Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Crystallization and preliminary X-ray diffraction analysis of the sialic acid-binding domain (VP8*) of porcine rotavirus strain CRW-8

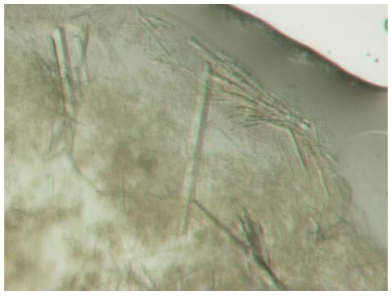
Stacy A. Scott,[a]‡ Gavan Holloway,[b] Barbara S. Coulson,[b] Alex J. Szyczew,[a] Milton J. Kiefel,[a] Mark von Itzstein[a] and Helen Blanchard[a]*‡

[a]Institute for Glycomics, Griffith University (Gold Coast Campus) PMB 50, Gold Coast Mail Centre, Queensland 9726, Australia, and [b]Department of Microbiology and Immunology, The University of Melbourne, Victoria 3010, Australia

‡ These authors made equal contributions.

Correspondence e-mail: h.blanchard@griffith.edu.au

Received 24 March 2005
Accepted 3 May 2005
Online 1 June 2005



Rotavirus recognition and attachment to host cells involves interaction with the spike protein VP4 that projects outwards from the surface of the virus particle. An integral component of these spikes is the VP8* domain, which is implicated in the direct recognition and binding of sialic acid-containing cell-surface carbohydrates and facilitates subsequent invasion by the virus. The expression, purification, crystallization and preliminary X-ray diffraction analysis of VP8* from porcine CRW-8 rotavirus is reported. Diffraction data have been collected to 2.3 Å resolution, enabling the determination of the VP8* structure by molecular replacement.

1. Introduction

Rotaviruses cause severe gastroenteritis in the young of most mammalian species, including humans, *via* infection of epithelial cells of the small intestine (Estes, 2001), resulting in an estimated 440 000 deaths of human infants per year (Parashar *et al.*, 2003). Rotaviruses are classified into serogroups A–E, with group A causing most human disease, and are further divided into G and P serotypes based on identification of outer capsid antigens (Coulson, 1993; Lundgren & Svensson, 2001). The rotavirus virion comprises a triple-layered protein structure surrounding a genome of 11 segments of double-stranded RNA (Estes, 2001). The outermost layer is composed of two structural proteins, the viral haemagglutinin VP4, which forms a spike (possibly as a dimer) that extends 100 Å from the virion surface, and the coat glycoprotein VP7, which is assembled into calcium-dependent trimers (Prasad & Chiu, 1994). Calcium chelation causes conformational changes in VP7 and enables conversion into a transcriptionally active double-layered particle. This uncoating occurs after the intact particle crosses the membrane and enters into the cell cytoplasm. Since the triple-layered particle binds gangliosides, whereas the double-layered particle cannot (Delorme *et al.*, 2001), it is clear that proteins forming the outer layer are involved in oligosaccharide binding. VP4 and VP7 are both implicated in initial host-cell interactions and are crucial to rotavirus infectivity (Coulson *et al.*, 1997; Estes & Cohen, 1989; Estes, 2001), with VP4 having specific roles in viral attachment, entry and neutralization (Estes, 2001; Prasad & Chiu, 1994).

Rotavirus infectivity requires proteolytic cleavage of VP4, giving an N-terminal sialic acid-recognizing domain (VP8*) and a C-terminal fragment (VP5*) that remain associated with the virion (Dormitzer *et al.*, 2001). It has been postulated that rotavirus initially interacts with a sialic acid-containing cellular recognition site prior to interacting with additional determinants, including integrins (Jolly *et al.*, 2000; Hewish *et al.*, 2000). The binding ability of the triple-layered particles of various rotavirus strains to host cells was investigated by analysing infectivity of cultured cells after treatment with sialidases that remove terminal (but not internal) sialic acids (such as *N*-acetylneuraminic acid, Neu5Ac) from the cell-surface glycoconjugates (Ciarlet & Estes, 1999; Fukudome *et al.*, 1989). This indicates whether infection is dependent upon recognition and binding of a terminal sialic acid residue on the glycoconjugate. Virus strains that are 'sialidase-insensitive' therefore could still be recog-

nizing a sialic acid, but only those located within the oligosaccharide chain. The absolute requirement for sialic acids to enable infection is controversial, however there is growing consensus that those strains of rotavirus that are insensitive to sialidase treatment, for example the human rotaviral strains, appear able to recognise sialic acids located internally within oligosaccharide chains (Delorme *et al.*, 2001; Guerrero *et al.*, 2000; Guo *et al.*, 1999; Jolly *et al.*, 2000). This is further exemplified by human rotavirus KUN and MO strains utilizing recognition of the sialic acid containing ganglioside GM1 on the host cell surface during infection (Guo *et al.*, 1999).

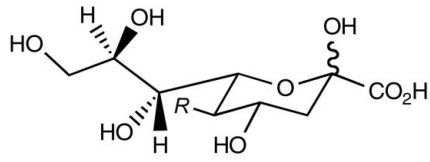
Rotavirus infections show remarkable host specificity and cross-species infections are rare events, although the species barrier is not absolute; for example, a bovine-like rotavirus has been isolated from symptomatic infants (Gerna *et al.*, 1992) and additional evidence suggests cross-species pathogenicity, with bovine rotavirus being a potential risk to pigs by the emergence of a pathogenic reassortant rotavirus (El-Attar *et al.*, 2001). During the binding of rotavirus to cells there are multiple cell-surface interactions, some being common to human and lower animal strains (Guo *et al.*, 1999). Importantly, there are substantiated differences in the structural components of the glycoconjugates recognized according to the rotavirus strain. Common sialic acids of mammalian cells are *N*-acetylneuraminic acid (Neu5Ac, R = $NHCOCH_3$; Fig. 1) and the hydroxylated form, *N*-glycolylneuraminic acid (Neu5Gc, R = $NHCOCH_2OH$; Fig. 1). Some rotavirus strains, including porcine OSU (which has the same serotype as porcine CRW-8), preferentially bind gangliosides containing Neu5Gc, although some biochemical data has shown that GM_3 gangliosides consisting of either Neu5Ac or Neu5Gc isolated from piglet intestine are relevant cell-surface receptors for OSU (Rolsma *et al.*, 1998). Rhesus rotavirus (RRV) VP8* exhibits a tenfold lower affinity for Neu5Gc than for Neu5Ac (K_d of 1.2 m*M*), demonstrating contrast to porcine rotavirus strains (Dormitizer, Sun, Blixt *et al.*, 2002). The mechanism by which different rotavirus strains exhibit host-specificity is undetermined. To investigate this phenomenon, and ultimately progress our research toward the design of rotavirus-specific therapeutics (Fazli *et al.*, 2001; Kiefel & von Itzstein, 2003) we have initiated structural studies on the VP8* domain from various rotavirus strains. Here we report the cloning, overexpression and protein purification of porcine CRW-8 $VP8^*_{64–224}$, as well as the preliminary X-ray crystallographic structure determination by molecular replacement.

2. Experimental procedures and results

2.1. Cloning, expression and purification of CRW-8 $VP8^*_{64–224}$

2.1.1. Cloning. Viral dsRNA was extracted from rotavirus strain CRW-8 (passage 30) as previously described (Dyall-Smith & Holmes, 1984). This dsRNA was used as a template for cDNA synthesis using the Superscript First-Strand Synthesis System (Invitrogen) for reverse transcription-polymerase chain reaction (RT-PCR). The primer used for RT was 5′-AAGTCCGTCGGTAATGAACC-3′ and products of the RT reaction were then used as templates for PCR using the Expand High Fidelity PCR system (Roche) to amplify the VP8* gene fragment for cloning into the bacterial expression vector pGEX-4T-1 (Amersham Biosciences). A single round of PCR was performed using the forward primer 5′-CGC**GGATCC**CTGC-CTGCTAGATGGTCCGTATC-3′ and the reverse primer 5′-G**GAATTC**TCACAGTCCATGATTGATATACTC-3′ containing *Bam*HI and *Eco*RI restriction sites (bold). Following digestion, PCR products were ligated into pGEX-4T-1, yielding pGEX-CRW8-VP8*, which encoded amino-acid residues 64–224 of CRW-8 VP4 fused to the C-terminus of GST. The integrity of the inserted VP8* gene fragment was assessed by DNA sequencing. The predicted amino-acid sequence of pGEX-CRW8-VP8* was identical to the published CRW-8 VP4 sequence except for the substitutions of glutamine for arginine (position 70), isoleucine for valine (position 82), asparagine for serine (position 151), serine for proline (position 157) and serine for leucine (position 201) (accession No. L07888, Entrez Database).

2.1.2. Expression and purification. *Escherichia coli* strain BL21 DE3 transformed with pGEX-CRW8-VP8* plasmid was grown at 310 K in Luria–Bertani (LB) medium supplemented with 150 µg ml^{-1} ampicillin. Once an OD_{600} of 0.6 had been reached, the cultures were incubated at 298 K for 1 h and then induced with 1 m*M* isopropyl α-D-thiogalactopyranoside (BioVectra DCL). Cells were harvested 4 h after induction by centrifugation at 8000*g* for 10 min. Frozen cell pellets were thawed in PBS (137 m*M* NaCl, 2.7 m*M* KCl, 10 m*M* Na_2HPO_4, 2 m*M* KH_2PO_4 pH 7.3) supplemented with 1 m*M* phenylmethylsulfonyl fluoride (PMSF, Roche Diagnostics). The cells were lysed with 1 mg ml^{-1} lysozyme supplemented with 1%(*w*/*v*) Triton X-100 and 20 µg ml^{-1} DNaseI. Cell debris was removed by centrifugation at 20 000*g* for 30 min and the supernatant was passed over a glutathione-Sepharose column (Amersham-Pharmacia Biotech). The column was washed with TNC (20 m*M* Tris pH 8.0, 100 m*M* NaCl, 1 m*M* $CaCl_2$) and bound GST-fusion protein was digested with 8 µg ml^{-1} TPCK-treated trypsin (Worthington Biochemical) for 2 h at room temperature. A benzamidine-Sepharose (Amersham-Pharmacia Biotech) column pre-equilibrated with TNC was connected in series with the glutathione-Sepharose column. The cleaved protein was eluted with 20 m*M* $NaPO_4$ pH 7.5, 1 *M* NaCl and 1 m*M* PMSF; 2.5 m*M* benzamidine was added to the eluant. The



Neu5Ac, R = $NHCOCH_3$
Neu5Gc, R = $NHCOCH_2OH$

Figure 1
N-Acetylneuraminic acid (Neu5Ac; R = $NHCOCH_3$) and the hydroxylated form *N*-glycolylneuraminic acid (Neu5Gc), where R = $NHCOCH_2OH$.

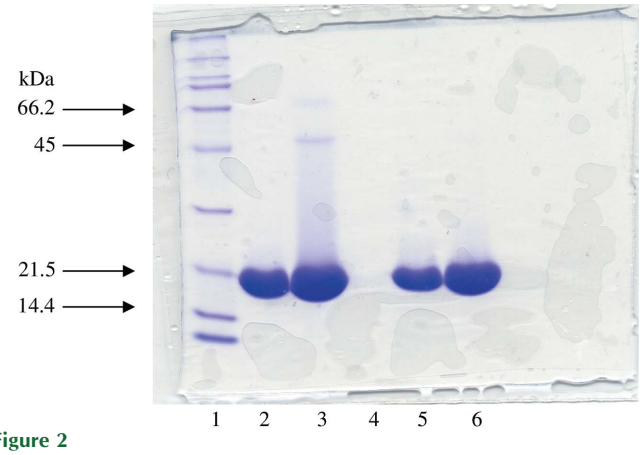


Figure 2
SDS–PAGE (12%, stained with Coomassie blue) analysis of porcine CRW-8 $VP8^*_{64–224}$ (calculated molecular weight of 18 150 Da). Lane 1, ladder. Lanes 2 and 3 are the protein sample (65 µg and 195 µg, respectively, protein concentrations were 13 mg ml^{-1} and 40 mg ml^{-1}) prior to gel-filtration. These lanes both show a higher molecular weight species (in lane 2 the band is very faint, but present). Lane 4, gap (no sample). Lanes 5 and 6 are protein samples (65 µg and 195 µg, respectively) after gel-filtration.

protein was dialysed against TNE (20 m*M* Tris–HCl, 100 m*M* NaCl, 1 m*M* EDTA pH 8.0), concentrated to ∼1 ml and further purified by means of gel-filtration chromatography using Sephacryl S100 resin (Sigma) equilibrated with TNE. SDS–PAGE analysis (Fig. 2) shows samples of CRW-8 VP8*$_{64–224}$ before and after the final gel-filtration purification step. Removal of a higher molecular-weight band (just greater than 45 kDa) is apparent and is illustrated most clearly by comparison of the pre- and post-gel-filtration samples in lanes 3 and 6, respectively. Analysis of these samples by dynamic light scattering [CoolBatch +90T instrument (Precision Detectors); 293 K, protein concentration of 40 mg ml^{-1}, 20 repeats] corroborates these findings. The pre-gel-filtrated sample reveals a bimodal distribution, representing one protein species with a hydrodynamic radius (R_h) of 1.74 nm, corresponding to VP8*$_{64–224}$ (calculated R_h = 1.47 nm), and one larger aggregate, whereas the post-gel-filtered sample gives a monomodal distribution of VP8*$_{64–224}$. The polydispersity of the purified CRW-8 VP8*$_{64–224}$ peak is 22.4%, which falls in the range (15–30%) indicative of good homogeneity.

2.2. Crystallization

Prior to crystallization, the CRW-8 VP8*$_{64–224}$ protein was concentrated to 20 mg ml^{-1} in TNE and mixed with methyl α-D-*N*-acetylneuraminide (Neu5Acα2Me; Fig. 3) at 61 m*M* concentration, giving a ratio of 1:50 protein:ligand. The ligand was synthesized following published protocols (Kononov *et al.*, 1998). Initial crystallization conditions were obtained by screening using a sparse-matrix approach (Jancarik & Kim, 1991) with Crystal Screen and Crystal Screen II (Hampton Research) at 293 K *via* the hanging-drop vapour-diffusion technique. Rod-shaped crystals (Fig. 4) formed overnight from combination of an equal volume of pre-mixed protein–ligand solution and reservoir solution comprising 70% MPD, 0.1 *M* HEPES pH 7.5. These crystals grew to typical dimensions of 0.2 × 0.02 × 0.02 mm after 14 d at 293 K and were reproducible; however, over time substantial surface nucleation occurred, resulting in rod clusters. Optimization trials with variation of precipitant concentrations, inclusion of additives and at additional temperatures of 277 and 303 K has not yet produced substantial improvement of crystal quality. Macroseeding led to a gradual increase in crystal size but with associated excessive surface nucleation, resulting in numerous smaller rods protruding from the surface. A combination of a number of macroseeding steps interdispersed with short soaks of the crystal in reservoir solution in order to disrupt smaller rods from the crystal surface led to the production of a few crystals that appeared to be predominantly comprised of one single lattice.

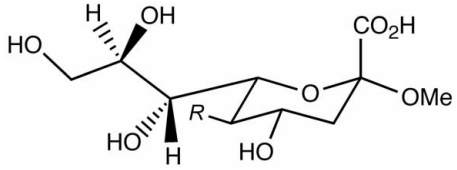


Figure 3
The ligand methyl α-D-*N*-acetylneuraminide (Neu5Acα2Me).

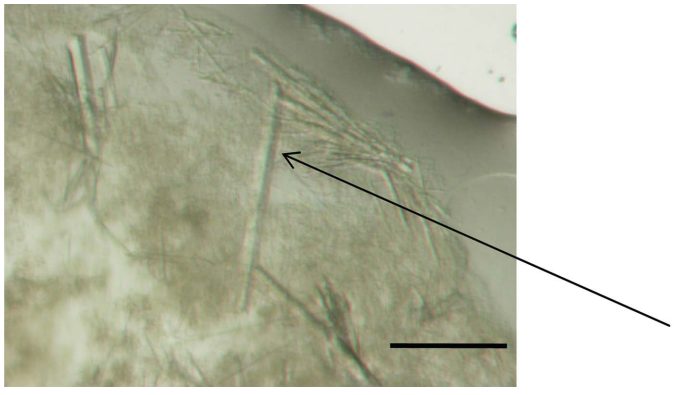


Figure 4
Crystals of porcine VP8*$_{64–224}$ grown *via* the hanging-drop vapour-diffusion method. The largest crystal depicted (a fragment of which was used in X-ray diffraction data collection) is 0.45 × 0.03 × 0.03 mm. The scale bar represents 0.2 mm.

Table 1
Diffraction data statistics.

Values in parentheses represent the highest resolution shell.

Resolution (Å)	20–2.3 (2.4–2.3)
Total No. of observations	343346 (17668)
No. of unique reflections	17704 (2545)
Redundancy	19.4 (6.9)
Completeness (%)	99.8 (99)
Average $I/\sigma(I)$	5.6 (2.1)
R_{merge} (%)†	12.5 (38.4)

† $R_{merge} = \sum|I - \langle I\rangle|/\sum\langle I\rangle$, where I is the intensity of each reflection.

2.3. X-ray diffraction analysis and structure determination

A rod-shaped crystal with dimensions ∼0.2 × 0.03 × 0.03 mm (a fragment from a crystal of dimensions 0.45 × 0.03 × 0.03 mm; Fig. 4) was briefly dipped into a cryoprotectant solution containing 15% glycerol, 70% MPD, 0.1 *M* HEPES pH 7.5 and flash-frozen in liquid nitrogen. X-ray diffraction data were collected (FIP-BM30A station of the European Synchrotron Radiation Facility; ESRF) at 100 K, with λ = 0.9794 Å and using a MAR CCD detector. The crystal-to-detector distance was 170 mm and 180 frames were collected with 1.0° oscillation and an exposure time of 60 s. The crystal diffracted to

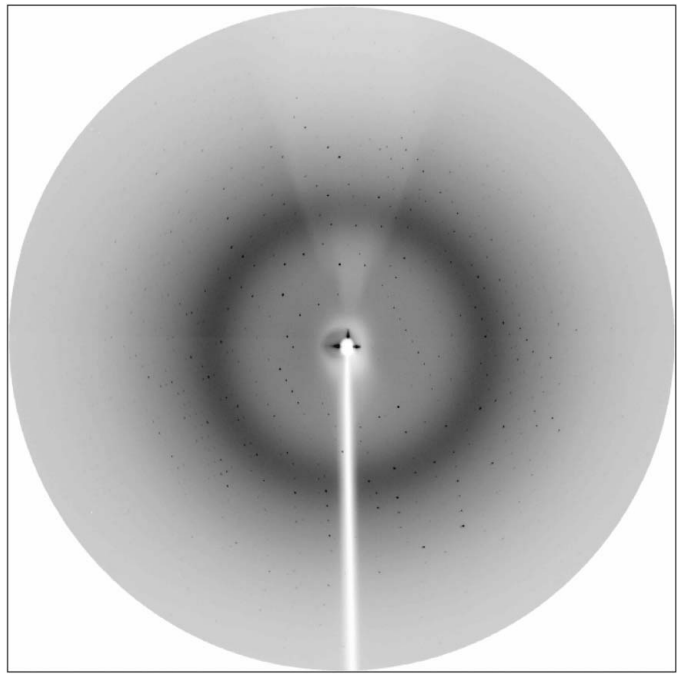


Figure 5
X-ray diffraction pattern from the crystal of CRW-8 VP8*$_{64–224}$. Synchrotron X-ray diffraction data were collected on a MAR CCD detector, set at distance of 170 mm. The edge of the image is 2.25 Å.

2.3 Å (Fig. 5) and the data were indexed and processed using *MOSFLM* (Leslie, 1992) and *SCALA* (*CCP*4 suite; Collaborative Computational Project, Number 4, 1994). Diffraction data statistics are summarized in Table 1. The crystal exhibits an orthorhombic system and belongs to space group $P2_12_12_1$, with unit-cell parameters a = 59.93, b = 64.66, c = 109.89 Å. Assuming two molecules per asymmetric unit, the Matthews coefficient (Matthews, 1968) V_M value is 2.6 Å^3 Da^{-1}, giving a solvent content of 53%.

A homology model of porcine CRW-8 VP8*$_{64–224}$ was built using *MODELLER* (Šali & Blundell, 1993) based on amino-acid sequence alignment with RRV VP8* (∼70% amino-acid sequence identity) and the three-dimensional protein structure of RRV VP8*$_{62–224}$ (Dormitzer, Sun, Wagner *et al.*, 2002). Using this as a molecular-replacement search model and *AMoRe* (Navaza, 1994), the rotation function calculated at 8–3 Å gave the highest two rotation peaks with R factors of 56.3 and 56.7% and correlation coefficients of 14.9 and 14.4%, respectively, with the next highest peak having an R factor of 56.9% and a correlation coefficient of 12.4%. Translation-function solutions were calculated for search models rotated to the top 20 rotation-function solutions, giving two solutions with R factors of 53.2 and 53.8% and correlation coefficients of 27 and 26.5%, respectively, significantly above the next peak with R factor 55.5% and correlation coefficient 16.8%. The top solution was fixed in position, rigid-body refinement was applied and the translation function was calculated for the remaining top five solutions. A clear solution (R factor 48.3% and correlation coefficient 41.5%) was obtained for the second molecule. Rigid-body refinement was applied to these two molecules, giving an R factor of 46.8% and a correlation coefficient of 49.1%. Examination of these solutions revealed good crystal packing and refinement of the model is in progress.

HB, BSC, MJK and MI gratefully acknowledge the financial support of the Australian Research Council (ARC). HB thanks Dr Jean-Luc Ferrer for assistance in the operation of X-ray diffraction data-collection equipment at beamline FIP-BM30A (ESRF). Fiona Fleming is acknowledged for providing assistance in cloning CRW-8 VP8*$_{62–224}$. BSC is supported by a Senior Research Fellowship from the National Health and Medical Research Council of Australia. MI also thanks the ARC for the award of an Australian Federation Fellowship.

References

Ciarlet, M. & Estes, M. K. (1999). *J. Gen. Virol.* **80**, 943–948.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Coulson, B. S (1993). *J. Clin Microbiol.* **31**, 1–8.
Coulson, B. S., Londrigan, S. H. & Lee, D. J. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 5389–5394.
Delorme, C., Brüssow, H., Sidoti, J., Roche, N., Karlsson, K.-A., Neeser, J.-R. & Teneberg, S. (2001). *J. Virol.* **75**, 2276–2287.
Dormitzer, P. R., Greenberg, H. B. & Harrison, S. C. (2001). *J. Virol.* **75**, 7339–7350.
Dormitzer, P. R., Sun, Z. Y., Blixt, O., Paulson, J. C., Wagner, G. & Harrison, S. C. (2002). *J. Virol.* **76**, 10512–10517.
Dormitzer, P. R., Sun, Z.-Y. J., Wagner, G. & Harrison, S. C. (2002). *EMBO J.* **21**, 885–897.
Dyall-Smith, M. L. & Holmes, I. H. (1984). *Nucleic Acids Res.* **12**, 3973–3982.
El-Attar, L., Dhaliwal, W., Howard, C. R. & Bridger, J. C. (2001). *Virology*, **291**, 172–182.
Estes, M. K. (2001). *Fields Virology*, 4th ed., edited by D. M. Knipe & P. M. Howley, p. 1747. Philadelphia: Lippincott–Raven.
Estes, M. K. & Cohen, J. (1989). *Microbiol. Rev.* **53**, 410–449.
Fazli, A., Bradley, S. J., Kiefel, M. J., Jolly, C., Holmes, I. H. & von Itzstein, M. (2001). *J. Med. Chem.* **44**, 3292–3301.
Fukudome, K., Yoshie, O. & Konno, T. (1989). *Virology*, **172**, 196–205.
Gerna, G., Sarasini, A., Parea, M., Arista, S., Miranda, P., Brussow, H., Hoshino, Y. & Flores, J. (1992). *J. Clin. Microbiol.* **30**, 9–16.
Guerrero, C. A., Zárate, S., Corkidi, G., López, S. & Arias, C. F. (2000). *J. Virol.* **74**, 9362–9371.
Guo, C. T., Nakagomi, O., Mochizuki, M., Ishida, H., Kiso, M., Ohta, Y., Suzuki, T. & Miyamoto, D. (1999). *J. Biochem.* **126**, 683–688.
Hewish, M. J., Takada, Y. & Coulson, B. S. (2000). *J. Virol.* **74**, 228–236.
Jancarik, J. & Kim, S.-H (1991). *J. Appl. Cryst.* **24**, 409–411.
Jolly, C. L., Beisner, B. M. & Holmes, I. H. (2000). *Virology*, **275**, 89–97.
Kiefel, M. J. & von Itzstein, M. (2003). *Methods Enzymol.* **363**, 395–412.
Kononov, L. O., Kornilov, A. V., Sherman, A. A., Zyrianov, E. V., Zatonskii, G. V., Shashkov, A. S. & Nifant'ev, N. E. (1998). *Bioorg. Khim.* **24**, 608–622.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.
Lundgren, O. & Svensson, L. (2001). *Microbes Infect.* **3**, 1145–1156.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Parashar, U. D., Hummelman, E. G., Bresee, J. S., Miller, M. A. & Glass, R. I. (2003). *Emerg. Infect. Dis.* **9**, 565–572.
Prasad, B. V. & Chiu, W. (1994). *Curr. Top. Microbiol. Immunol.* **185**, 9–29.
Rolsma, M. D., Kuhlenschmidt, T. B., Gelberg, H. B. & Kuhlenschmidt, M. S. (1998). *J. Virol.* **72**, 9079–9091.
Šali, A. & Blundell, T. L. (1993). *J. Mol. Biol.* **234**, 779–815.