Design and analysis of post-fusion 6-helix bundle of heptad repeat regions from Newcastle disease virus F protein

Jieqing Zhu1, Pengyun Li2, Tinghe Wu1, Feng Gao2, Yi Ding2, Catherine W.-H. Zhang3,4, Zihe Rao2, George F. Gao5,6 and Po Tien1,6

1Department of Molecular Virology, Institute of Microbiology, Chinese Academy of Sciences, Zhongguancun Beiyitiao, Beijing 100080,
2Laboratory of Structural Biology and MOE Laboratory of Protein Sciences, School of Life Sciences and Bio-Engineering, Tsinghua University, Beijing 100084, China,
3Laboratory of Immunobiology, Dana-Farber Cancer Institute and Department of Medicine, Harvard Medical School, Boston, MA 02115, USA
4Present address: Oxagen Ltd, Milton Park, Oxfordshire OX14 4RY, UK
5Department of Medical Genetics and Department of Biochemistry, School of Medicine, The Chinese University of Hong Kong, Shatin, NT, Hong Kong
6To whom correspondence should be addressed.

E-mail: gga066@yahoo.com; tiempo@sun.im.ac.cn

Fusion of paramyxovirus to the cell involves receptor binding of theHN glycoprotein and a number of conformational changes of F glycoprotein. The F protein is expressed as a homotrimer on the virus surface. In the present model, there are at least three conformations of F protein, i.e. native form, pre-hairpin intermediate and the post-fusion state. In the post-fusion state, the two highly conserved heptad repeat (HR) regions of F protein form a stable 6-helix coiled-coil bundle. However, no crystal structure is known for this state for the Newcastle disease virus, although the crystal structure of the F protein native form has been solved recently. Here we deployed an Escherichia coli in vitro expression system to engineer this 6-helix bundle by fusion of either the two HR regions (HR1, linker and HR2) or linking the 6-helix [3 × (HR1, linker and HR2)] together as a single chain. Subsequently, both of them form a stable 6-helix bundle in vitro judging by gel filtration and chemical cross-linking and the proteins show salient features of an α-helix structure. Crystals diffracting X-rays have been obtained from both protein preparations and the structure determination is under way. This method could be used for crystallization of the post-fusion state HR structures of other viruses. Keywords: construct design/fusion/heptad repeat/6-helix/Newcastle disease virus

Introduction

It is emerging that conformational changes of a given protein are important for the complex biological functions from whole genome sequencing of a number of organisms, including humans, based on the fact that the actual numbers of proteins (open reading frame) are far fewer than we ever expected. Interaction of the virus envelope proteins and the host cell receptors is the key step for enveloped-virus fusion in the process of virus entry in viral infections. Conformational changes of the viral envelope proteins are believed, for some viruses, if not all, to play a critical role in this entry/fusion process [review (Eckert and Kim, 2001)]. It is striking that a number of viruses contain highly conserved heptad repeat (HR) regions in their fusion proteins, in that the HR regions form α-helical structures (called HR1 or HR-A or HR-N or N-peptide and HR2 or HR-B or HR-C or C-peptide) (Bentz, 2000; Skelhel and Wiley, 2000; Eckert and Kim, 2001). The fusion peptide, responsible for insertion of the envelope protein into the host cell membrane, is located immediately before the first HR region. As a number of viral envelope protein structures, especially the ‘trimer of 2HR’ or so-called ‘6-helix bundle’, are now available, a consensus fusion mechanism of conformational changes has been proposed for some enveloped viruses, such as retrovirus, orthomyxovirus, paramyxovirus and filovirus (Bentz, 2000; Skelhel and Wiley, 2000; Eckert and Kim, 2001). In the current model, at least three conformations of the fusion proteins are involved in this process, i.e. pre-fusion native state, intermediate pre-hairpin state and post-fusion state (Bentz, 2000; Skelhel and Wiley, 2000; Eckert and Kim, 2001). In the post-fusion state, the two HR sequences form an anti-parallel 6-helix coiled-coil structure (3 × 2HR) (Lu et al., 1995; Eckert and Kim, 2001). In this structure, the ‘trimer of 2HR’ or so-called ‘6-helix bundle’ is formed by three HR1 in the center, surrounded by three HR2 owing to the strong interaction between HR1 and HR2, and is therefore called a trimer of HR1 and HR2 heterodimer.

Fusion and entry of paramyxoviruses require two glycoproteins, fusion (F) protein and haemagglutinin-neuraminidase (HN) protein (HN or homologue H and G) (Lamb, 1993), although some of them in the family do not absolutely require HN or its homologues (Horvath et al., 1992; Bagai and Lamb, 1995; Karron et al., 1997). Newcastle disease virus (NDV) is a member of the genus Rubulavirus in the family Paramyxoviridae, which is composed of enveloped negative-stranded RNA viruses (Morrison and Portner, 1991). It was recently proposed that NDV and related avian paramyxoviruses should be assigned as a new genus, Avulavirus, as they have important enough differences compared with the genus Rubulavirus (Chang et al., 2001). Like all the members in the family, the NDV F protein is synthesized initially as a precursor, F0, and then cleaved into F1 and F2 by a furin-like enzyme of the host cells (Homma and Ouchi, 1973; Scheid and Choppin, 1974). During the attachment and subsequent fusion process, the F protein undergoes conformational changes that expose the fusion peptide in F1 and result in the fusion peptide embedding in the host cell membrane. In the F1 of NDV, the HR1 region is located at the N-terminus, immediately followed by the fusion peptide, while the HR2 region is located adjacent to the N-terminus of the transmembrane domain. Previously, we have shown that separately expressed HR1 and HR2 of the NDV F protein assembled into a 6-helix coiled-coil bundle in vitro (Yu et al., 2002), although the actual structure is not known. Recently, the crystal structure of the pre-fusion native state of NDV F protein has been solved (Chen et al., 2001), although there is a lack of some important domains; however, it
represents the second such complex structure after the influenza HA (Chen et al., 1998). Therefore, there is a great need for the crystal structure and biophysical analysis of the NDV post-fusion 6-helix coiled-coil bundle for the comparison of conformational changes, which could give a clear idea of the molecular mechanism of paramyxovirus fusion in the process of virus entry.

Here we have deployed a GST fusion Escherichia coli expression system to express the HR1 and HR2 as a GST fusion protein of single-chain of 2-helix (HR1±2) or 6-helix (HR1±2±1±2±1±2). Both the expressed, GST-removed proteins form a stable 6-helix complex, representing the post-fusion state, and X-ray diffracting crystals have been obtained. This methodology would facilitate the future crystal structure analysis of the post-fusion 6-helix bundle of this group of viruses.

Materials and methods
Cloning and expression of the recombinant constructs
The NDV F gene was cloned from the Chinese virulent isolate F48E9 (GeneBank no. AF 079172). The HR1 region used was derived from amino acids 115–152 and HR2 from 442–477 (see Figure 1). The regions chosen were based on the computer program CoilsScan (Lupas et al., 1991) and the known crystal structure of SV5 HR1/2 (Baker et al., 1999). Two different constructs were made by PCR, 2-helix and 6-helix, as shown in Figure 1. The 2-helix construct was made by linking the HR1 and HR2 with a six amino acid linker (SGGRGG; single amino acid abbreviations used here). The 6-helix construct was generated by sequentially linking the HR1 and HR2 using two different linkers (linker 1 as SGGRGG or linker 2 as GGSGG) (Figure 1B). All the constructs were cloned into the BamHI and XhoI sites (introduced by synthetic PCR primers) of GST fusion expression vector pGEX-6p-1 (Pharmacia), in which there is a rhinovirus 3C protease cleavage site for the fusion protein (the same as the commercial PreScission protease cleavage site). The positive plasmids were verified by direct DNA sequencing.

Protein expression and purification
Escherichia coli strain BL21(DE3) transformed with the recombinant pGEX-6p-1 plasmid was grown at 37°C in 2× YTA medium to an optical density of 0.8–1.0 (OD 590 nm) before induction with 1 mM IPTG for 4 h at 37°C. Bacterial cells were harvested and lysed by sonication in phosphate-buffered saline (PBS, 10 mM sodium phosphate, pH 7.3; 150 mM NaCl). TritonX-100 was then added to a final concentration of 1% and the lysate was incubated for 30 min at 0°C and was subsequently clarified by centrifugation at 12 000 g for
30 min at 4 °C. The clarified supernatants were passed through a glutathione–Sepharose 4B column (equalized by PBS). The GST fusion protein-bound column was washed with PBS over 10 column volumes and eluted with reduced glutathione (5 mM) for three column volumes. The GST fusion proteins were then cleaved by GST-fusion rhinovirus 3C protease (kindly provided by Drs K. Hudson and J. Heath) at 5 °C for 16 h in the cleavage buffer (50 mM Tris–HCl, pH 7.0; 150 mM NaCl; 1 mM DTT; 1 mM EDTA, pH 8.0). The free GST and the GST-3C protease were removed by passage through the glutathione–Sepharose 4B column again. The resultant protein 2-helix and 6-helix were dialyzed against PBS and concentrated to a suitable concentration by ultrafiltration and stored at −70 °C for further analysis. Proteins were analyzed by Tris–tricine SDS–PAGE or 12% SDS–PAGE. Protein concentrations were determined by the BCA protein determination assay (Pierce Biochemicals).

**Gel-filtration analysis**

The GST column-purified 2-helix and 6-helix proteins were separately loaded on a HiLoad Superdex G75 column with an Akta FPLC system (Pharmacia). The fractions of the peak were collected and subjected to Tris–tricine SDS–PAGE or 12% SDS–PAGE. Protein concentrations were determined by the BCA protein determination assay (Pierce Biochemicals).

**Chemical cross-linking**

The gel filtration-purified 2-helix protein was dialyzed against cross-linking buffer (50 mM HEPES, pH 8.3; 100 mM NaCl) and concentrated to ~2 mg/ml by ultrafiltration (10K cut-off). Proteins were cross-linked with ethylene glycol bis(succinimidyl succinate) (EGS) (Sigma). The reaction mixtures were incubated for 1 h on ice at concentrations of 0, 0.1, 0.5, 2.0, 4.0 and 6.0 mM EGS and stopped with 50 mM glycine. The cross-linked products were analyzed under reducing conditions by 14% SDS–PAGE.

**Circular dichroism (CD) spectroscopy**

CD spectra were measured on a Jasco J-715 spectrophotometer in PBS (10 mM sodium phosphate, pH 7.3; 150 mM NaCl). Wavelength spectra were recorded at 25 °C using a 0.1 cm pathlength cuvette. Thermodynamic stability was measured at 222 nm by recording the CD signals in the temperature range 25–95 °C with a scan rate of 1 °C/min.

**Protein crystallization**

Proteins (in 20 mM Tris–HCl, pH 8.0) were concentrated to 10 mg/ml. The initial crystallization conditions were screened by using the hanging-drop vapor diffusion method with sparse matrix crystallization kits (Crystal Screen I and II; Hampton Research, Riverside, CA). The 2-helix protein crystals were obtained in at least four conditions and all the conditions were optimized in order to obtain the best diffracting crystals. The 6-helix crystals were also obtained under one set of conditions. The X-ray diffraction data for 2-helix crystals were collected and analyzed.

**Results**

**GST-2-helix and GST-6-helix were both expressed as soluble proteins**

The constructs of 2-helix and 6-helix are shown schematically in Figure 1 and they were both expressed as soluble GST-fusion proteins in *E. coli*, despite the high hydrophobicity of the HR1 amino acids. In the construct design, to avoid the insolubility problem that we encountered earlier (Yu et al., 2002), we chose shorter HR1 regions that lack some amino acids near the fusion peptide. Those amino acids would not be covered by HR2 based on the crystal structure of SV5 HR1/2 (Baker et al., 1999). It was
confirmed from this result that this construct design is important to achieve highly soluble protein. The GST affinity-purified proteins from the supernatants of cell lysates were subsequently cleaved by using the GST-3C rhinovirus protease (kindly provided by Drs K.Hudson and J.Heath) as described in Materials and methods (Figure 2).
After the removal of the fusion partner GST, the cleaved proteins of both 2-helix and 6-helix are soluble at high concentrations (up to 30–40 mg/ml) in PBS or Tris buffer. Both 2-helix and 6-helix protein preparations form trimers of HR1–HR2

The purified 2-helix and 6-helix proteins were analyzed by gel filtration and chemical cross-linking (2-helix only) for estimation of the molecular weight as described in Materials and methods. In the gel filtration, the purified proteins of both 2-helix and 6-helix gave rise to a peak with a molecular weight of 26 kDa (Figure 3), close to the theoretical values for the trimer of HR1–HR2. In chemical cross-linking of the 2-helix preparations (Figure 4), both dimers and trimers could be seen on the reducing gel. Monomers/dimers tended to fade as the cross-linker EGS concentration was increased, whereas the trimer bands became thicker. All of this indicated a trimer state of the 2-helix, although the monomer/dimer bands did not disappear even at high concentrations of the cross-linker. It also showed that the 6-helix protein preparation ran in a similar position to the 2-helix trimer (Figure 3).

**CD spectra**

To investigate the intrinsic α-helix nature of the 6-helix bundle, the secondary structure of the 2-helix and 6-helix proteins was examined by CD spectroscopy as described in Materials and methods. The resulting spectra at 25°C (Figure 5A and C) show that both 2-helix and 6-helix protein preparations were stably folded and formed typical α-helix structures, with obvious double minima at 208 and 222 nm. Regarding the thermal stability (Figure 5B and D), both 2-helix proteins show a typical α-helix secondary structure (A and C) and thermal stability (B and D), although the 6-helix protein seems slightly more stable than the 2-helix trimer. Thermal denaturation curves were recorded at 222 nm with a scan rate of 1°C/min.

**Fig. 4.** Chemical cross-linking of 2-helix protein. Cross-linked products were separated on 14% SDS-PAGE followed by Coomassie Brilliant Blue staining. Protein markers are shown in kDa. The numbers 0, 0.1, 0.5, 2.0, 4.0 and 6.0 indicate the concentration (mM) of ethylene glycol bis(succinimidylsuccinate) (EGS) used. The 6-helix protein was run as a control. Bands corresponding to monomer, dimer and trimer are indicated.

**Fig. 5.** CD spectra at 25°C and thermal stability measurement of 2-helix and 6-helix proteins in PBS. Both 2-helix and 6-helix show a typical α-helix secondary structure (A and C) and thermal stability (B and D), although the 6-helix protein seems slightly more stable than the 2-helix trimer. Thermal denaturation curves were recorded at 222 nm with a scan rate of 1°C/min.
and 6-helix constructs showed the formation of a very stable complex. The melting temperatures were very high, 89±90°C for 2-helix (Figure 5B) and 90±91°C for 6-helix (Figure 5D).

Protein crystallization

Both 2-helix and 6-helix protein preparations were crystallizable under several conditions (Figure 6), again indicating stable complex formation of the 6-helix bundle. X-ray diffraction data from two crystals were collected up to 2.5 Å resolution and the structure determination is under way. Similarly, we have successfully obtained 2-helix crystals with other paramyxoviruses, e.g. measles virus, Menangle virus and Nipah virus.

Discussion

Great efforts have been made to elucidate the molecular mechanism of paramyxovirus fusion in the process of virus entry; however, the true nature of this process has been poorly revealed so far. In comparison with other enveloped viruses, the conformational change of the viral envelope proteins involved in virus fusion of entry is a key step for some of the enveloped viruses, e.g. retrovirus, filovirus and orthomyxovirus [review (Eckert and Kim, 2001)]. For paramyxovirus, there are two envelope proteins: fusion protein F and hemagglutination–neuraminidase (HN) (or homologue H, G, depending on the genus in the Paramyxoviridae family) (Lamb, 1993). F protein is believed to be directly responsible for the fusion between the cellular membrane and the virus membrane, with or without the help of HN protein (Horvath et al., 1992; Bagai and Lamb, 1995; Karron et al., 1997). In the current model (Eckert and Kim, 2001), there are at least three conformational states of the F protein, i.e. pre-fusion native state, pre-hairpin intermediate state and post-fusion state. In the post-fusion state, the two conserved heptad repeat regions bind each other in an anti-parallel way as a coiled-coil structure and form the so-called 6-helix trimer (Lu et al., 1995; Eckert and Kim, 2001). In this structure, the 6-helix bundle coiled-coil is centered by three HR1 coils and surrounded by three HR2 coils. The two HR regions are highly conserved in all paramyxoviruses sequenced to date. Two 6-helix HR coiled-coil bundle structures have been seen so far at the atomic level for paramyxovirus, Simian virus 5 (SV5) (Baker et al., 1999) and human respiratory syncytial virus (hRSV) (Zhao et al., 2000). They basically confirmed the 6-helix coiled-coil bundle trimer formed by HR1 and HR2 and imply that the molecular fusion model proposed for other enveloped viruses (Eckert and Kim, 2001) most likely would be suitable for the paramyxovirus. Recently, the three-dimensional structure of NDV F protein has been solved (Chen et al., 2001) and it shows that the HR1 region forms an α-helix trimer, although the HR2 region is not seen in the structure. In comparison with the only available crystal structure of the native form of fusion proteins (HA) from influenza virus, the respective HR1 region should form an α-helix trimer in the native form and the HR1 and HR2 regions should fold together to form a 6-helix bundle trimer in the post-fusion state (Wilson et al., 1981; Bullough et al., 1994; Chen et al., 1998). Therefore, seeing a crystal structure of the HR 6-helix bundle of NDV F protein in comparison with its available native crystal structure (Chen et al., 2001) would give a second example of such conformational changes for this group of viruses that adopt a similar fusion/entry mechanism. Our data and those of others (Young et al., 1997, 1999; San Roman et al., 2002; Yu et al., 2002) have shown that HR1 and HR2 from NDV would form a stable 6-helix bundle structure in vitro. It is noteworthy that the orientations of the central triple-stranded coiled coil in the NDV F protein structure and influenza HA structure are in opposite directions. Therefore, the actual fusion mechanisms between paramyxovirus and orthomyxovirus might be slightly different although they might adopt a similar one overall. In this sense, structural analysis of

Fig. 6. Different crystals of 2-helix trimer and 6-helix. Crystals of 2-helix trimer were grown in (A) 0.1 M sodium acetate, pH 4.6, 0.2 M calcium chloride, 5–15% 2-propanol, (B) 0.1 M Tris–HCl, pH 8.5, 10–15% PEG 4000, (C) 0.1 M sodium cacodylate, pH 6.5, 0.2 M calcium acetate, 10–15% PEG 8000 and (D) 35% dioxane. 6-Helix crystals were grown in 0.1 M sodium acetate, pH 4.6, 1.0–2.0 M sodium chloride (E).
the F protein of NDV in different conformations would shed light on the molecular mechanism of enveloped virus entry.

Previously, we have used the E.coli system to express separately the F protein HR1 and HR2 of the NDV, subsequently assembling the purified HR1 and HR2 into a 6-helix bundle (Yu et al., 2002). However, we failed to obtain crystals of good quality for structural analysis. In this study, we designed two constructs by linking together the heptad repeat regions of the NDV F protein using some flexible amino acids. By using a series of biochemical and biophysical measures, our results show that both 2-helix and 6-helix proteins form soluble 6-helix trimers and are crystallizable using this method. We have since succeeded in the crystallization of 6-helix bundle LR trimers derived from some other virus fusion proteins, including those of measles virus, Menangle virus and Nipah virus. This indicates that this is a reliable method for obtaining crystallizable proteins for crystal structure analysis of such a group of proteins and will inevitably help in the study of the molecular mechanism of virus fusion with the accumulation of more structural data.

Acknowledgements

We thank K.Hudson and J.Heath for the rhinovirus 3C protease GST fusion construct. 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).

References


Received August 13, 2002; revised January 20, 2003; accepted March 13, 2003