Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray diffraction analysis of post-fusion six-helix bundle core structure from Newcastle disease virus F protein

Peng-Yun Li,[a] Jie-Qing Zhu,[b] Bei-Li Wu,[a] Feng Gao,[a] Po Tien,[b] Zihe Rao[a] and George F. Gao[b,c]*

[a]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, [b]Department of Molecular Virology, Institute of Microbiology, Chinese Academy of Sciences, Zhongguancun Beiyitiao, Beijing 100080, People's Republic of China, and [c]Nuffield Department of Clinical Medicine, John Radcliffe Hospital, University of Oxford, Oxford OX3 9DU, England

Correspondence e-mail: george.gao@ndm.ox.ac.uk

Received 21 March 2003
Accepted 6 May 2003

Fusion of virus members from the *Paramyxoviridae* family involves two glycoproteins. They are termed attachment glycoprotein (HN, H or G) and fusion protein (F). The F protein contains two highly conserved heptad-repeat (HR) regions, HR1 and HR2. Through conformational changes in the F protein, HR1 and HR2 are believed to form a stable six-helix coiled-coil bundle during the membrane-fusion process. However, no crystal structure has yet been documented for this state in the Newcastle disease virus (NDV, a member of the *Paramyxoviridae* family) F protein, despite the recent success on its F_0 crystal structure (Chen *et al.*, 2001), which was thought to represent the pre-fusion conformation of F glycoprotein. In this study, a single-chain polypeptide constructed by linking two truncated HR regions of the NDV F protein has been expressed, purified and crystallized by means of the hanging- or sitting-drop vapour-diffusion method. Crystals in hexagonal and trapezoid forms with a resolution limit of 2.6 Å were obtained. These crystals belonged to space group $C2$, with unit-cell parameters a = 66.4, b = 38.2, c = 102.0 Å, β = 100.2°. Crystals in the rhombic form with a resolution limit of 2.5 Å were also obtained. These crystals belonged to space group $P2_1$, with unit-cell parameters a = 59.0, b = 31.9, c = 62.3 Å, β = 117.0°. This will add to the repertoire of viral fusion protein post-fusion state structures and help further the understanding of the molecular mechanism of enveloped virus fusion.

1. Introduction

Membrane fusion is a critical process during enveloped virus infection, through which the virus envelope membrane and cellular membrane fuse together to become a single membrane and the viral genetic materials (in the form of nucleocapsid) are introduced into the cell (Richardson *et al.*, 1986; Yeagle *et al.*, 1991; Blissard & Wenz, 1992; Lamb, 1993). The fusion is initiated *via* binding of the viral envelope protein to the cellular receptor, mediated directly by the viral envelope fusion protein (Lamb, 1993). The receptor-binding envelope protein and fusion envelope protein can be either two different glycoproteins or a single glycoprotein with bipartisan functions depending on the virus family. It has been shown that the majority of the fusion proteins of the enveloped viruses contain two highly conserved heptad-repeat (HR) regions: HR1 and HR2. Previous studies of several viral envelope fusion proteins using biochemical and structural techniques indicate that many enveloped viruses contain a 'trimer of hairpins' motif formed by HR1 and HR2 polypeptides (Lamb *et al.*, 1999; Weissenhorn *et al.*, 1999; Bentz, 2000; Skehel & Wiley, 2000; Eckert & Kim, 2001). This motif is comprised of a homotrimeric coiled coil, formed by the N-terminal region of the protein (HR1) surrounded by three C-terminal regions (HR2) that pack against the coiled coil in an oblique antiparallel manner (a trimer of HR1/HR2 heterodimers). The function of this structure is to bring the viral and cellular membranes together to allow fusion to take place. This is because they exist as an intermediate pre-hairpin conformation when the fusion peptide, located adjacent to the HR1, penetrates into the cellular membrane (Lamb *et al.*, 1999; Eckert & Kim, 2001). At present, the HR1 and HR2 regions are virus-fusion inhibition targets for inhibitor screening, as under this model the introduced exogenous HR1 or HR2 would compete for binding to the endogenous counterpart in the intermediate pre-hairpin state, thereby preventing formation of the stable six-helix bundle.

Newcastle disease virus (NDV) belongs to the *Paramyxoviridae* family of enveloped negative-stranded RNA viruses (Lamb *et al.*, 2000). A surface glycoprotein termed 'F' or fusion protein mediates membrane fusion during the NDV infection process, while the hemagglutinin-neuramidase (termed 'HN', the attachment protein) is responsible for receptor binding (Lamb, 1993; Lamb *et al.*, 2000). The F

protein is synthesized as a precursor F_0, which is activated upon proteolytic cleavage to form two disulfide-bond-linked chains: F_1 (C-terminus) and F_2 (N-terminus) (Homma & Ohuchi, 1973). The F_1 polypeptide contains two heptad-repeat domains: HR1, located at the N-terminus, and HR2, located at the C-terminus (Young *et al.*, 1997, 1999). The HR1 and HR2 peptides can interact with each other to form a six-helix bundle (Young *et al.*, 1997, 1999; Yu *et al.*, 2002; Zhu, Li *et al.*, 2003), which represents the post-fusion conformation of the F protein. Conformational changes in the F protein are believed to be the key process in virus fusion (Lamb, 1993). There are at least three conformations under the current model: the pre-fusion native state, the pre-hairpin intermediate state and the post-fusion six-helix bundle state (Eckert & Kim, 2001). Recently, the crystal structure of NDV F_0 has been resolved (Chen *et al.*, 2001), although the HR2 region is lacking from the structure, and it is believed that this represents the pre-fusion conformation. To better understand the molecular mechanism of NDV fusion and facilitate fusion-inhibitor design, the crystal structure of the six-bundle of HR1/HR2 is of great importance. This will be the second resolved structure of its type following the influenza haemagglutinin protein (Wilson *et al.*, 1981; Bullough *et al.*, 1994) and will allow a detailed comparison between two different conformations of the same protein. Here, we report the crystallographic study of a single-chain polypeptide constructed by linking two truncated HR regions of the NDV F protein. This will add to the repertoire of viral 'trimer of hairpins' structures for the elucidation of the molecular mechanism of enveloped virus entry.

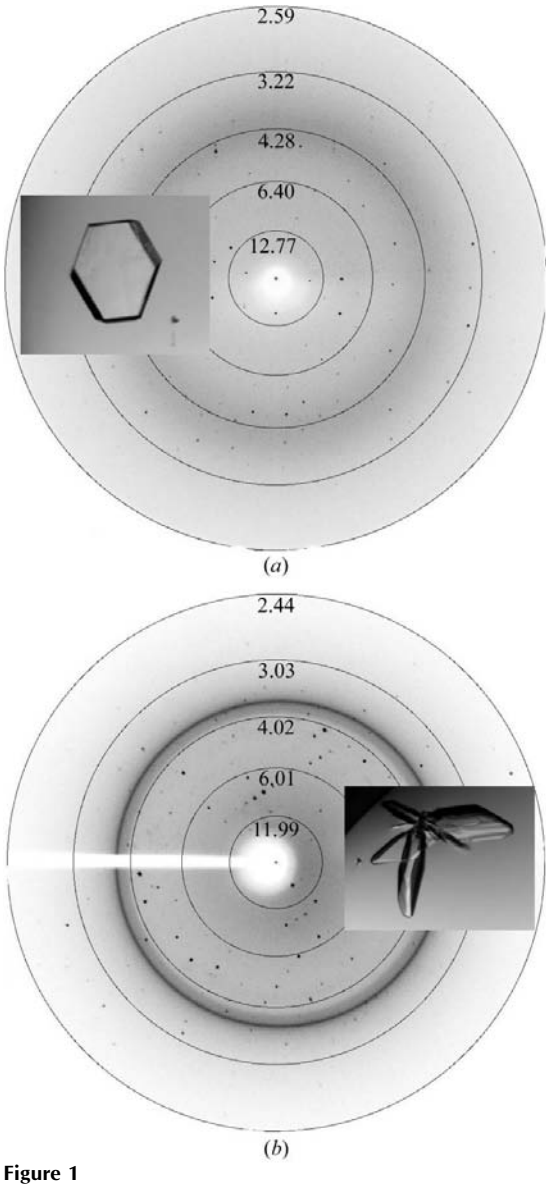

Figure 1
Diffraction patterns and photos (inset) of crystals obtained from 0.1 *M* sodium acetate pH 4.6, 0.2 *M* calcium chloride, 10–15% 2-propanol (*a*) and 0.1 *M* sodium cacodylate pH 6.5, 0.2 *M* calcium acetate, 10–15% PEG 8000 (*b*).

2. Materials and methods

2.1. Protein expression and purification

The construct production and protein expression/purification have been reported previously (Zhu, Li *et al.*, 2003). Briefly, the NDV F gene was cloned from the Chinese virulent isolate F48E9 (GenBank No. AF079172). The HR1 region used was derived from amino acids 115–152 and HR2 was derived from amino acids 442–477 of the F glycoprotein. The 2-Helix construct was made by linking the HR1 and HR2 with a six amino-acid linker (SGGRGG using single-letter amino-acid abbreviations). The expression vector pGEX-6p-1 (Pharmacia) was used to clone the 2-Helix PCR product by unique *Bam*HI and *Xho*I restriction sites introduced by PCR primers. The GST-fusion protein was purified by GST-affinity chromatography and the GST tag was removed by GST-fusion rhinovirus 3C protease cleavage (kindly provided by Drs J. Heath and K. Hudson). The GST-removed protein was further purified by sequential Mono Q and Superdex G75 columns. In both columns the 2-Helix proteins appeared as a single peak. The column-purified protein had a purity of >98% and was used for further crystallization experiments. This 2-Helix protein was shown to form a six-helix bundle (Zhu, Li *et al.*, 2003).

2.2. Crystallization

The purified protein (in 20 m*M* Tris–HCl pH 8.0) was concentrated to 10–15 mg ml^{-1}. Initial crystallization conditions were screened using crystallization screening kits (Hampton Research, Riverside CA, USA). The protein was crystallized under several conditions. Conditions yielding small crystals were further optimized by varying the precipitant, protein concentration and buffer pH. Crystals of good quality could be obtained from two conditions: (i) 0.1 *M* sodium acetate pH 4.6, 0.2 *M* calcium chloride, 10–15% 2-propanol and (ii) 0.1 *M* sodium cacodylate pH 6.5, 0.2 *M* calcium acetate, 10–15% PEG 8000. Crystallization was performed by the sitting-drop vapour-diffusion method at 291 K to produce further diffractable crystals using the above two conditions. 1 µl of protein solution was mixed with 1 µl of reservoir solution and the mixture was equilibrated against 500 µl reservoir solution at 291 K. Hexagonal and trapeziform crystals appeared in 3 d from the first condition (Fig. 1*a*) and rhombic crystals appeared in two weeks from the second condition (Fig. 1*b*).

2.3. X-ray data collection and processing

Data were collected using a 345 mm MAR Research image-plate system mounted on a Rigaku RU-2000 rotating-anode generator operated at 48 kV and 98 mA (Cu $K\alpha$; λ = 1.5418 Å). During the data collection, the crystal was maintained at 100 K using an Oxford Cryosystem with a cryoprotectant prepared by adding 20% glycerol to the mother buffer (see above). Data were indexed and scaled using *DENZO* and *SCALEPACK* programs (Otwinowski & Minor, 1997).

3. Results and discussion

The 2-Helix protein of NDV F glycoprotein could form a stable six-helix bundle and could also be crystallized under several conditions (Zhu, Li *et al.*, 2003). Nevertheless, good diffractable crystals could only be obtained under two sets of conditions: 0.1 *M* sodium acetate pH 4.6, 0.2 *M* calcium

Table 1
Diffraction data statistics.

Values in parentheses correspond to the highest resolution shell.

Space group	$P2_1$	$C2$
Unit-cell parameters		
a (Å)	59.0	66.4
b (Å)	31.9	38.2
c (Å)	62.3	102.0
β (°)	117.0	100.2
Resolution limits (Å)	40–2.5 (2.59–2.5)	4.0–2.6 (2.66–2.6)
Total observations	35641	23003
Unique reflections	7041 (719)	8769 (509)
Average $I/\sigma(I)$	15.0 (10.2)	14.3 (8.4)
Redundancy	5.1 (2.2)	2.9 (2.6)
Completeness (%)	96.9 (100)	98.9 (96.4)
$R_{\rm merge}$† (%)	4.4 (18.2)	9.6 (20.4)

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

chloride, 10–15% 2-propanol and 0.1 *M* sodium cacodylate pH 6.5, 0.2 *M* calcium acetate, 10–15% PEG 8000. In this study, better crystals could be obtained using the sitting-drop method rather than the hanging-drop method under both conditions. Crystals from the first set of conditions belong to the space group *C*2, with unit-cell parameters a = 66.4, b = 38.2, c = 102.0 Å, β = 100.2°. Assuming there to be three molecules in the asymmetric unit, the solvent content is calculated to be about 50%. The best crystals were able to diffract X-rays to 2.6 Å (Fig. 1*a*). Crystals from the second set of conditions belong to space group $P2_1$, with unit-cell parameters a = 59.0, b = 31.9, c = 62.3 Å, β = 117°. The best crystals from this group diffracted X-rays to 2.5 Å resolution (Fig. 1*b*). For these crystals, there were three molecules in the asymmetric unit and the solvent content was calculated to be about 40%. The data statistics for these two crystal forms are reported in Table 1. Determination of the structure of the protein is under way using the molecular-replacement method. The resolution of this six-helix core structure will provide a second example of its type after the influenza haemagglutinin (HA structures in neutral pH and low pH; Wilson *et al.*, 1981; Bullough *et al.*, 1994) and will provide the basis for a detailed comparison of different conformations of the NDV F glycoprotein owing to the recent breakthrough of the F_0 crystal structure (Chen *et al.*, 2001).

This report, together with our earlier studies on HR1/2 trimer core crystallization of measles virus F glycoprotein (Zhu *et al.*, 2002; Zhu, Ding *et al.*, 2003), demonstrate that the 2-Helix construct strategy by linking the HR1 and HR2 together using a flexible linker (SGGRGG) is a plausible method of obtaining crystallizable six-helix trimer core of paramyxovirus F glycoproteins. Therefore, it will add to the repertoire of the trimer core structure if more and more trimer cores can be approached. This will inevitably illuminate the molecular mechanism of paramyxovirus fusion as the structure data are accumulated together with the two 'trimer of hairpins' structures that have already been resolved for human respiratory syncytial virus (hRSV; Zhao *et al.*, 2000) and simian parainfluenza virus 5 (SV5; Baker *et al.*, 1999) from the *Paramyxoviridae* family.

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 No. G1999011902). Dr Yiwei Liu and Yi Ding of the Rao Lab are acknowledged for their help in data analysis. We are grateful to Professor Qing-Ge Xie, the principal scientist of Project 973 (G19990119), for his support. We thank David Cole and Mark Bartlam for critical reading of the manuscript.

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.
Blissard, G. W. & Wenz, J. R. (1992). *J. Virol.* **66**, 6829–6835.
Bullough, P. A., Hughson, F. M., Skehel, J. J. & Wiley, D. C. (1994). *Nature (London)*, **371**, 37–43.
Chen, L., Gorman, J. J., McKimm-Breschkin, J., Lawrence, L. J., Tulloch, P. A., Smith, B. J., Colman, P. M. & Lawrence, M. C. (2001). *Structure*, **9**, 255–266.
Eckert, D. M. & Kim, P. S. (2001). *Annu. Rev. Biochem.* **70**, 777–810.
Homma, M. & Ohuchi, M. (1973). *J. Virol.* **12**, 1457–1465.
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.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Richardson, C., Hull, D., Greer, P., Hasel, K., Berkovich, A., Englund, G., Bellini, W., Rima, B. & Lazzarini, R. (1986). *Virology*, **155**, 508–523.
Skehel, J. J. & Wiley, D. C. (2000). *Annu. Rev. Biochem.* **69**, 531–569.
Weissenhorn, W., Dessen, A., Calder, L. J., Harrison, S. C., Skehel, J. J. & Wiley, D. C. (1999). *Mol. Membr. Biol.* **16**, 3–9.
Wilson, I. A., Skehel, J. J. & Wiley, D. C. (1981). *Nature (London)*, **289**, 366–373.
Yeagle, P. L., Epand, R. M., Richardson, C. D. & Flanagan, T. D. (1991). *Biochim. Biophys. Acta*, **1065**, 49–53.
Young, J. K., Hicks, R. P., Wright, G. E. & Morrison, T. G. (1997). *Virology*, **238**, 291–304.
Young, J. K., Li, D., Abramowitz, M. C. & Morrison, T. G. (1999). *J. Virol.* **73**, 5945–5956.
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., Ding, Y., Gao, F., Wu, T., Zhang, C. W.-H., Tien, P., Rao, Z. & Gao, G. F. (2003). *Acta Cryst.* D**59**, 587–590.
Zhu, J., Li, P., Wu, T., Gao, F., Ding, Y., Zhang, C. W.-H., Rao, Z., Gao, G. F. & Tien, P. (2003). In the press.
Zhu, J., Zhang, C. W.-H., Qi, Y., Tien, P. & Gao, G. F. (2002). *Biochem. Biophys. Res. Commun.* **299**, 897–902.