INTRODUCTION
Classical swine fever (CSF), also known as hog cholera is a long-standing disease, being described as early as 1833, in Ohio USA (1). CSF continues be of importance presently (2). The disease causes major damage to the pig populations in the Chinese provinces and in other countries worldwide (3). CSF is a highly contagious disease that mainly spreads by contact between live pigs or by feeding pigs with contaminated pig meat. Though many countries (e.g. European Union member states) pursue a non-vaccination eradication policy, massive vaccination with attenuated vaccines, such as hog cholera lapinized vaccine (HCLV) (also known as C-strain), developed in China in mid-1950s has been implemented routinely as a major control strategy in China as well as many other developing countries. The characteristic of CSF is the marked immuno-suppression induced on infected animals, including the depletion of B-lymphocytes and T-lymphocytes (4). Both the domestic pig and wild boar can become infected and CSF continues to cause serious problems in various parts of the world (5).

CSF is caused by classical swine fever virus (CSFV), a member of the genus Pestivirus within the family Flaviviridae. Other members of this group include bovine viral diarrhea virus (BVDV) and border disease virus (BDV) (6). The genus Pestivirus in the family Flaviviridae includes economically important, widespread pathogens causing postnatal and prenatal infections in a wide range of host animals, such as cattle, swine, and sheep (7). The CSFV virion is particle hexagonally shaped, which has an electron-dense inner core structure of about 30 nm, surrounded by a spherical envelope with diameters ranging between 40 and 60 nm (7). The genome of CSFV comprises a single open reading frame (ORF), approximately 12.3 kb in length (1, 8). This ORF, flanked by a 5’-nontranslated region (NTR) and a 3’-NTR, encodes a polyprotein composed of about 3898 amino acids, which is processed by viral and cellular enzymes into four structural (C, Erns, E1 and E2) and eight non-structural
(Npro, p7, NS2, NS3, NS4A, NS4B, NS5A and NS5B) proteins (1, 2, 9) (Fig 1). Phylogenetic analysis indicates CSFV mainly fall into one of two groups: group 1 comprising mainly the modified live vaccines and of highly virulent strains and group 2 mainly comprising of recent moderately virulent isolates (2).

To date, much progress on the CSFV molecular characterization have been achieved, but the genome structure and its protein function remains to be elucidated. We present here a review CSFV molecular characterization and potential future directions for research in this area.

**STRUCTURAL PROTEINS**

**C protein**

The core protein C is a small protein which contains many basic amino acids (lysine and arginine) (1, 8, 10). The formation of a core protein–RNA complex inside the virion suggests a protective function of the core protein (11). C protein can activate the promoter of heat shock protein 70 gene, and suppress the SV40 early promoter. These findings indicate that the C protein appears to function not only as a viral structural protein but also as a regulator of gene expression (12).

**E<sup>ns</sup> protein**

E<sup>ns</sup> (E0) glycoprotein was formally termed gp44 with an apparent size of 41-44 kDa and composed of 227 amino acids. Approximately half of the molecular mass of the mature E<sup>ns</sup> glycoprotein is made up of carbohydrates (13). Unlike E1 and E2 proteins, E<sup>ns</sup> glycoprotein lacks the membrane anchor and is secreted from infected cells (14), but it is associated with mature virions and has been reported to be involved in CSFV entry into the cell (15). E<sup>ns</sup> may take part in the initial attachment process of viral entry, rather than in the specific binding or fusion process (13). E<sup>ns</sup> may also play an important role in the post-entry stages, which may be a possible reason causing the CSFV with E<sup>ns</sup> deletion to be non-transmissible (16, 17). E2 is the major neutralizing antigen for CSFV infection, while E<sup>ns</sup> is considered to be the secondary glycoprotein that mediates neutralization (18). Both E<sup>ns</sup> and E2 are known to induce viral neutralizing antibodies which give protective immunity in the natural host (19). E<sup>ns</sup> glycoprotein represents a second determinant for the induction of protective immunity against CSF. It is reported that animals vaccinated with recombinant vaccinia virus expressing E<sup>ns</sup> developed neutralizing antibodies against CSFV (13, 20). Compared to other viruses in the family *Flaviviridae*, the N-terminal of the E<sup>ns</sup> is unique to *Pestiviruses*, and has been implicated in the evasion of host interferon (IFN) responses (20, 21). An important feature is that E<sup>ns</sup> is heavily glycosylated with N-linked glycans and representing up to half of the apparent molecular weight (22). It contains ribonuclease activity and its C-terminal domain controls translocation across eukaryotic cell membranes and has a critical role in the inhibition of double-stranded RNA-induced cell responses. N-glycan of CSFV E<sup>ns</sup> is also essential for E<sup>ns</sup> blocking of IFN-β induction (23, 24, 25). CSFV E<sup>ns</sup> can bind to exogenous double strand RNA (dsRNA) and inhibit dsRNA-induced IFN-β production, however it failed to inhibit TRIF (Toll/IL-1 receptor domain-containing adapter inducing IFN-β) but instead triggered IFN-β production. Deglycosylation of E<sup>ns</sup> rendered it unable to bind to dsRNA, and thus unable to inhibit dsRNA-induced IFN-β production (26, 27).

**E1 protein**

Glycosylation is one of the most common types of protein modifications, and N-linked oligosaccharides are added to specific asparagines residues in the context of the consensus sequence Asn-X-Ser/Thr (28). E1, with its molecular weight 33KDa and containing 195 amino acids, is the smallest envelope and type I trans-membrane proteins with an N-terminal ectodomain and a C-terminal hydrophobic anchor, which has been implicated in viral adsorption to host cells (18, 21). E1 has been shown to contain three putative N-linked glycosylation sites. Glycosylation sites (at...
CSFV amino acid residue position N500, N513 and N594) are highly conserved among CSFV isolates and two of them (N513 and N594) also conserved among other Pestiviruses (21). Removal of all three putative glycosylation sites in E1, at CSFV positions N500, N513 and N594, yielded non-viable progeny, while single or dual site mutants excluding N594 were viable. Virus glycoproteins are crucial to the virus cycle such as attachment to host cell receptors, entry, assembly of newly produced viral progeny, and exit (29). In vivo, viral glycoproteins have been shown to influence infectivity, virulence, and host immune responses. Modifications of glycosylation sites in envelope proteins E1 yielded viruses with a variable degree of attenuation (21). Modification of E1 glycosylation patterns could be used for developing CSFV live-attenuated vaccines (3).

**E2 protein**

Envelope protein E2 (373 amino acids, 51-55KDa) is the major envelope glycoprotein exposed on the outer surface of the virion and represents an important target for induction of the immune responses during infection (30). The protein contains four antigenic domains (A, B, C, and D), which are located within the N-terminal half of the protein. E2 forms a heterodimer with E1 as well as homodimers, and it may play a major role in virus attachment and entry (18). On the basis of amino acid sequence analysis and determination of the N-terminal, it was shown that processing of E2 is mediated by a host cell signalase. Typical hydrophobic signal sequences were identified upstream of the E2 N-terminus within E1-coding sequences. A transmembrane anchor of about 40 hydrophobic amino acids was demonstrated at the C-terminus of E2. The E2 glycoprotein contains sequential neutralizing epitopes, which are responsible for eliciting neutralizing antibodies which confer protective immunity and is therefore frequently used for designing DNA vaccines against CSF (31, 32).

**NON-STRUCTURAL PROTEINS**

**Npro protein**

Npro (168 amino acids, 23KDa) is the first non-structural protein encoded in the ORF protein. Npro (N-terminal auto-proteinase) exerts two known functions: an auto-protease activity for co-translational cleavage from the nascent downstream nucleocapsid protein C and as an antagonistic effect on the IFN-α/β induction pathway (22, 24, 33). Npro is a cysteine proteinase, that has similarities to subtilisin-like proteinases and which is not found in other viral systems (27). Mutants lacking the Npro gene, as opposed to wild-type CSFV, induced type I interferon in cell culture, which suggests a function in the interference with the primary cellular antiviral defense (34). Previous studies indicated that the N-terminal proteinase Npro is required for virulence of CSFV. Replacement of the Npro gene of CSFV by the murine ubiquitin gene only slightly affects the characteristics of virus replication in the porcine kidney cell line SK-6. Mutant vA187-Npro-Ubi obtained by replacement of the Npro gene by the murine ubiquitin gene in the CSFV genome, did not induce any clinical symptoms in experimentally infected pigs (25). Npro deletion mutants are attenuated and induce protective immunity in SPF (Specific pathogen free) pigs after single oronasal inoculation (35).

**P7 protein**

Downstream of E2, the small hydrophobic protein p7 (70 amino acids, 7KDa), flanked by signal peptidase cleavage sites, is essential for infectious virus production, but is not associated with virus particles (36). In cell extracts, p7 is partially associated with E2 protein. Separation of the E2 and p7 genes with an internal ribosome entry site (IRES) resulted in a viable virus, demonstrating that the E2-p7 precursor protein is not required for virus replication (10, 37). Generation of p7 requires microsomal membranes, but p7 is not a major structural component of the virion. p7 forms the junction between the structural and the non-structural genes in pestiviruses, and the same is probably true for hepatitis C virus. The nature of p7 and E2-p7 and the peculiar processing of the E2-p7-NS2 region are well conserved among pestiviruses (10).

**NS2 protein**

Except for the leader protein Npro, all the mature processing products and the precursor protein NS2-3 are essential for the viral life cycle, while the proteins NS3 to NS5B are sufficient for RNA replication (37, 38). NS2 is an auto-protease associated with a cellular chaperone termed JIV (J-domain protein interacting with viral protein) and cleaves the NS2-3 protein between NS2 and NS3 (39). NS2-mediated cleavage at the NS2-3 site is essential for the Pestivirus life cycle and
is temporally modulated by a 1:1 stoichiometric association of NS2 with the cellular cofactor JIV (40). Consumption of the intracellular JIV pool that remains tightly associated with NS2, and the restriction of NS2 to cis-cleavage limits the NS2 activity to shortly after infection, resulting in the virtual absence of NS2-3 cleavage at later times post-infection and, consequently, in the reduction of viral RNA replication, which is crucial for the virus to persist (37, 41). Nevertheless, NS2 might exert other functions not required for the viral life cycle. NS2 was shown to inhibit gene expression from different cellular promoters as well as to interfere with cell proliferation (40, 41).

**NS3 protein**

Non-structural protein NS3 is a multi-functional protein involved in polyprotein processing (42). Previous reports showed that NS3 protein can induce apoptosis in host cells that present with cytopathic effects (CPE) (41). NS3 is believed to possess three enzyme activities that are likely to be essential for virus replication: a serine protease located in the N-terminus of one-third of the protein and RNA-stimulated nucleotide triphosphatase (NTPase) as well as RNA helicase activities located in the C-terminal portion (43, 44). The presence of these activities, initially inferred from the identification of amino acid motifs and alignments with related proteins, has subsequently been demonstrated for several members of the family Flaviviridae. Circumstantial evidence suggested that the protease and helicase/NTPase domains of the NS3 protein were functionally interdependent (45). Bioinformatics analysis showed that the CSFV NS3 protein contains canonical amino acid motifs present in all super-family II RNA helicases (43). The NTPase/helicase activities of NS3, together with the NS5B (RNA-dependent RNA polymerase, RdRp), should be essential for viral replication (42). Wen et al. reported that NS3 helicase activity was dependent on the presence of NTP and divalent cations, with a preference for ATP and Mn²⁺, and required the substrates possessing a 3’-un-base-paired region on the RNA template strand (46). NS3 protease domain enhanced the NS3 helicase activity but had no effect on its NTPase activity. For the truncated NS3 helicase domain (NS3h), both NTPase and helicase activities were up-regulated by NS5B (47). The NS3h specifically interacted with the plus- and minus-strand 3’-NTR. Characterization of the NS3 NTPase activity was confirmed and it was shown that both the full length and truncated NS3 enhanced IRES-mediated and cellular translation (42, 48).

**NS4A and NS4B proteins**

NS4A associates with the N-terminus of NS3 and acts as a cofactor for the serine protease activity (37). A role in cytopathogenicity was associated with NS4B (49). NS4A is an indispensable cofactor of uncleaved NS2-3 in the formation of infectious particles (50). Importantly, NS4A can be supplied in trans form separately from NS2-3 without loss of efficiency (37). NS4A is important for the correct conformation, topology and functionality of NS3 within the infected cell (42). Another function of NS4A could be to recruit other viral or cellular proteins, as for example NS4B and NS5A for which an interaction with NS4A has been reported (51).

**NS 5A and NS5B proteins**

Cleavage between NS5A and NS5B is slow, accordingly, the precursor protein NS5A-5B can be detected in cell extracts and is processed in NS5A and NS5B with a half-life of less than 60 minutes (6). NS5A is phosphorylated and was shown to be the only protein of the replication complex that can be complemented in trans form (37). NS5A was also found to interact with a subunit of the translation elongation factor 1A (TEF 1A), suggesting a role of this protein in genome replication and translation (52). NS5B is located at the extreme C-terminus of the polyprotein and binds to the 3’-NTR of the viral RNA and initiates its replication. Binding to this part of the genome will cause the degradation of the entire viral RNA, likely reducing viral replication (53). NS5B contains motifs shared by RNA-dependent RNA polymerase (RdRp), such as the Gly-Asp-Asp (GAA) motif, which is highly conserved among RdRps and has been demonstrated to possess RdRp activity (54). RdRp is a key enzyme which initiates RNA replication by a de novo mechanism without a primer and is a potential target for anti-virus therapy (55). A three-dimensional model has been built by homology modeling based on the alignment with several virus RdRps and the results indicated that the fingers domain contains an N-terminal region that plays an important role in conformational changes (56).
NON-TRANSLATED REGION (NTR)

The 3’-NTR of plus-strand RNA of CSFV genome is believed to be the first entry site for viral replicases for initiation of RNA genome replication. The helicase activity is needed to unwind the secondary structure present in the 3’-NTR for the replication. The 3’-NTR is most likely involved in initiation of the pestiviral genome replication, also involved in the coordination of the viral translation and replication. The 3’-NTR is an important site for interaction between proteins and viral RNA genome (57). 3’-NTR mutants with deletion of 3’-terminal sequences is unable to be bound to NS5B proteins and even inhibits RNA synthesis by viral replicase (58). The evidence suggested that 3’-terminal sequence of 3’-NTR is essential to replicase and helicase activities (56). The 3’-terminal sequence of 3’-NTR might be also the first interactive site between helicase and the viral genome (44, 59). The intact secondary structure of 3’-terminal sequence of 3’-NTR is important in helicase binding, which has been demonstrated to be necessary for replicase binding and RNA synthesis. Moreover, the 3’-NTR may contact the 5’-NTR by RNA-RNA interactions (42). 3’-NTR also is the site for the interaction between NS5B replicase and viral genome. The 5’-NTR contains the IRES for cap-independent translation of the viral polyprotein. The 3’-NTR may contain replication signals involved in minus-strand RNA synthesis (54). The 5’-NTR contains cis-elements required for replication (59) and an IRES for cap-independent translation initiation of the viral polyprotein which is co- and post- translationally processed by host cell and viral proteases, whereas the 3’-NTR contains signals involved in the RNA replication process (60).

CONCLUSION

Much progress in CSFV molecular characterization has been made, which will undoubtedly accelerate understanding CSFV genome and protein structure. These attempts also provide many feasible strategies and methods for the development the new vaccines against CSF.

ACKNOWLEDGEMENT

The study was supported by the Technology Research Foundation of Education Department of HeiLongJiang Province, China. Fund No. (12511352)


