, for practical assistance; R. Bryan, R. Esnouf and others from the LMB for their support; E. Y. Jones for critical reading of the manuscript; SL was supported by the AFRC and MRC.

#### References

- ACHARYA, R., FRY, E., STUART, D., FOX, G., ROWLANDS, D. & BROWN, F. (1989). *Nature (London)*, **337**, 709–716.
- BRAGG, W. & BRAGG, W. (1933). *The Crystalline State*. London: G. Bell and Sons.
- BRICOGNE, G. (1976). *Acta Cryst.* **A32**, 832–847.

- BRICOGNE, G. (1991). *Acta Cryst.* **A47**, 803–829.
- BRICOGNE, G. (1993). *Acta Cryst.* **D49**, 37–60.
- BRITTON, D. (1972). *Acta Cryst.* **A28**, 296–297.
- BUERGER, M. (1960). *Crystal Structure Analysis*. New York: John Wiley.
- CASPAR, D. & KLUG, A. (1962). *Cold Spring Harbour Symp. Quant. Biol.* **27**, 12219–12223.
- CURRY, S., ABU-GHAZALEH, R., BLAKEMORE, W., FRY, E., JACKSON, T., KING, A., LEA, S., LOGAN, D., NEWMAN, J. & STUART, D. (1992). *J. Mol. Biol.* **228**, 1263–1268.
- FISHER, R. & SWEET, R. (1980). *Acta Cryst.* **A36**, 755–760.
- FOX, G., STUART, D., ACHARYA, K., FRY, E., ROWLANDS, D. & BROWN, F. (1987). *J. Mol. Biol.* **196**, 591–597.
- FRY, E., ACHARYA, R. & STUART, D. (1993). *Acta Cryst.* **A49**, 45–55.
- GRANT, R., FILMAN, D., FUJINAMI, R., ICENOGLU, J. & HOGLE, J. (1992). *Proc. Natl Acad. Sci. USA*, **89**, 2061–2065.
- HOGLE, J., CHOW, M. & FILMAN, D. (1985). *Science*, **229**, 1358–1365.
- KITSON, J., MCCAHOON, D. & BELSHAM, G. (1990). *Virology*, **179**, 26–34.
- LOGAN, D., ABU-GHAZALEH, R., BLAKEMORE, W., CURRY, S., JACKSON, T., KING, A., LEA, S., LEWIS, R., NEWMAN, J., PARRY, N., ROWLANDS, D., STUART, D. & FRY, E. (1993). *Nature (London)*, **362**, 566–568.
- LUO, M., HE, C., TOTH, K., ZHANG, C. & LIPTON, H. (1992). *Proc. Natl Acad. Sci. USA*, **89**, 2409–2413.
- LUO, M., VRIEND, G., KAMER, G., MINOR, I., ARNOLD, E., ROSSMANN, M., BOEGÉ, U., SCRABA, D., DUKE, G. & PALMENBERG, A. (1987). *Science*, **235**, 182–191.
- LUO, M., VRIEND, G., KAMER, G. & ROSSMANN, M. (1989). *Acta Cryst.* **B45**, 85–92.
- MCCAHOON, D., CROWTHER, J., BELSHAM, G., KITSON, J., DUCHESNE, M., HAVE, P., MELOEN, R., MORGAN, D. & DE SIMONE, F. (1989). *J. Gen. Virol.* **70**, 639–645.
- REES, D. (1980). *Acta Cryst.* **A36**, 578–581.
- ROSSMANN, M., ARNOLD, E., ERICKSON, J., FRANKENBURGER, E., GRIFFITH, J., HECHT, H., JOHNSON, J., KAMER, G., LUO, M., MOSSER, A., RUECKERT, R., SHERRY, B. & VRIEND, G. (1985). *Nature (London)*, **317**, 145–153.
- SMYTH, M., FRY, E., STUART, D., LYONS, C., HOEY, E. & MARTIN, S. (1993). *J. Mol. Biol.* **231**, 930–932.