

## Biochemical and biophysical characterization of the transmissible gastroenteritis coronavirus fusion core

Guangpeng Ma<sup>a,1</sup>, Youjun Feng<sup>b,c,1</sup>, Feng Gao<sup>b,d</sup>, Jinzi Wang<sup>d</sup>,  
Cheng Liu<sup>d</sup>, Yijing Li<sup>a,\*</sup>

<sup>a</sup> Department of Preventive Veterinary, College of Veterinary Medicine, Northeast Agriculture University, 150030 Harbin, PR China

<sup>b</sup> Laboratory of Molecular Immunology and Molecular Virology, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100080, PR China

<sup>c</sup> Graduate School of the Chinese Academy of Sciences, PR China

<sup>d</sup> China Agricultural University, Beijing 100094, PR China

Received 29 September 2005

Available online 10 October 2005

### Abstract

Transmissible gastroenteritis coronavirus (TGEV) is one of the most destructive agents, responsible for the enteric infections that are lethal for suckling piglets, causing enormous economic loss to the porcine fostering industry every year. Although it has been known that TGEV spiker protein is essential for the viral entry for many years, the detail knowledge of the TGEV fusion protein core is still very limited. Here, we report that TGEV fusion core (HR1-SGGRGG-HR2), in vitro expressed in GST prokaryotic expression system, shares the typical properties of the trimer of coiled-coil heterodimer (six  $\alpha$ -helix bundle), which has been confirmed by a combined series of biochemical and biophysical evidences including size exclusion chromatography (gel-filtration), chemical crossing, and circular diagram. The 3D homologous structure model presents its most likely structure, extremely similar to those of the coronaviruses documented. Taken together, TGEV spiker protein belongs to the class I fusion protein, characterized by the existence of two heptad-repeat (HR) regions, HR1 and HR2, and the present knowledge about the truncated TGEV fusion protein core may facilitate in the design of the small molecule or polypeptide drugs targeting the membrane fusion between TGEV and its host.

© 2005 Elsevier Inc. All rights reserved.

**Keywords:** Transmissible gastroenteritis coronavirus; Fusion core; Heptad-repeat regions

Porcine transmissible gastroenteritis (TGE) is an acute, highly prevalent enteric infectious disease, which is associated with high morbidity in animals of all ages and with high mortality in suckling piglets, causing extremely enormous economic loss to piglet cultivation in the world every year [1,2]. Transmissible gastroenteritis virus of swine (TGEV), a member of group I of coronaviruses, has been identified to be the key causative agent responsible for the TGE [1,3]. To date, the only solution to prevent the disease only is to inoculate ordinary vaccines, which are either inactivated or tissue culture ones with lot of disadvantages,

such as high cost, low effectiveness, etc. [2,3]. Since none of the effective vaccines or drugs could be available, it seems to be very urgent to develop some genetic engineering vaccines or drugs specific for TGE [2].

Coronaviruses are enveloped, positive-strand RNA viruses with largest genomes among the RNA viruses and are characterized by 3–4 enveloped proteins which embedded on the surface [3,4]. Both the receptor binding and the subsequent membrane fusion process of coronavirus are mediated by the spiker glycoprotein (S protein) [3–5]. It is generally believed that enveloped virus might adopt a similar molecular apparatus of virus membrane fusion in which two types have been proposed [3]. In type I, human immunodeficiency virus (HIV) [6–8], influenza virus [9], Ebola virus [10], and human respiratory syncytial virus

\* Corresponding author. Fax: +86 0451 5113336.

E-mail address: [yijingli@163.com](mailto:yijingli@163.com) (Y. Li).

<sup>1</sup> These authors contributed equally to this work.

(hRSV) [11,12] may be several typical ones of those examples with deeply structural insights, while in type II, flavivirus is an example lacking much knowledge in molecular mechanism of virus membrane fusion [3].

The fusion with the host cell membrane is the crucial step in the life cycle of all enveloped viruses, because it is necessary to facilitate the intracellular deposition of the viral genome followed by its replication [3,5]. It has been known that the envelope protein undergoes a series of conformational changes during the virus fusion process [3–6,9–11]. Two highly conserved heptad-repeat regions (HR1 and HR2) of S protein function as important modules/domains in this process and show different conformations in different fusion states [3,6,9]. Under the model given above, there are at least three conformational states of the envelope fusion protein, which include pre-fusion native state, pre-hairpin intermediate state, and post-hairpin state [3,9,10]. During these state transitions, the HR1 and HR2 are exposed to an intermediate conformational state but bind to each other to form the coiled-coil structure in an anti-parallel mode in the post-fusion stage. This coiled-coil bundle conformation is proposed to be important for bringing two lipid membranes (cellular and viral) into proximity with each other allowing the membrane fusion for viral entry into host cells [6,9,10].

The HR1/HR2 coiled-coil bundle is so called the virus fusion core [11]. In this structure, as presented by several crystal structures of fusion cores, including HIV [6–8], influenza virus A [9], Ebola virus [10], and hRSV [11,12], three HR1 bind to each other to form a trimeric core whereas three HR2 surround this core. As both HR1 and HR2 are structurally  $\alpha$ -helical in the fusion core, the structure is also called 6-helix coiled-coil bundle [3].

Recently, two research papers have separately presented the crystal structure of the fusion protein cores of the severe acute respiratory syndrome coronavirus (SARS-CoV) [16,17] and murine coronavirus (mouse hepatitis virus, MHV) [18], which both belongs to the Family *coronaviridae*. The two above 3D structures drew the following picture: a 6-helix bundle with three HR2 helices packed against the hydrophobic grooves on the surface of central coiled-coil formed by three parallel HR1 helices in an oblique anti-parallel manner, indicating that both SARS-CoV and MHV adopt the so-called type I virus membrane fusion mechanism. Moreover, the soluble HR2 derived from SARS-CoV and MHV are demonstrated to possess the inhibitory activities for viral fusion, extremely similar to the peptide inhibitor for HIV, Enfuvirtide or T20 [4,8,13].

To our knowledge, none of any experimental evidences have been presented to support that the TGEV fusion core shares the same features as those of well-known coronaviruses, although it is a member of group I of Family *coronaviridae* differing from both SARS-CoV and MHV [3,5]. In this study, we intended to investigate the structural basis of TGEV fusion through providing the biochemical and biophysical traits of its fusion core, especially the possibility of the molecular apparatus of TGEV fusion applied in fusion

inhibitor design for the treatment of TGE. Here, we deployed bio-engineering technique to design and prepare the protein of the TGEV fusion core (denoted as 2-Helix). The results of gel-filtration combined with circular dichroism (CD), chemical cross-linking, indicated that it is of trimer of heterodimer, coiled-coil bundle, implying that TGEV may adopt type I membrane fusion mechanism furthermore, the 3D structure model of TGEV fusion core clearly represented its most likely stereo configuration extremely similar to those of coronaviruses including HIV, etc. [6–12,17,18]. In conclusion, the presented knowledge about the truncated TGEV fusion protein core will facilitate to design the small molecules or polypeptide drugs targeting the crucial step of TGEV membrane fusion, similar to T20 specific for HIV which has been successfully applied for the treatment of HIV infection.

## Materials and methods

*Prediction of the heptad-repeat regions and construction of the TGEV fusion core.* The porcine transmissible gastroenteritis virus (TGEV) spiker gene used in this work was cloned from the Chinese isolate TH-98 (GenBank Accession No. AF494337). As shown in Fig. 1, the TGEV S protein is a typical type I membrane protein. The HR1 and HR2 regions were predicted by using the computer software of LearnCoil-VMF, freely available from the website (<http://night-ingale.lcs.mit.edu/cgi-bin/vmf>) [19]. The predicted HR1 region covers amino acids 1045–1184, whereas the relevant HR2 includes the amino acids 1339–1378 (Fig. 1). Considering the feasibility of the soluble expression of the fusion core generated in this experiment, the HR1 and HR2 regions of TGEV were adequately truncated and extended, respectively, on the basis of the multiple alignment of TGEV with the other coronaviruses in the conserved regions. Finally, the TGEV fusion core (2-Helix construct) was made by linking the modified HR1 (1057–1119) and HR2 (1326–1383) with a flexible linker (SGGRGG, single amino acid abbreviation used here), and then was inserted directionally into the prokaryotic expression vector pGEX-6P-1 (Pharmacia) via the restriction sites *Bam*HI and *Xho*I (introduced by PCR primers). The acquired recombinant expression plasmid, which harbored the interested DNA fragment of the TGEV fusion core was verified by direct DNA sequencing.

*Protein expression and purification.* The candidate positive recombinant clones were transformed into *Escherichia coli* strain BL21 (DE3) competent cells and the single colony was inoculated into Luria–Bertani (LB) medium containing 50 mg/L ampicillin (Sigma, USA) at 37 °C for overnight. Following, the overnight culture was transferred into the fresh LB medium for large-scale protein production at 37 °C. When the culture density ( $OD_{600}$ ) added up to 0.6–0.8, the culture was induced with 0.15 mM isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG) (Sigma, USA) and grown for another ~12 h at 16 °C until the bacterial cells were harvested.

The harvested culture was centrifuged at 5000 rpm for 12 min at 4 °C, and the bacterial cell pellet was resuspended in the iced PBS (140 mM NaCl, 2.7 mM KCl, 10 mM  $Na_2HPO_4$ , and 1.8 mM  $KH_2PO_4$ ) and homogenized by sonication. The lysate was centrifuged at 18,600 rpm for 20 min at 4 °C and subsequently filtered through 0.22  $\mu$ m membrane for clarification. Then the supernatant was loaded onto a glutathione–Sephacrose 4B column (Pharmacia). When the protein-loaded column was then washed with six times of column volume of PBS, the GST-fusion protein was eluted by 20 mM reduced glutathione (Pharmacia). To obtain the GST-removed protein, the GST-3C rhinovirus protease (kindly provided by Drs. J. Heath and K. Hudson) was added into the resin and then the mixture was incubated with gentle agitation for about 10 h at 4 °C. The target protein was eluted with 10 ml PBS.

*Gel-filtration analysis.* The target protein (2-Helix) loaded on a Superdex 75 column (Pharmacia) with an Akta Purifier System

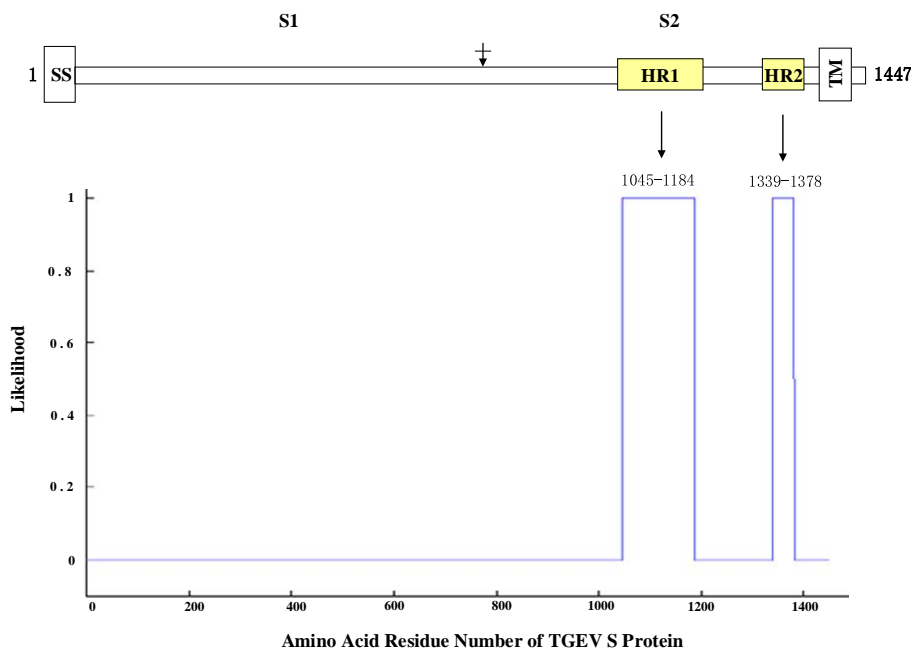


Fig. 1. Prediction of the HR regions of TGEV S protein. Schematic diagram of S protein (amino acids 1–1447 for the full length) is shown in the upper panel. The cleavage of S1 and S2 is indicated by a vertical arrow. SS, signal peptide; HR1 and HR2, heptad-repeat regions 1, 2; TM, transmembrane region. In the lower panel, the likelihood of HR1 and HR2 predicted by LearnCoil-VMF program [19] is represented.

(Pharmacia) after it was concentrated by ultra-filtration (10 kDa cut-off) and exchanged from PBS buffer into the exclusion buffer. The fraction of the peak was collected and analyzed by a 17% SDS-PAGE, and the molecular weight of the interested peak was estimated by comparison with the GST protein run on the same gel.

**Circular dichroism spectroscopic analysis.** Circular dichroism (CD) spectra were performed on a Jasco J-715 spectrophotometer in PBS. Wavelength spectra were recorded at 25 °C using a 0.1 cm path-length cuvette. Thermodynamic stability was measured at 222 nm by recording the CD signals at the temperature which varied from 25 to 90 °C with a scan rate of 5 °C/min.

**Chemical cross-linking of the fusion core.** The purified 2-Helix protein after the gel-filtration was dialyzed against cross-linking buffer (50 mM Herpes, pH 8.3; 100 mM NaCl) and concentrated to approximately 5 mg/L by ultra-filtration (10 kDa cut-off). The resultant proteins were subjected to chemical cross-linking reaction with ethylene glycol bis-succinimidylsuccinate (EGS) (Pierce). The reactions were incubated for 1 h at room temperature at different concentrations of EGS, respectively (0.0, 0.2, 0.5, 1.0, 1.5, and 2.0 mM EGS), and quenched with 50 mM glycine. Eventually, the cross-linked samples were analyzed by 17% SDS-PAGE.

**3D structural model building of the fusion core.** The deduced amino acid sequence of the TGEV fusion core was sent into the CPHmodels 2.0 Server [23] and then was processed. Finally, the acquired coordinates were used to generate the 3D structure of the TGEV fusion core with the aid of the program of DeepView/SwissPdb-Viewer 3.7 (SPS).

## Results and discussion

### Design of the TGEV fusion core

Based on the prediction of the LearnCoil-VMF program (Fig. 1) and the multiple alignments of the putative heptad-repeat regions of TGEV with those of several other coronaviruses (Fig. 2), the final version of the TGEV fusion core used in this study has been determined. The fusion core consisted of HR1, a truncated type of the Learn-

Coil-VMF predicted HR1, connected with HR2, an extended version of the LearnCoil-VMF predicted HR2, by a flexible linker (SGGRGG). The fusion core designed here shared the feature of the helix wheels, a typical characteristic of class I fusion protein (Fig. 3).

### Soluble expression of the TGEV fusion core

The TGEV fusion core was synthesized by using overlapping PCR and then cloned into the prokaryotic expression vector, pGEX-6P-1, through the restriction sites (*Bam*HI and *Xho*I) introduced by PCR. Ideally, the fusion core should be 127 aa in length with an ideal molecular mass of about 13.0 kDa. Luckily, the GST fused fusion core protein in the soluble form was observed in the supernatant of the bacteria lysate, and the GST removed fusion core protein run on the 17% SDS-PAGE showed at the position of the expected size (Fig. 4). The availability of much soluble TGEV fusion core protein made it possible to perform the subsequent experiments to classify and characterize the TGEV fusion core.

### Characteristic 6-helix bundle formed by the TGEV fusion core

The purified TGEV fusion core proteins (2-Helix) were concentrated to 10–20 mg/ml in the size exclusion buffer and analyzed by gel-filtration and chemical cross-linking for estimation of the molecular weight. The 2-Helix protein was eluted at the volume of ~10 ml which followed the position of GST dimer (52 kDa) presented by the Superdex 75 Column (Fig. 5). In comparison, the

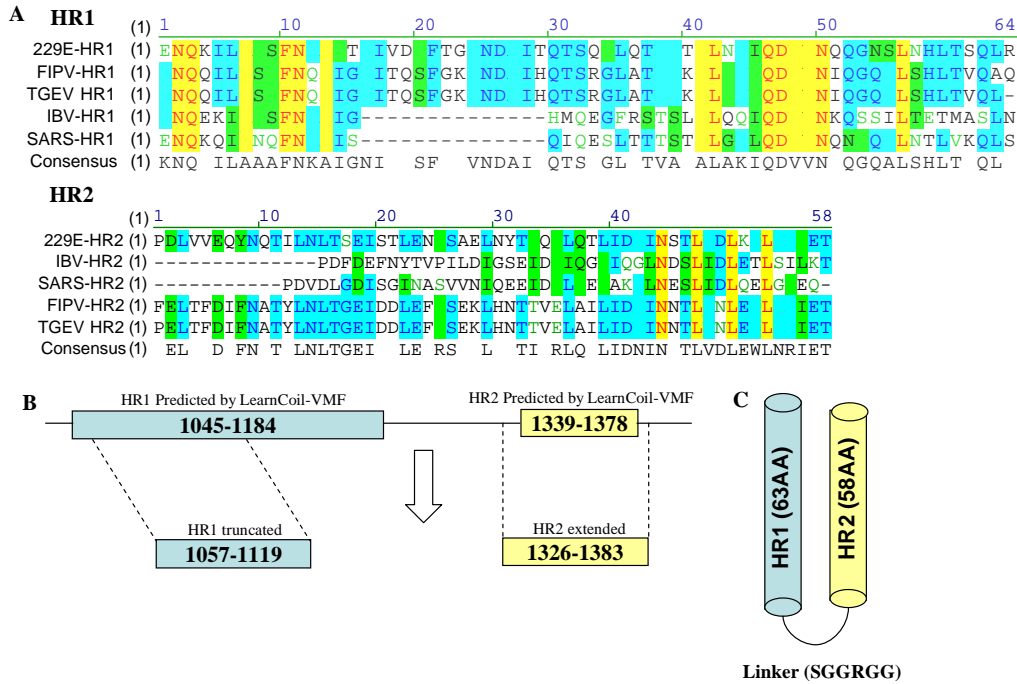


Fig. 2. Development strategy for the TGEV fusion core. (A) Multiple alignments of the partial sequences of the coronavirus heptad-repeat regions (HR) in amino acid level. (B) The modification of the final version of the HR1 and HR2 of TGEV applied in this study. (C) Schematic representation of the TGEV fusion core (2-Helix).

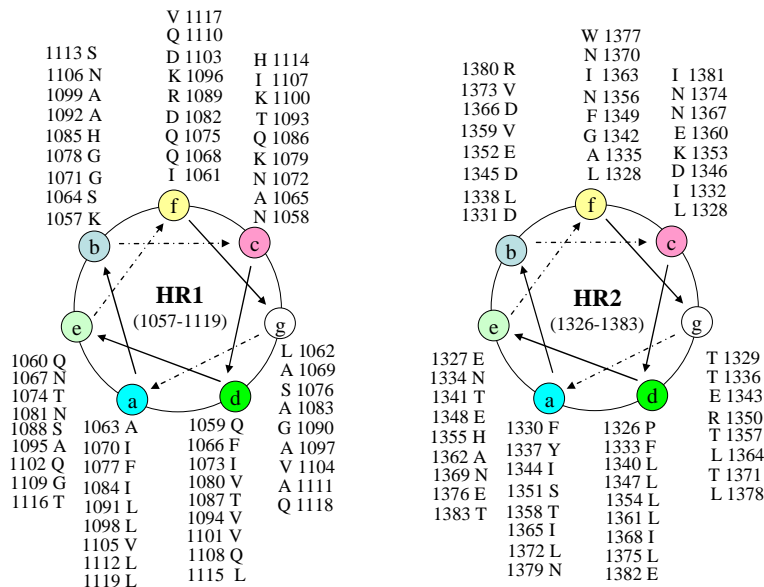


Fig. 3. The helical-wheel representation of the final type of HR regions of TGEV spiker protein.

computed molecular mass of the 2-Helix protein was about 13.0 kDa, and then it indicated that the 2-Helix might form oligomers (~40 kDa). Subsequently, the chemical cross-linking experiment demonstrated the 2-Helix protein oligomer to be a trimer (Fig. 7), and at the same time, that the transitional states (monomer and dimer) could be observed clearly. In addition, the content of the trimer increased with the concentration increase of the chemical cross-linker (EGS).

CD spectroscopic profile of the fusion core (2-Helix) presented an absorption curve of the typical  $\alpha$ -helix structure, with double minima at 208 and 222 nm (Fig. 6A), which was completely consistent with the previously published data of some other virus fusion cores [4–6,11,15,21]. Moreover, the thermodynamic measurement of the fusion core protein indicated that it could keep its advance structure up to above 85 °C (Fig. 6B), suggesting that the 2-Helix formed trimer represents the core structure

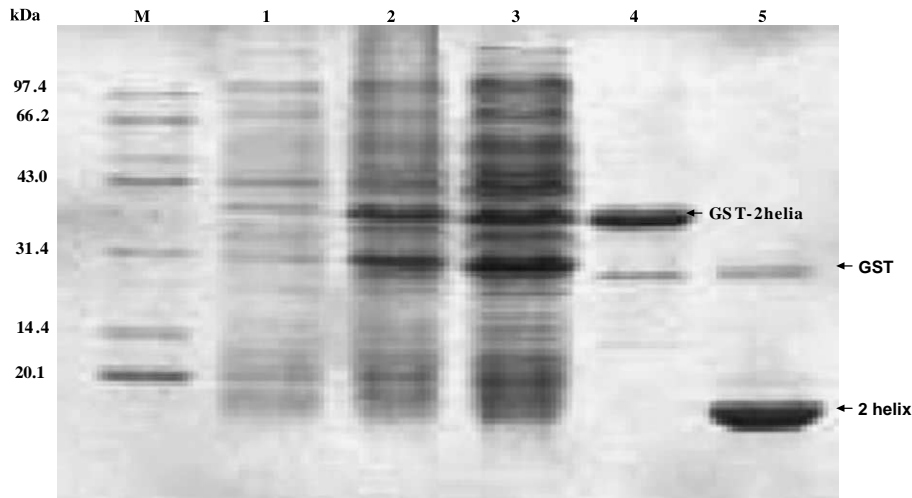


Fig. 4. SDS-PAGE analysis of the expression and purification of the TGEV fusion core (2-Helix). Lane M, molecular weight standards in kDa; lane 1, non-induced; lane 2, supernatant of the IPTG-induced bacteria BL21(DE3); lane 3, pellet of the IPTG-induced bacteria BL21(DE3); lane 4, the GST fused 2-Helix protein purified by GST beads; lane 5, GST-removed 2-Helix protein.

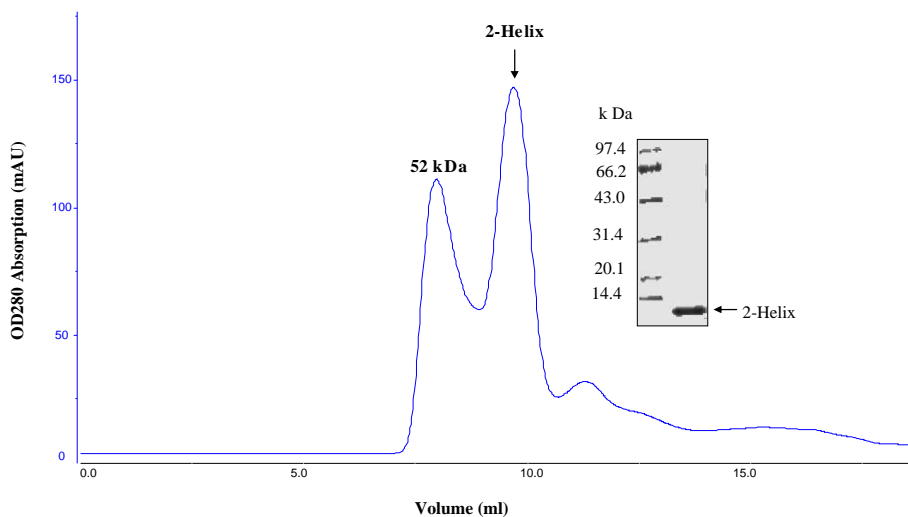


Fig. 5. Gel-filtration analysis of the purified TGEV fusion core (2-Helix). The first peak was formed by the GST dimer (52 kDa), and the inset picture is 17% SDS-PAGE analysis of the protein collected from the second peak. The profile of the TGEV fusion core shows clearly it exists in a complex of about 40 kDa, implying it possibly forms a trimer.

of the post-fusion state of the TGEV coiled-coil bundle, which is extraordinarily stable.

Finally, 3D structural model building of the TGEV fusion core was conducted with the program of Swiss-Pdb-Viewer 3.7 (SPS) and showed clearly the typical characteristic of the 2-Helix in molecular level (Fig. 8).

In summary, the TGEV fusion core actually formed a 6-helix bundle, a trimer of heterodimer, implying that TGEV may adopt the type I membrane fusion mechanism.

#### Potential implication for antiviral strategies

Similar to human coronavirus 229E (hCoV 229E), TGEV has also been identified as a member of group I in Family *coronaviridae*, which do not include SARS-CoV,

MHV, etc. [3,5]. Importantly, it has been confirmed to be a severe pathogen responsible for the porcine transmissible gastroenteritis (TGE), an acute and highly epidemic enteric infectious disease, which is associated with high morbidity in animals of all ages and with high mortality in suckling piglets [1,2]. There is no doubt that it may result in an extremely enormous economic loss to piglet cultivation in the world every year, and thereby has attracted the attention of several virologists in the world to research the pathogenic mechanism of TGEV, and even interaction between TGEV and the host, porcine. In fact, so far, the only way used to prevent the disease is to inoculate ordinary vaccines, which are either inactivated or tissue culture ones with much disadvantages, such as high cost, low effectiveness, etc. Just due to lacking of effective vaccines or drugs,





entry into its host, perturbing the successful infection of virus. It will be of interest to test whether HR1 or HR2 has the inhibitory activity in the next work. Anyway, it did first provide us the biochemical and biophysical basis of the TGEV fusion core, pointing out a novel direction to design the polypeptide or small molecule drugs, perturbing TGEV fusion with its host membrane, for the prevention and therapeutics of TGE.

### Acknowledgments

We thank Professor George F. Gao because the work was completed in the laboratory of Molecular Immunology and Molecular Virology in the charge of Professor George F. Gao, and we are also grateful to Dr. Yueyong Liu for his technical assistances, Ms. Huimin Zhang for her devotion in preparing the manuscript.

### References

- [1] J.C. Yin, X.F. Ren, Y.J. Li, Molecular cloning and phylogenetic analysis of ORF7 region of chinese isolate TH-98 from transmissible gastroenteritis virus, *Virus Genes* 30 (2005) 395–401.
- [2] C. Schwegmann-Wessels, G. Zimmer, B. Schroder, G. Breves, G. Herrler, Binding of transmissible gastroenteritis coronavirus to brush border membrane sialoglycoproteins, *J. Virol.* (2003) 11846–11848.
- [3] X.J. Wang, W.H. Zhang, M. Wang, G.F. Gao, The molecular mechanism of the enveloped virus-cell membrane fusion, *Prog. Biochem. Biophys.* 31 (2004) 482–491.
- [4] J.Q. Zhu, G.F. Xiao, Y.H. Xu, F. Yang, C.Y. Zheng, Y.Y. Liu, H.M. Yan, D.K. Cole, J.I. Bell, Z.H. Rao, P. Tien, G.F. Gao, Following the rule: formation of the 6-helix bundle of the fusion core from severe acute respiratory syndrome coronavirus spike protein and identification of potent peptide inhibitors, *Biochem. Biophys. Res. Commun.* 319 (2004) 283–288.
- [5] B.J. Bosch, R. van der Zee, C.A. de Haan, P.J. Rottier, The coronavirus spiker protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex, *J. Virol.* 77 (2003) 8801–8811.
- [6] D.C. Chan, D. Fass, J.M. Berger, P.S. Kim, Core structure of gp41 from the HIV envelope glycoprotein, *Cell* 89 (1997) 263–273.
- [7] D.C. Chan, P.S. Kim, HIV entry and its inhibition, *Cell* 93 (1998) 681–684.
- [8] M.J. Root, M.S. Kay, P.S. Kim, Protein design of an HIV-1 entry inhibitor, *Science* 291 (2001) 884–888.
- [9] P.A. Bullough, F.M. Hughson, J.J. Skehel, D.C. Willey, Structure of influenza haemagglutinin at the pH of membrane fusion, *Nature* 371 (1994) 37–43.
- [10] W. Weissenhorn, A. Carfi, K.H. Lee, J.J. Skehel, D.C. Willey, Crystal structure of the Ebola virus membrane fusion subunit, GP2, from the envelope glycoprotein ectodomain, *Mol. Cell* 2 (1998) 605–611.
- [11] X. Zhao, M. Singh, V.N. Malashkevich, P.S. Kim, Structural characterization of the human respiratory syncytial virus fusion protein core, *Proc. Natl. Acad. Sci. USA* 97 (2000) 14172–14177.
- [12] E. Wang, X. Sun, Y. Qian, L. Zhao, P. Tien, G.F. Gao, Both heptad repeats of human respiratory syncytial virus fusion protein are potent inhibitors of viral fusion, *Biochem. Biophys. Res. Commun.* 302 (2003) 469–475.
- [13] C. Cianci, D.R. Dischino, Y.X. Sun, K.L. Yu, A. Stanley, J. Roach, Z.F. Li, R. Dalterio, R. Colonno, N.A. Meanwell, M. Krystal, Targeting a binding pocket within the trimer-of-hairpins: small-molecule inhibition of viral fusion, *Proc. Natl. Acad. Sci. USA* 101 (2004) 15046–15051.
- [14] G.F. Gao, Filling the hole: evidence of a small molecule binding to the fusion core pocket in human respiratory syncytial virus, *Expert Opin. Investig. Drugs* 14 (2005) 195–197.
- [15] Y.Y. Xu, D.K. Cole, Z.Y. Lou, Y.W. Liu, L. Qin, X. Li, Z.H. Bai, F. Yuan, Z.H. Rao, G.F. Gao, Construct design, biophysical, and biochemical characterization of the fusion core from mouse hepatitis virus (a coronavirus) spike protein, *Protein Expr. Purif.* 38 (2004) 116–122.
- [16] Y.Y. Xu, J.Q. Zhu, Z.Y. Lou, Y.Y. Liu, D.K. Cole, L. Ni, N. Su, L. Qin, X. Li, J.I. Bell, H. Pang, P. Tien, G.F. Gao, Z.H. Rao, Characterization of the heptad repeat regions, HR1 and HR2, and design of a fusion core structure model of the spike protein from severe acute respiratory syndrome (SARS) Coronavirus, *Biochemistry* 43 (2004) 14064–14071.
- [17] Y.Y. Xu, Z.Y. Lou, Y.W. Liu, H. Pang, P. Tien, G.F. Gao, Z.H. Rao, Crystal structure of severe acute respiratory syndrome coronavirus spike protein fusion core, *J. Biol. Chem.* 279 (2004) 49414–49419.
- [18] X. Xu, Y. Liu, Z.Y. Lou, L. Qin, X. Li, Z. Bai, H. Pang, P. Tien, G.F. Gao, Z.H. Rao, Structure basis for coronavirus mediated membrane fusion: crystal structure of MHV spider protein fusion core, *J. Biol. Chem.* 279 (2004) 30514–30522.
- [19] M. Singh, B. Berger, P.S. Kim, Learn Coil-VMF: computational evidence for coiled-coil-like motifs in many viral membrane-fusion proteins, *J. Mol. Biol.* 290 (1999) 1031–1041.
- [20] T.F. Wild, R. Buckland, Inhibition of measles virus infection and fusion with peptides corresponding to the leucine zipper region of the fusion protein, *J. Gen. Virol.* 78 (1997) 107–111.
- [21] M. Yu, E. Wang, Y. Liu, D. Gao, N. Jin, C.W.-H. Zhang, M. Bartlam, Z.H. Rao, P. Tien, G.F. Gao, Six-helix bundle assembly and characterization of heptad repeat regions from the F protein of Newcastle disease virus, *J. Gen. Virol.* 83 (2002) 623–629.
- [22] X.H. Xu, S. Gao, D.K. Cole, J.J. Zhu, N. Su, H. Wang, G.F. Gao, Z.H. Rao, Basis for fusion inhibition by peptides: analysis of the heptad repeat regions of the fusion proteins from Nipah and Hendra virus, newly emergent zoonotic paramyxovirus, *Biochem. Biophys. Res. Commun.* 315 (2004) 664–670.
- [23] Lund, M. Nielsen, C. Lundegaard, P. Worning, X3M a Computer Program to Extract 3D Models, Abstract at the CASP5 conference A102, 2002.