Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray crystallographic analysis of the trimer core from measles virus fusion protein

Jieqing Zhu,[a] Yi Ding,[b] Feng Gao,[b] Tinghe Wu,[a] Catherine W.-H. Zhang,[c] Po Tien,[a] Zihe Rao[b] and George F. Gao[a,d]*

[a]Department of Molecular Virology, Institute of Microbiology, Chinese Academy of Sciences, Zhongguancun Beiyitiao, Beijing 100080, People's Republic of China, [b]Laboratory of Structural Biology and MOE Laboratory of Protein Sciences, School of Life Sciences and Bio-Engineering, Tsinghua University, Beijing 100084, People's Republic of China, [c]Oxagen Ltd, 91 Milton Park, Oxfordshire OX14 4RY, England, and [d]Nuffield Department of Clinical Medicine, John Radcliffe Hospital, University of Oxford, Oxford OX3 9DU, England

Correspondence e-mail: ggao66@yahoo.com

Received 26 November 2002
Accepted 10 January 2003

Two heptad-repeat regions (HR1 and HR2) are highly conserved in paramyxovirus fusion proteins and form a stable helical trimer of heterodimers [$(HR1–HR2)_3$] after the fusion between viral and cellular membranes. In this study, two HR regions of the fusion protein of measles virus, a member of the paramyxoviruses, were selected and overexpressed as a single chain (named 2-Helix) connected by an amino-acid linker using a GST-fusion expression system in *Escherichia coli*. Crystals of 2-Helix protein (GST removed) could be obtained from many conditions using the sitting- or hanging-drop vapour-diffusion method. A complete data set was collected in-house to 1.9 Å resolution from a single crystal. The crystal belongs to space group $P6$, with unit-cell parameters $a = b = 51.637$, $c = 67.058$ Å. To facilitate the crystal structure solution, SeMet-substituted 2-Helix crystals, grown under similar conditions to the native, were also obtained and diffracted X-rays to 1.8 Å using synchrotron radiation.

1. Introduction

Recent studies have shown that many enveloped viruses may adopt a similar molecular mechanism of virus entry (reviewed by Weissenhorn *et al.*, 1999; Lamb *et al.*, 1999; Skehel & Wiley, 2000; Bentz, 2000; Eckert & Kim, 2001). Fusion (F) proteins of the enveloped viruses undergo conformational changes to become active in mediating virus–cell fusion after binding to the cellular receptor(s), either by themselves or through another envelope-attachment protein. A large number of crystal structures of viral F proteins in the post-fusion state, including those of influenza haemagglutinin, human and simian inmmunodeficiency virus gp41, Moloney murine leukaemia virus TM, Ebola virus GP2, human T-cell leukaemia virus type 1 gp21, simian parainfluenza virus 5 F_1 and human respiratory syncytial virus F_1, indicate that high α-helix-content trimer or six-helix bundle formation of two heptad-repeat regions (HR1, also called HR-A, N-peptide, and HR2, also called HR-B, C-peptide) is a common character in F-protein-mediated cell fusion (Bullough *et al.*, 1994; Fass *et al.*, 1996; Weissenhorn *et al.*, 1997, 1998; Chan *et al.*, 1997; Malashkevich *et al.*, 1998, 1999; Caffrey *et al.*, 1998; Yang *et al.*, 1999; Kobe *et al.*, 1999; Baker *et al.*, 1999; Zhao *et al.*, 2000). In this structure, the fusion core of HR1 and HR2 forms a stable six-helix coiled coil centred by three HR1 and surrounded by three HR2 (a trimer of HR1–HR2 heterodimers). It has also been demonstrated that either HR1 or HR2, or both, will inhibit virus fusion/entry (Weissenhorn *et al.*, 1999; Lamb *et al.*, 1999; Eckert & Kim, 2001; Yu *et al.*, 2002). Hence, study of the crystal structure of the HR1/HR2 trimer will give us insight into the interaction of HR1 and HR2, thereby leading to the discovery of new drugs to block virus entry, which may include peptides, peptide analogues or small molecules.

Measles virus (MeV) is a member of the genus *Morbillivirus* in the family Paramyxoviridae (Lamb *et al.*, 2000). Fusion of MeV is mediated by the two envelope glycoproteins on the virus surface: H protein binds cellular receptors and triggers conformational change of F protein, leading to cell–virus membrane fusion (Horvath *et al.*, 1992; Lamb, 1993). The F protein is synthesized initially as a precursor F_0, which is cleaved into a disulfide-like heterodimer of F_1 and F_2 by the furin-linked enzyme in the host cell (Homma & Ohuchi, 1973; Scheid & Choppin, 1974). Two heptad-repeat domains in F_1, HR1, which is located at the carboxyl terminus of the fusion peptide, and HR2, which is located adjacent to the transmembrane domain, have been identified and characterized, consistent with the common features of paramyxoviruses (Lambert *et al.*, 1996; Wild & Buckland, 1997; Zhu *et al.*, 2002). Peptides corresponding to the HR2 region can also inhibit MeV fusion (Lambert *et al.*, 1996; Wild & Buckland, 1997).

Knowledge of the atomic structure of the MeV HR1/HR2 six-helix bundle would help further understanding of the molecular-fusion mechanism of paramyxoviruses and would help to facilitate the discovery of MeV fusion/entry inhibitors. Here, the predicted MeV

HR1/HR2 domains were expressed as a single chain (named 2-Helix) connected by a flexible amino-acid linker in the *Escherichia coli* system. The 2-Helix protein has been crystallized and preliminary X-ray analysis has been carried out.

2. Materials and methods

2.1. Gene construction, expression and protein purification

The expression, purification and preparation of soluble 2-Helix protein has been reported previously (Zhu *et al.*, 2002); a schematic 2-Helix construct showing the relative positions of the MeV F proteins is presented in Fig. 1. Briefly, *E. coli* strain BL21(DE3) transformed with the recombinant pGEX-6p-1 plasmid was grown at 310 K in 2×YTA medium to an optical density ($OD_{590\ nm}$) of 0.8–1.0 prior to induction with 1 m*M* IPTG for 4 h. Bacterial cells were harvested and lysed by sonication in phosphate-buffered saline (PBS, 10 m*M* sodium phosphate pH 7.3, 150 m*M* NaCl). Triton X-100 was then added to a final concentration of 1% and the lysate was incubated for 30 min at 273 K and subsequently clarified by centrifugation at 12 000*g* for 30 min at 277 K. The clarified supernatants were passed through a glutathione-Sepharose 4B column (equilibrated with PBS). The GST-fusion protein-bound column was washed with ten column volumes of PBS and eluted with three column volumes of reduced glutathione (10 m*M*). The GST-fusion proteins were then cleaved by GST-fusion rhinovirus 3C protease (kindly provided by Drs K. Hudson and J. Heath) at 278 K for 16 h in cleavage buffer (50 m*M* Tris–HCl pH 7.0, 150 m*M* NaCl, 1 m*M* DTT, 1 m*M* EDTA pH 8.0). GST and the GST-fusion rhinovirus 3C protease were removed by passage through a glutathione-Sepharose 4B column. 2-Helix protein was concentrated by ultrafiltration and further purified by gel filtration (Hiload Superdex G75, Pharmacia). The protein was dialysed into 20 m*M* Tris–HCl pH 8.0 and concentrated.

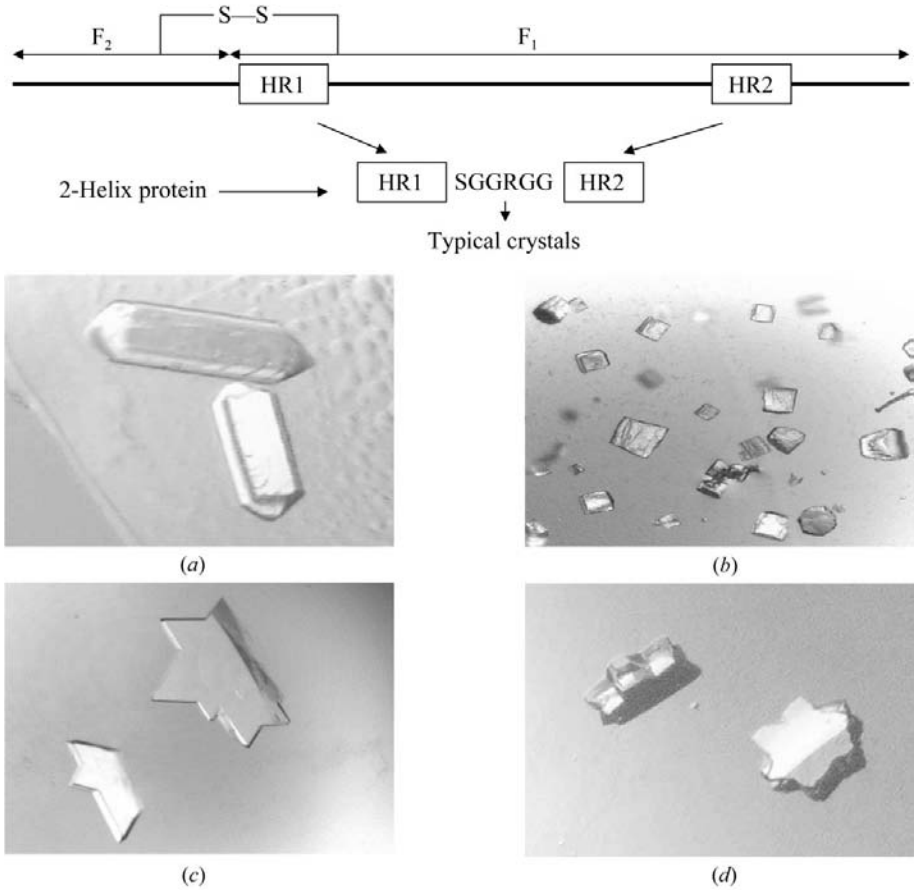

Figure 1
Schematic structure of the MeV fusion protein (F_1 and F_2), 2-Helix construct and typical crystals of the 2-Helix protein grown under different conditions. 'SGGRGG' is an amino-acid linker linking HR1 and HR2 in this trimer core. F_1 and F_2 are linked by a disulfide bond on the virus surface, indicated by 'S—S'. (*a*) Native crystals grown in 30% PEG 4000, 0.2 *M* $Li_2SO_4{\cdot}H_2O$, 0.1 *M* Tris–HCl pH 8.5. (*b*) Native crystals grown in 28% PEG 400, 0.2 *M* $CaCl_2{\cdot}2H_2O$, 0.1 *M* Na HEPES pH 7.5. (*c*) Native crystals grown in 19% PEG 400, 0.2 *M* $CaCl_2{\cdot}2H_2O$, 0.1 *M* Na HEPES pH 7.5. (*d*) SeMet crystals grown in 26% PEG 400, 0.2 *M* $CaCl_2{\cdot}2H_2O$, 0.1 *M* Na HEPES.

2.2. Preparation of selenomethionine-labelled proteins

SeMet-labelled 2-Helix protein was expressed and purified using the same GST-fusion expression system. *E. coli* strain BL21(DE3) transformed with the recombinant pGEX-6p-1 plasmid was grown overnight at 310 K in LB medium containing 100 µg ml^{-1} ampicillin and then diluted with adaptive medium (20% LB medium plus 80% M9 medium plus 5% glucose) and grown at 310 K to an $OD_{600\ nm}$ of 0.6–0.8. The bacterial cells were harvested and resuspended with 5–10 ml expression medium (1% YNB, 5% glucose in M9 medium), transferred to a large volume of expression medium and grown to an $OD_{600\ nm}$ of 0.6–0.8. L-Selenomethionine at 60 mg l^{-1}, lysine, threonine and phenylalanine at 100 mg l^{-1} and leucine, iosleucine and valine at 50 mg l^{-1} were added and incubated at 310 K for 15 min before induction with 1 m*M* IPTG at 301 K for 10 h. The bacterial cells were harvested and the SeMet 2-Helix protein was purified using the same method as for the native protein. Molecular weights of the native and SeMet-labelled proteins were compared by MALDI–TOF mass spectroscopy using a Bruker Daltonics Biflex III MALDI–TOF Mass Spectrometer.

2.3. Crystallization

GST-removed 2-Helix protein in 20 m*M* Tris–HCl pH 8.0 was concentrated to 10 mg ml^{-1}. Initial crystallization conditions were screened using the hanging-drop vapour-diffusion method with sparse-matrix crystallization kits (Crystal Screen I and II, Hampton Research, Riverside, CA, USA). 1.5 µl of protein solution was mixed with 1.5 µl of reservoir solution and equilibrated against 0.3 ml of reservoir solution. 2-Helix protein crystals were obtained from at least four conditions and all these conditions were optimized in order to obtain the best diffracting crystals in both hanging and sitting drops.

2.4. X-ray data collection and analysis

X-ray diffraction data for a single crystal grown in 19% PEG 400, 0.2 *M* $CaCl_2$, 0.1 *M* Na HEPES pH 7.5 using the sitting-drop method were collected at 100 K with a MAR Research image-plate detector using Cu $K\alpha$ radiation from a Rigaku rotating-anode

Table 1
Diffraction data statistics.

Values in parentheses correspond to the highest resolution shell.

	Native	Peak	Inflection	Remote
Space group	*P*6			
Unit-cell parameters (Å)	$a = b = 51.6$, $c = 67.1$			
Resolution (Å)	50–1.9 (1.97–1.90)	50–1.8 (1.86–1.80)		
Total observations	57580	88961	88514	91147
Unique reflections	7997	9100	9015	9175
Redundancy	7.2 (7.0)	9.8 (7.3)	9.8 (7.8)	9.9 (8.8)
Average $I/\sigma(I)$	20.4 (16.1)	19.1 (6.1)	19.3 (7.3)	20.4 (9.4)
$R_{\rm merge}$† (%)	0.076 (0.153)	0.079 (0.342)	0.073 (0.319)	0.070 (0.251)
Completeness (%)	99.0 (95.3)	95.3 (89.4)	95.6 (90.8)	96.0 (95.6)

† $R_{\rm merge} = \sum |I - \langle I \rangle| / \sum I$.

generator. The SeMet protein crystal diffraction data were collected using synchrotron radiation at the Advanced Photon Source (APS), Argonne National Laboratory, USA. The data were processed with *DENZO* and *SCALEPACK* (Otwinowski, 1993; Otwinowski & Minor, 1997).

3. Results and discussion

The 2-Helix protein was soluble and stable in PBS or Tris–HCl buffers. Crystals were obtained from at least four conditions in 7 d at 291 K: (i) 20% PEG 8000, 0.2 *M* $Mg(OAc)_2{\cdot}4H_2O$, 0.1 *M* sodium cacodylate pH 6.5; (ii) 8% PEG 4000, 0.1 *M* $NaOAc{\cdot}3H_2O$ pH 4.6; (iii) 30% PEG 4000, 0.2 *M* $Li_2SO_4{\cdot}H_2O$, 0.1 *M* Tris–HCl pH 8.5; (iv) 28% PEG 400, 0.2 *M* $CaCl_2{\cdot}2H_2O$, 0.1 *M* Na HEPES pH 7.5. However, the crystals grown from conditions (i) and (ii) diffracted X-ray very poorly as they were too small or poorly shaped. Rod-shaped crystals were obtained from condition (iii) (Fig. 1*a*), but did not diffract well, possibly because of instability. Crystals grown from condition (iv) (Fig. 1*b*) were slightly mosaic according to X-ray diffraction. The best crystals were obtained by decreasing the concentration of PEG 400 to 18–19% in condition (iv) and grew in 4 d at 291 K (Fig. 1*c*). The crystals diffract X-rays to 1.9 Å resolution (Fig. 2) and belong to space group *P*6, with unit-cell parameters $a = b = 51.637$, $c = 67.058$ Å. Data-collection statistics are given in Table 1. The 00*l* reflections are listed in Table 2, indicating that there are no systematic absences along the screw axes. There is one molecule in the asymmetric unit. The Matthews coefficient is about 2.7 $Å^3\,Da^{-1}$ and the solvent content is about 55% (Matthews, 1968). As demonstrated above, the expected active form is a trimer. The triple axis of this trimer should superpose with the crystallographic triple axis in the *P*6 space group. The SeMet crystals (Fig. 1*d*) could also be obtained in 4 d from condition (iv) with a PEG 400 concentration of 25–27% at 291 K. The SeMet crystals diffracted to 1.8 Å resolution using synchrotron radiation. Data-collection statistics of the SeMet crystals are given in Table 1. Mass-spectroscopy analysis showed that the molecular weight of 2-Helix protein was 9310 Da, whereas the molecular weight of the SeMet-substituted 2-Helix protein was 9403 Da, indicating that two methionines are replaced by SeMet as expected in the 2-Helix protein (Fig. 3) [molecular weight increases by $2 \times (79 - 32) = 94$ Da].

Table 2
The 00*l* reflections.

0 0 4	13679.9	891.5
0 0 5	32001.8	1931.4
0 0 6	30367.9	1944.2
0 0 7	2032.7	133.7
0 0 8	82622.4	4991.7
0 0 9	36801.5	2311.8
0 0 10	230570	14170
0 0 11	74558.8	4364.3
0 0 12	154061	9324
0 0 13	54951.6	3324.3
0 0 14	52896.4	3149.0
0 0 15	22496.9	1306.0
0 0 16	36664.6	2241.6
0 0 17	52951.4	5429.5
0 0 18	75339.3	7673.2
0 0 19	23401.6	2469.7
0 0 20	31243.7	3146.3
0 0 21	49314.2	5008.2
0 0 22	83154.1	8594.6
0 0 23	121125	12216
0 0 24	615.6	138.4
0 0 25	21486.2	2625.3
0 0 26	50507.4	6136.6
0 0 27	15019.2	1820.7
0 0 28	2378.4	344.9
0 0 29	2900.4	446.0
0 0 30	31536.3	4479.1
0 0 31	3944.0	590.2
0 0 32	5569.9	817.4
0 0 33	5609.5	1031.9
0 0 34	19371.9	3515.8
0 0 35	391.2	171.1

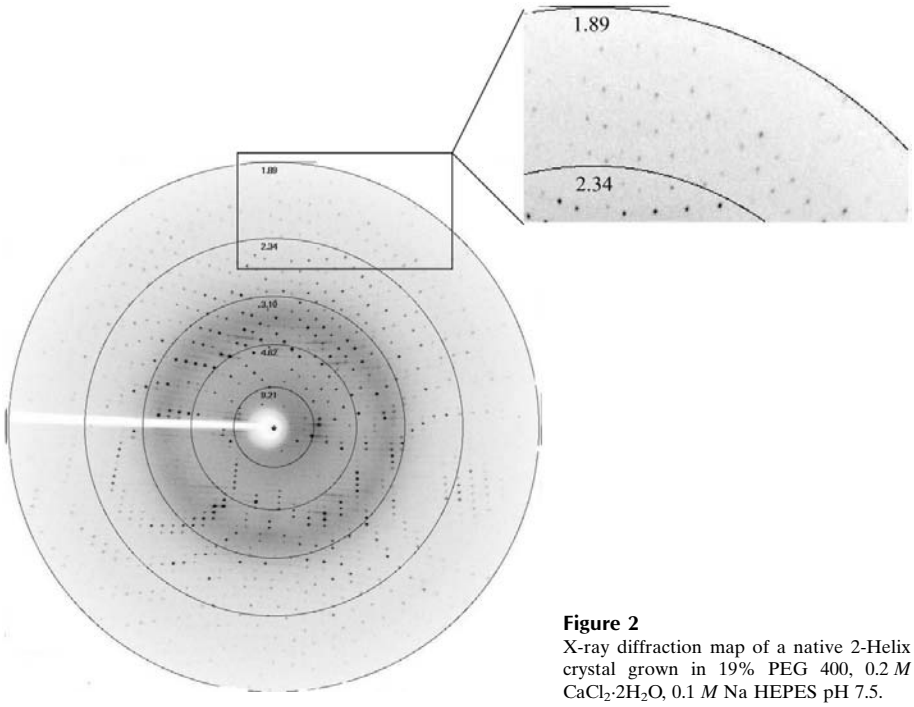

Figure 2
X-ray diffraction map of a native 2-Helix crystal grown in 19% PEG 400, 0.2 *M* $CaCl_2{\cdot}2H_2O$, 0.1 *M* Na HEPES pH 7.5.

Initial molecular-replacement calculations were performed using the program *AMoRe*, (Navaza, 1994) with the crystal structure of paramyxovirus SV5 N1/C1 trimer as a search model (PDB code 1svf), but no distinct peaks were obtained. Different models and different programs are being tried in order to solve the structure.

To study the interaction and structure of the HR regions from fusion proteins, we could use chemical synthesis or express HR1 and HR2 separately in the *E. coli* system and then mix the synthetic or purified peptides *in vitro*. However, the cost of chemical synthesis was too high and purification of HR1 and HR2 separately was problematic, as the single HR1 peptides had a strong tendency

to aggregate, possibly because of their hydrophobic nature. In this study, we expressed HR1 and HR2 as a single chain (2-Helix) connected by a flexible amino-acid linker in high yield using a GST-fusion expression system. The 2-Helix protein was soluble and could easily be purified. The 2-Helix protein was also very stable owing to the formation of a trimer during the expression or purification step. From this study, it is clearly shown that another advantage of using the 2-Helix protein is its easy crystallization under many conditions, which provides a convenient method to study the crystal structure of the HR1/HR2 trimer from viral fusion proteins.

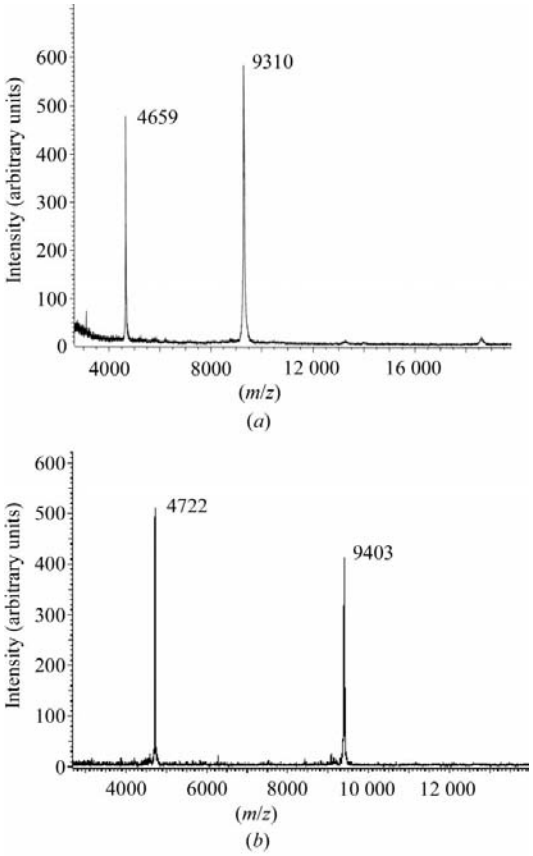

Figure 3
Mass spectra of native and SeMet-labelled 2-Helix proteins. The molecular weight of native 2-Helix protein was 9310 Da (*a*). The molecular weight of SeMet 2-Helix protein was 9403 Da (*b*). The first peak in both cases indicates that the proteins carry two charges.

This work was supported by a grant from the National Frontier Research Programme (Project 973) of the Ministry of Science and Technology of the People's Republic of China (grant number G1999011902). We thank Drs K. Hudson and J. Heath for their generous supply of the rhinovirus 3C protease GST fusion plasmid and Pengyun Li and Beili Wu for their help in crystallization trials. We are grateful to Professor Qing-Ge Xie, the principal scientist of Project 973 (G19990119), for his support.

References

Baker, K. A., Dutch, R. E., Lamb, R. A. & Jardetzky, T. S. (1999). *Mol. Cell*, **3**, 309–319.
Bentz, J. (2000). *Biophys. J.* **78**, 886–900.
Bullough, P. A., Hughson, F. M., Skehel, J. J. & Wiley, D. C. (1994). *Nature (London)*, **371**, 37–43.
Caffrey, M., Cai, M., Kaufman, J., Stahl, S. J., Wingfield, P. T., Covell, D. G., Gronenborn, A. & Clore, G. M. (1998). *EMBO J.* **17**, 4572–4584.
Chan, D. C., Fass, D., Berger, J. M. & Kim, P. S. (1997). *Cell*, **89**, 263–273.
Eckert, D. M. & Kim, P. S. (2001). *Annu. Rev. Biochem.* **70**, 777–810.
Fass, D., Harrison, S. C. & Kim, P. S. (1996). *Nature Struct. Biol.* **3**, 465–469.
Homma, M. & Ouchi, M. (1973). *J. Virol.* **12**, 1457–1465.
Horvath, C. M., Paterson, R. G., Shaughnessy, M. A., Wood, R. & Lamb, R. A. (1992). *J. Virol.* **66**, 4564–4569.
Kobe, B., Center, R. J., Kemp, B. E. & Poumbouris, P. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 4319–4324.
Lamb, R. A. (1993). *Virology*, **197**, 1–11.
Lamb, R. A., Collins, P. L., Kolakofsky, D., Melero, J. A., Nagai, Y., Oldstone, M. B. A., Pringle, C. R. & Rima, B. K. (2000). *Virus Taxonomy Classicification and Nomenclature of Viruses, 7th Report*, edited by M. H. V. van Regenmortel, C. M. Fauquet, D. H. L. Bishop, E. B. Carstens, M. K. Estes, S. M. Lemon, J. Maniloff, M. A. Mayo, D. J. McGeoch, C. R. Pringle & R. B. Wickner, pp. 549–561. San Diego: Academic Press.
Lamb, R. A., Joshi, S. B. & Dutch, R. E. (1999). *Mol. Membr. Biol.* **16**, 11–19.
Lambert, D. M., Barney, S., Lambert, A. L., Guthrie, K., Medinas, R., Davis, D. E., Bucy, T., Erickson, J., Merutka, G. & Petteway, S. R. Jr (1996). *Proc. Natl Acad. Sci. USA*, **93**, 2186–2191.
Malashkevich, V. N., Chan, D. C., Chutkowski, C. T. & Kim, P. S. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 9134–9139.
Malashkevich, V. N., Schneider, B. J., McNally, M. L., Milhollen, M. A., Pang, J. X. & Kim, P. S. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 2662–2667.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Otwinowski, Z. (1993). *Proceedings of the CCP4 Study Weekend. Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 56–62. Warrington: Daresbury Laboratory.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Scheid, A. & Choppin, P. W. (1974). *Virology*, **57**, 475–490.
Skehel, J. J. & Wiley, D. C. (2000). *Annu. Rev. Biochem.* **69**, 531–569.
Weissenhorn, W., Carfi, A., Lee, K.-H., Skehel, J. J. & Wiley, D. C. (1998). *Mol. Cell*, **2**, 605–616.
Weissenhorn, W., Dessen, A., Calder, L. J., Harrison, S. C., Skehel, J. J. & Wiley, D. C. (1999). *Mol. Membr. Biol.* **16**, 3–9.
Weissenhorn, W., Dessen, A., Harrison, S. C., Skehel, J. J. & Wiley, D. C. (1997). *Nature (London)*, **387**, 426–430.
Wild, T. F. & Buckland, R. (1997). *J. Gen. Virol.* **78**, 107–111.
Yang, Z. N., Mueser, T. C., Kaufman, J., Stahl, S. J., Wingfield, P. T. & Hyde, C. C. (1999). *J. Struct. Biol.* **126**, 131–144.
Yu, M., Wang, E., Liu, Y., Cao, D., Jin, N., Zhang, C. W., Bartlam, M., Rao, Z., Tien, P. & Gao, G. F. (2002). *J. Gen. Virol.* **83**, 623–629.
Zhao, X., Singh, M., Malashkevich, V. N. & Kim, P. S. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 14172–14177.
Zhu, J., Zhang, C. W.-H., Qi, Y., Tien, P. & Gao, G. F. (2002). *Biochem. Biophys. Res. Commun.* **299**, 897–902.