An intermolecular salt bridge linking substrate binding and P1 substrate specificity switch of arterivirus 3C-like proteases

Qian Chen\textsuperscript{a,b}, Junwei Zhou\textsuperscript{a,b}, Zhixiang Yang\textsuperscript{a,b}, Jiahui Guo\textsuperscript{a,b}, Zimin Liu\textsuperscript{a,b}, Xinyi Sun\textsuperscript{a,b}, Qingshi Jiang\textsuperscript{a,b}, Liurong Fang\textsuperscript{a,b}, Dang Wang\textsuperscript{a,b},\textsuperscript{a}\textsuperscript{a}\textsuperscript{a}\textsuperscript{a}\textsuperscript{b}, Shaobo Xiao\textsuperscript{a,b}

\textsuperscript{a}State Key Laboratory of Agricultural Microbiology, College of Veterinary Medicine, Huazhong Agricultural University, Wuhan 430070, China
\textsuperscript{b}Key Laboratory of Preventive Veterinary Medicine in Hubei Province, The Cooperative Innovation Center for Sustainable Pig Production, Wuhan 430070, China

1. Introduction

The family \textit{Arteriviridae\textsuperscript{a}}, which is a member of the order \textit{Nidovirales}, includes six subfamilies \textit{Crocarterivirinae\textsuperscript{a}, Eqarterivirinae\textsuperscript{a}, Heroarterivirinae\textsuperscript{a}, Simarterivirinae\textsuperscript{a}, Variarterivirinae\textsuperscript{a} and Zealarterivirinae\textsuperscript{a} \cite{1}. Some arteriviruses are known as important veterinary pathogens while others infect certain species of wild rodents or African non-human primates and can only be characterized only by genome sequencing \cite{2,3}. Of these arteriviruses, porcine reproductive and respiratory syndrome virus (PRRSV, belongs to the subfamily \textit{Variarterivirinae\textsuperscript{a}}) and equine arteritis virus (EAV, which belongs to the subfamily \textit{Eqarterivirinae\textsuperscript{a}}) are important etiological agents in the area of veterinary research because of the economic losses that they cause in the horse- and swine-breeding industries worldwide \cite{4–6}. Arterivirus infections are both cytological and cause outbreaks of either acute or persistent infections in their hosts \cite{7}. Although most EAV infections are subclinical, some virulent strains periodically cause significant diseases outbreaks and may be associated with abortions, neonatal mortality, and establishing persistent infection in stallions \cite{8,9}. Pigs of all ages in immunologically naïve herds are vulnerable to being infected with PRRSV \cite{10}. Highly virulent PRRSV strains can cause severe clinical diseases in susceptible herds, which are economically important pathogens of pigs \cite{11–13}. Currently, available drugs and vaccines do not provide adequate protection against these diseases \cite{10,12,14}, thus it is important to find an appropriate strategy to prevent arterivirus from proliferating in the host.

Arteriviruses represent a family of positive-sense, single-stranded, enveloped RNA viruses whose viral RNA genomes range in length from about 12 to 16 kb \cite{15,16}. Arteriviruses represent a...
family of positive-sense, single-stranded, enveloped RNA viruses whose viral RNA genomes range in length from about 12 to 16 kb [7,17]. Three-fourths of the 5′-proximal length of the genome is taken up by two large open reading frames (ORFs), ORF1a and ORF1b, that encode two long nonstructural replicase polyproteins, pp1a and pp1ab [18,19]. Both polyproteins are autoproteolytically processed by various papain-like cysteine proteases and a 3C-like serine protease (3CLpro) at conserved sites to generate 13–14 mature nonstructural proteins (nsps) for gene transcription and replication [20–24].

Arterivirus 3CLpro plays an indispensable role in the life cycle of arteriviruses that cleaves the polyproteins at nine conserved sites [20,21,25]. Considering its essential function in viral replication, arterivirus 3CLpro is a highly prospective target for anti-arterivirus drug design [26–29]. In addition to viral polyprotein processing, we and others demonstrated that arterivirus 3CLpro, specifically EAV 3CLpro and PRRSV 3CLpro, cleave host NF-κB essential modulator (NEMO) at multiple sites, thereby impairing NF-κB activation and IFN-β production. Importantly, cleavage of NEMO by PRRSV 3CLpro or EAV 3CLpro led to suppression of the type I IFN signaling pathway [30,31]. Therefore, targeting 3CLpro would hinder viral replication and could also boost host immune response [32].

Many 3CLpro inhibitors for COVID-19 treatment have been proposed recently and some new drug candidates have been successful in preclinical studies [33]. Understanding the substrate recognition mechanism of 3CLpro is a key step in establishing substrate-specific inhibitors of 3CLpro. Meanwhile, substrate envelope-guided design has achieved better potency and resistance profile of 3CLpro inhibitors. The 3CLpro families of arteriviruses are highly variable among themselves and are reasonably conserved only in the active site region. Like most chymotrypsin-like serine proteases, PRRSV 3CLpro employs a canonical Ser-His-Asp triad responsible for recognizing cleavage sites with a glutamic acid (Glu, E) at the P1 position to target viral polyprotein [34,35]. Nevertheless, EAV 3CLpro uses the same catalytic triad but recognizes both Glu and glutamine (Gln, Q) at the P1 position [20,36]. Similar substrate preferences were obtained when arterivirus 3CLpro cleaved host protein, NEMO. Our previous study demonstrated that both EAV 3CLpro and PRRSV 3CLpro-induced NEMO cleavage at the conserved residues E166, E171, and E349 at the P1 position, whereas only EAV 3CLpro could target NEMO at Q205 [30]. The substrate preferences of PRRSV 3CLpro and EAV 3CLpro cleaved host proteins might be consistent with the 3CLpro preferences for viral substrates. However, why the cleavage of NEMO at site Q205 is unique to EAV 3CLpro has not been investigated in previous studies. In this study, we characterized the key amino acids in the S1 subsite of arterivirus 3CLpro for P1 substrate specificity by molecular dynamics (MD) simulations and site-directed mutagenesis. The major differences between the S1 subsites manifest at PRRSV 3CLpro Lys138 residue, where EAV 3CLpro contains Thr139 residue. The K138T mutant of PRRSV 3CLpro converts the P1-Glu Substrate Specificity of this protease to that toward P1-Gln seen in the EAV 3CLpro. Moreover, Thr139 in EAV 3CLpro is responsible for the preference for P1-Gln, since the T139K mutant more efficiently recognizes substrates containing P1-Glu than the wild type (WT) EAV 3CLpro. Hence, we identified a potential intermolecular salt bridge linking substrate binding and P1 substrate specificity shift of arterivirus 3CLpro, which might be important when designing specific inhibitors of arterivirus 3CLpro.

2. Materials and methods

2.1. Sequence analysis of the PRRSV and EAV genomes

The complete PRRSV and EAV genomes were selected from the National Center for Biotechnology Information (NCBI, https://www.ncbi.nlm.nih.gov) for substrate profiling of their 3CLpro. Briefly, we obtained 1,037 complete PRRSV genome and 96 complete EAV genome (obtained 1 November 2021). Nine putative cleavage sites of PRRSV 3CLpro and EAV 3CLpro were inferred from alignments of pp1ab polyprotein amongst PRRSV and EAV isolates. Details on the data set are summarized in Supporting Information Fig. S1.

2.2. Plasmids

The eukaryotic expression plasmids for PRRSV 3CLpro, EAV 3CLpro, and NEMO (E166A-E171A-E349A) were described previously [30]. PRRSV 3CLpro and EAV 3CLpro mutants were constructed by overlap extension PCR using specific mutagenic primers and cloned into the C-terminal hemagglutinin (HA) tag-encoding pCAGGS-HA-C plasmid. The luciferase-based biosensor plasmid 233D, which contained oligonucleotides corresponding to ENLFPQ1YS (cleaved by TEV 3Cpro), were used as the biosensor control as described previously [37]. To monitor the activity of PRRSV 3CLpro and EAV 3CLpro oligonucleotides corresponding to the amino acid sequences of substrate peptides were cloned into the 233D biosensor vector, including PRRSV-nsp3/4, EAV-nsp3/4, EAV-nsp10/11, PRRSV-nsp3/4-Q, and NEMO-Q205. All the constructs were to be validated by DNA sequencing.

2.3. Dual-luciferase assays

Human embryonic kidney (HEK293T) cells were obtained from the China Typical Culture Collection Center (Wuhan, China) and cultured at 37 °C, 5% CO2 in Dulbecco's modified Eagle medium (DMEM, Invitrogen, Madison, WI, USA), with 10% fetal bovine serum. HEK293T cells in 24-well plates were transfected with indicated luciferase biosensor plasmids and pRL-TK (Promega), along with PRRSV 3CLpro or EAV 3CLpro expression plasmids or the empty control plasmid. The pRL-TK plasmid was used as a control for transfection efficiency. Cells were lysed 28 h post-transfection and firefly luciferase and Renilla luciferase activities were determined using the dual-luciferase reporter assays (Promega, USA). The data represent relative firefly luciferase activity normalized to Renilla luciferase activity and are representative of three independently conducted experiments. All results were presented as means ± standard deviations. The p value of < 0.05 was regarded as significant. All results were analyzed using GraphPad Prism Version 9.3 (www.graphpad.com).

2.4. Western blot

HEK293T cells grown in 6-well plates were transfected with indicated plasmids. After 28 h transfection, cell lysates were prepared in RIPA buffer (Beyotime, China), quantified for protein concentrations, and subject to SDS-PAGE and immunoblotting as previously described [30]. The following antibodies were used to identify the corresponding proteins: anti-β-actin antibody (Antigen, Wuhan, China), anti-Luciferase antibody (Promega), anti-HA antibody (MBL, Nagoya, Japan), and anti-Flag antibody (Macgene, China). Expression of β-actin was detected with an anti-β-actin antibody to confirm loading of equal protein amounts. To confirm the expression levels of HA-tagged WT arterivirus 3CLpro and its mutants, an anti-HA antibody was used for immunoblotting. The expression of luciferase or Flag-tagged NEMO proteins was analyzed using anti-Luciferase antibody or an anti-Flag antibody, respectively. Following incubation with appropriate secondary antibodies, protein bands were imaged with a ChemiDoc Imaging System (Bio-Rad, USA).
2.5. Molecular dynamics simulations

Molecular dynamics simulations were performed as described previously [37]. Briefly, the PRRSV 3CL\textsuperscript{pro} (PDB id: 5Y4L) and EAV 3CL\textsuperscript{pro} (PDB id: 1MBM) complexed with different substrate peptides (np3/np4 auto-cleavage sequence of PRRSV 3CL\textsuperscript{pro} LGSLLE/GAFRTQ, np3/np4 auto-cleavage sequence of EAV 3CL\textsuperscript{pro} GMVFE/GLFRSP, and ns10/ns11 cleavage sequence of EAV 3CL\textsuperscript{pro} CGWEKQ/SNKISC) were generated utilizing the structure of Glu-SPG in complex with a tetrapeptide (PDB id: 1HPG) as the reference template, because EAV 3CL\textsuperscript{pro} and PRRSV 3CL\textsuperscript{pro} exhibited a high similarity with Glu-SPG structure among all available crystal complex structure. For consistency, these crystal structures were also adopted as the template for the construction of the single mutant complexes using SYBYL-X v.2.0 (v.2.0; https://omictools.com/sybyl-x-tool). MD simulations were conducted using the Amber ff14SB force field (with the TIP3P water model) implemented in the GROMACS 2018 software package as formerly described [38–41]. The distance between the edge of the water box and protein was at least 1.2 nm. Molecular systems were neutralized by the addition of a certain number of counterions sodium (Na\textsuperscript{+}) or chloride (Cl\textsuperscript{−}), and NaCl at 150 mM was used to mimic physiological conditions [42]. For energy minimization and relaxation of the systems, each system was energy minimized by using a 2-step, extensive energy minimization process based on the steepest descent method followed by the conjugate gradient algorithm. Following minimization, each system was gradually heated from 0 to 310 K in 500 ps, then equilibrated at that temperature for another 500 ps. The standard temperature is kept constant at 310 K, then equilibration was performed under constant pressure for 150 ns with no position restrictions on protein. During the MD simulation, the SHAKE algorithm was applied to constrain all bonds involving hydrogen atoms [43]. The temperature was controlled using the modified Berendsen thermostat (V-rescale algorithm) with a collision frequency of 0.2 ps [44]. The pressure was maintained at 1 bar using the Parrinello-Rahman barostat with a compressibility of 4.5 × 10\textsuperscript{−5} bar\textsuperscript{−1} and a coupling constant of 2.0 ps. The non-bonded cutoff distances and van der Waals (vdW) interaction were 15 Å, and the Particle Mesh Ewald (PME) method was utilized to treat long-range electrostatic interactions [45]. A 2 fs step was applied to calculate the motion equations using the Leap-Frog integrator for equilibration steps [43]. The coordinates of the atoms were taken at every 10 ps and used for the final analysis. The backbone root mean square deviations (RMSD) values, root mean square fluctuation (RMSF) values, vDW interactions and electrostatic interactions were calculated using the GROMACS package. In addition, the RMSD values suggest that the different conformations remain stable during 150 ns simulation times (Fig. S2). The 3D-structure of proteins and peptides was visualized using PyMOL software Version 2.5 (Schrödinger, LLC) [46].

2.6. Hydrogen bond calculations

Hydrogen bonds (H-Bonds) were identified based on donor–acceptor distance and donor–acceptor angle (X–Y distance < 3.5 Å and X–H–Y angle < 40°), whereas only the pairs of polar atoms belonging to different hydrophobic groups were involved. The H-Bonds Plugin from Visual Molecular Dynamics was used to perform hydrogen bond analyses for each complex. H-bond occupancy indicated the percentage of indicated H-bonds that are maintained throughout the MD simulation. The data represented all H-bonds with occupancy greater than 10% between P1 residue of the substrate peptides and S1 subsite of WT arterivirus 3CL\textsuperscript{pro}.

2.7. Clustering of MD trajectories using principal component analysis

Principal component analysis (PCA) was used to determine the conformational changes in protein movement by calculation and diagonalization of the covariance matrix of the C-alpha (Cα) atoms. The calculated orthogonal vectors or eigenvectors with the highest eigenvalues are named principal components (PCs). PCA was performed using the Bio3d package (v2.4.1; http://thegrantlab.org/bio3d/) in R environment (v4.0.4; https://mirror.fcaglp.unlp.edu.ar/CRAN) [47,48]. Trajectories are obtained from dynamic simulation. The 100 ns MD trajectory consists of 10,000 frames after equilibration. One conformation out of ten is sampled. By sampling each conformation, a single long trajectory is obtained, and thus it represents 10% of the 10,000 frames MD trajectory. To determine the internal motions of a protein, alignment is necessary before building the C-matrix using Cartesian coordinates. Cα atoms were selected for analysis and in each frame were aligned by removing the overall translation and rotation, and then corresponding PCs were calculated under default parameters using the Bio3d package. The predominant movements during the simulation were plotted using these PCs.

3. Results

3.1. Different substrates preference at the P1 position between PRRSV 3CL\textsuperscript{pro} and EAV 3CL\textsuperscript{pro}

To clarify the substrate specificity of arterivirus 3CL\textsuperscript{pro}, we obtained 1,364 complete genome sequences of 25 different species of arteriviruses from the National Center for Biotechnology Information (NCBI). PRRSV and EAV, as two well-characterized arteriviruses, contributed most of the complete genome sequences of arteriviruses (1037 and 96, respectively). Thus, we selected PRRSV and EAV as representatives to investigate the substrates preference of arterivirus 3CL\textsuperscript{pro}. Based on previous studies [34,35], nine putative cleavage sites of PRRSV 3CL\textsuperscript{pro} and EAV 3CL\textsuperscript{pro} were predicted based on alignments of pp1ab polyproteins amongst all PRRSV and EAV isolates (Fig. 1a-b). The sequence logos of PRRSV 3CL\textsuperscript{pro} and EAV 3CL\textsuperscript{pro} substrates show considerable similarities, but there is a subtle difference at the P1 position. Consistent with previous studies, PRRSV 3CL\textsuperscript{pro} confers an extremely high specificity for the P1 residue of Glu in the substrates. Interestingly, while Glu is frequently found specific to the P1 position in EAV 3CL\textsuperscript{pro}, Gln at P1 appears in one of 9 cleavage sites (nsp10/nsp11) [21]. This cleavage site (nsp10/nsp11) with the preference of Gln at the P1 position is well conserved amongst all EAV isolates (Fig. S1). In addition, the P1′ position, on the other hand, is mainly occupied by various small-sized residues such as Gly, Ser and Ala, among which Gly is the most frequent one [21,22,34]. The residue preference at the remaining Pn’s residues of PRRSV 3CL\textsuperscript{pro} and EAV 3CL\textsuperscript{pro} is less demanding. These results revealed that the residue specificity of the P1 position varies in the substrates for PRRSV 3CL\textsuperscript{pro} and EAV 3CL\textsuperscript{pro}.

3.2. Structure comparison of S1 subsites in arterivirus 3CL\textsuperscript{pro} using molecular dynamics simulation

Viral 3CL\textsuperscript{pro} is broadly believed to undertake conformational changes after substrate binding. These changes allow them to interact with substrate proteins and to initiate catalytic activity [45]. To further determine whether structural changes in arterivirus 3CL\textsuperscript{pro} in response to substrate binding and hence potentially regulate arterivirus 3CL\textsuperscript{pro} activity, we performed molecular dynamics analysis of apo and substrate-bound arterivirus 3CL\textsuperscript{pro}.
S1 subsites of PRRSV 3CLpro and EAV 3CLpro showed significantly different amino acid residue compositions. In detail, the S1 subsite in the PRRSV 3CLpro crystallographic structure in our and others’ previous studies [49]. Surprisingly, such flexibility reduced substrate-bound PRRSV 3CLpro, suggesting that the presence of a substrate could stabilize the loops around the active site (Fig. 2d). B-factors analysis of all deposited arterivirus 3CLpro crystal structures completely verified these conclusions. It is worth adding that the residues on this loop differ greatly in 138 and 139 positions between PRRSV 3CLpro and EAV 3CLpro (Fig. S3), suggesting that these two amino acids in this loop are likely to be involved in the P1 substrate specificity switch of arterivirus 3CLpro.

To understand the essential motion changes of apo and substrate-bound conformations of arterivirus 3CLpro, PCA was performed on MD trajectories. For the apo and substrate-bound forms of PRRSV 3CLpro, the first two eigenvectors (EVs) contributed 33.66% and 56.6% of the global motion, respectively (Fig. S4). These results showed that the eigenvalues of the apo conformation were lower than those of the substrate-bound form. Furthermore, the scatter plots generated by the apo and substrate-bound forms of PRRSV 3CLpro indicated a significant difference between the two systems, which also suggested that the apo form of PRRSV 3CLpro showed less dynamic tertiary structure conformation than the substrate-bound form. Similar motion changes were obtained from MD trajectories of the apo and substrate-bound forms of EAV 3CLpro, suggesting that the structural rearrangement by substrate binding might contribute to the increased essential motion in the substrate-bound form of arterivirus 3CLpro.

3.3. Establishment of a biosensor assay for evaluating the protease – substrate specificity of arterivirus 3CLpro in vivo

The traditional approach to studying protease-substrate specificity is usually performed through in vitro proteolytic reactions with a purified protease and its predicted peptide or protein substrates in a fluid buffer. In vitro characterization is always labour- and time-demanding, and it might not accurately reflect the physiological characteristics of proteases. Here, we have established a fast, reliable, and efficient luciferase-based biosensor to quantify the substrate specificity of arterivirus 3CLpro in vivo, which has been used in previous studies to identify cleavage sites recognized by viral protease [37,50–52]. To further examine the flexibility of the S1 subsites in 3CLpro, we concentrated on the movements of loops enveloping the S1 subsites by RMSF analysis. We found that one of the analyzed loops of apo PRRSV 3CLpro, loop S136-G140, lined a wall of the S1 subsite and was more flexible than the corresponding loops of apo EAV 3CLpro structure, while the neighbouring loops had little flexibility (Fig. 2d). This might be deduced indirectly from the lack of the S136-G140 loop in the PRRSV 3CLpro crystallographic structure in our and others’ previous studies [49]. Surprisingly, such flexibility reduced substrate-bound PRRSV 3CLpro, suggesting that the presence of a substrate could stabilize the loops around the active site (Fig. 2d). B-factors analysis of all deposited arterivirus 3CLpro crystal structures completely verified these conclusions. It is worth adding that the residues on this loop differ greatly in 138 and 139 positions between PRRSV 3CLpro and EAV 3CLpro (Fig. S3), suggesting that these two amino acids in this loop are likely to be involved in the P1 substrate specificity switch of arterivirus 3CLpro.

To understand the essential motion changes of apo and substrate-bound conformations of arterivirus 3CLpro, PCA was performed on MD trajectories. For the apo and substrate-bound forms of PRRSV 3CLpro, the first two eigenvectors (EVs) contributed 33.66% and 56.6% of the global motion, respectively (Fig. S4). These results showed that the eigenvalues of the apo conformation were lower than those of the substrate-bound form. Furthermore, the scatter plots generated by the apo and substrate-bound forms of PRRSV 3CLpro indicated a significant difference between the two systems, which also suggested that the apo form of PRRSV 3CLpro showed less dynamic tertiary structure conformation than the substrate-bound form. Similar motion changes were obtained from MD trajectories of the apo and substrate-bound forms of EAV 3CLpro, suggesting that the structural rearrangement by substrate binding might contribute to the increased essential motion in the substrate-bound form of arterivirus 3CLpro.

As shown in Fig. 2a, the 3D structure of PRRSV 3CLpro is remarkably like that of EAV 3CLpro, with the corresponding RMSD of alignment of 1.653 Å, even though PRRSV 3CLpro has low amino acid sequence homology with EAV 3CLpro (about 37%). Both PRRSV 3CLpro and EAV 3CLpro consists of a chymotrypsin-like domains fold (domains I and II) at the N-terminus, an extra C-terminal domain III, and a lengthy loop linking domains II and III [23,34]. A shallow-surfaced gap between domains I and II, to which 6 and 7 antiparallel sheets contribute, forms the native substrate-binding site including the catalytic Ser-His-Asp triad and the Gly-Ser side-chain points away from the P1-Gln side-chain in EAV 3CLpro/Thr139 positions between PRRSV 3CLpro and EAV 3CLpro (Fig. S3), suggesting that the residues on this loop differ greatly in 138 and 139 positions between PRRSV 3CLpro and EAV 3CLpro (Fig. S3), suggesting that these two amino acids in this loop are likely to be involved in the P1 substrate specificity switch of arterivirus 3CLpro.

To understand the essential motion changes of apo and substrate-bound conformations of arterivirus 3CLpro, PCA was performed on MD trajectories. For the apo and substrate-bound forms of PRRSV 3CLpro, the first two eigenvectors (EVs) contributed 33.66% and 56.6% of the global motion, respectively (Fig. S4). These results showed that the eigenvalues of the apo conformation were lower than those of the substrate-bound form. Furthermore, the scatter plots generated by the apo and substrate-bound forms of PRRSV 3CLpro indicated a significant difference between the two systems, which also suggested that the apo form of PRRSV 3CLpro showed less dynamic tertiary structure conformation than the substrate-bound form. Similar motion changes were obtained from MD trajectories of the apo and substrate-bound forms of EAV 3CLpro, suggesting that the structural rearrangement by substrate binding might contribute to the increased essential motion in the substrate-bound form of arterivirus 3CLpro.
used as a negative control. In principle, arterivirus 3CL\textsuperscript{pro} effectively cleaves its favored substrates between residues P1 and P1’, leading to an interaction of the two firefly luciferase domains, resulting in an active form of the luciferase. Conversely, the unpreferred substrates will not be efficiently cleaved by arterivirus 3CL\textsuperscript{pro}, resulting in a non-active form of the luciferase. As shown...
in Fig. 3b, the luciferase activity of PRRSV-nsp3/4 was markedly induced in cells cotransfected with PRRSV 3CL<sup>pro</sup>, while no activity was detected in negative controls (233D). Moreover, the luciferase activity of both EAV-nsp3/4 and EAV-nsp10/11 was also induced in EAV 3CL<sup>pro</sup>-cotransfected cells. Interestingly, no activity of EAV-nsp10/11 was detected in PRRSV 3CL<sup>pro</sup>-cotransfected cells, suggesting that the substrate with P1-Gln was a nonpreferred substrate for PRRSV 3CL<sup>pro</sup>, which is consistent with previous studies on the P1 substrate specificity of EAV 3CL<sup>pro</sup> and PRRSV 3CL<sup>pro</sup> [21]. Western blotting also showed that PRRSV 3CL<sup>pro</sup>/EAV 3CL<sup>pro</sup> could split the recombinant firefly luciferase with the preferred cleavage sequence, resulting in a more rapidly migrating protein band (Fig. 3c). Consistency of cleavage and fold induction of the luciferase supports that there is a correspondence between the luciferase activity assay and cleavage of the biosensor construct by arterivirus 3CL<sup>pro</sup>, pointing to its possible use in evaluating the protease-substrate specificity of arterivirus 3CL<sup>pro</sup> in vitro.

3.4. Molecular determinants of arterivirus 3CL<sup>pro</sup> substrate specificity

Arterivirus 3CL<sup>pro</sup> employs a conserved Ser-His-Asp triad as the principal nucleophile and the general acid-base catalyst [23,34,53], which is important for substrate binding and hydrolysis. We considered the possibility that the additional amino acid residues of S1 subsites are also involved in arterivirus 3CL<sup>pro</sup> activity and P1 substrate preference. To this end, several amino acids in the S1 subsites of PRRSV 3CL<sup>pro</sup> and EAV 3CL<sup>pro</sup> were mutated to alanine and assessed the P1 substrate preference of all mutants using three cyclized luciferase-based biosensors (contained Glu or Gln at the P1 position to enable assessment of substrate preference at this position). As shown in Fig. 4a, the putative residues in S1 subsite of PRRSV 3CL<sup>pro</sup>, Thr113, Cys115, G116, His133, Ser136, and Lys138, were mutated to Ala and tested with a PRRSV-nsp3/4 biosensor (with P1-Glu). In contrast to WT PRRSV 3CL<sup>pro</sup>, the mutants (T113A, C115A, H133A, K138A) significantly less-efficiently cleaved substrate containing Glu at P1, but the mutants (G116A, S136A) almost completely abolished the ability of PRRSV 3CL<sup>pro</sup> to cleave substrates having P1-Glu. Using an EAV-nsp3/4 biosensor (with P1-Glu), we found that the mutants of EAV 3CL<sup>pro</sup> retained (T116A, S117A), significantly reduced (T115A, H134) or eliminated (G118A, S137A) the ability to cleave substrate containing Glu at P1 (Fig. 4b). These data suggest that both G116/H133/ S136 of PRRSV 3CL<sup>pro</sup> and G118/H134/S137 of EAV 3CL<sup>pro</sup> are important for recognition of P1-Glu. Interestingly, the mutants G118A and H134A of EAV 3CL<sup>pro</sup> lost P1-Gln cleavage activity, but the mutant S137A retained this ability cleaving P1-Gln substrate as did WT EAV 3CL<sup>pro</sup> (Fig. 4c), suggesting that although Ser137 is important for its ability to cleave substrates having P1-Glu, this residue is nonessential for recognition of P1-Gln.

To further investigate the underlying molecular mechanism of the alanine scan results, we performed MD simulations used by different PRRSV 3CL<sup>pro</sup> and EAV 3CL<sup>pro</sup> mutants with substrates. The RMSD values of the Ca atoms for each complex converged and remained stable during 150 ns MD simulations. In the last 100 ns of simulations, we calculated the selected H-bond (occupancy greater than 10%) during MD simulations to better capture the electrostatic interactions between the substrate and the protease. Gly116/Glu118 in PRRSV 3CL<sup>pro</sup>/EAV 3CL<sup>pro</sup> mainchain amides constitute the oxoanion hole by providing one H-bond to the cleavable P1-Glu and P1-Gln carbonyl oxygen that stabilizes the resulting negative charge during covalent catalysis. As expected, both G116A mutant in PRRSV 3CL<sup>pro</sup> and G118A mutant in EAV 3CL<sup>pro</sup> disrupted the engagement of H-bond network around the active sites (Fig. 4d-f), which also eliminated the ability to cleave the substrate (Fig. 4a-c). Moreover, His133/His134 and Thr113/Thr115 in PRRSV 3CL<sup>pro</sup>/EAV 3CL<sup>pro</sup> make H-bonds with the invariant side chain of P1-Glu and P1-Gln. Indeed, we found that alanine substitution of each of these two residues significantly reduced the H-bond occupancy between arterivirus 3CL<sup>pro</sup> and substrates. These conserved H-bond interactions between arterivirus 3CL<sup>pro</sup> and substrates provide an explanation for the reduced catalytic efficiency of the mutants (H133A/H134A, T113A/T115A) in PRRSV 3CL<sup>pro</sup>/EAV 3CL<sup>pro</sup> for both P1-Glu and P1-Gln substrates. By contrast, alanine mutation of other residues that cannot form stable H-bonds with the P1 substrate, such as C115 in PRRSV 3CL<sup>pro</sup> and T116/S117 in EAV 3CL<sup>pro</sup>, had less or no significant effect on P1-Glu/P1-Gln cleavage activity. Importantly, Ser136/Ser137 in PRRSV 3CL<sup>pro</sup> /EAV 3CL<sup>pro</sup> contacts P1-Glu, but not P1-Gln, through two conserved H-bonds (S136/S137-OG(Side)...P1E-OE1(Side) and S136/S137-OG (Side)...P1E-OE2 (Side)), S136A mutant in PRRSV 3CL<sup>pro</sup> and S137A mutant in EAV 3CL<sup>pro</sup> caused a loss of H-bonds between side chains of S136/S137 and the P1-Glu (Fig. 4d column 1 vs 6, Fig. 4e column 1 vs 7), whereas S137A mutant in EAV 3CL<sup>pro</sup> had no significant effect on H-bond interactions toward P1-Gln (Fig. 4f column 1 vs 7). Further confirming this notion, the S137A mutant in EAV 3CL<sup>pro</sup> lost the ability in cleaving substrate containing P1-Glu but had a similar cleavage activity toward P1-Gln as did WT EAV 3CL<sup>pro</sup> (Fig. 4b-c). These results suggested that substrate interactions with the Ser136/Ser137 side chains in PRRSV 3CL<sup>pro</sup> / EAV 3CL<sup>pro</sup> distinguish P1-Glu and P1-Gln recognition.

Interestingly, in the flexible loop S136-G140 of PRRSV 3CL<sup>pro</sup>, a salt bridge Lys138-P1-Glu interaction is observed in PRRSV 3CL<sup>pro</sup> -P1-Glu complex and contributes to binding P1-Glu of the substrate, which might explain why the K138A mutant lost the ability to cleave substrate with P1-Glu. Surprisingly, neither EAV 3CL<sup>pro</sup> -P1-Glu nor EAV 3CL<sup>pro</sup>-P1-Gln complexes show the salt bridge observed with Lys138-P1-Glu in the PRRSV 3CL<sup>pro</sup> substrate complexes. Considering that PRRSV 3CL<sup>pro</sup> exhibited P1 preference only toward Glu and EAV had both P1-Glu and P1-Gln preference, we speculate that the salt bridge might lead to a P1-substrate specificity shift from P1-Glu/Gln to P1-Glu preference during the evolution of arterivirus 3CL<sup>pro</sup>.

3.5. Identification of an intermolecular salt bridge linking P1 substrate specificity shift of arterivirus 3CL<sup>pro</sup>

To further test whether the Lys138-P1-Glu salt bridge links the P1 substrate specificity shift of arterivirus 3CL<sup>pro</sup>, we attempted to...
Fig. 3. Exploiting the biosensor assay to evaluate the protease-substrate specificity of arterivirus 3CLpro in vivo. (a) Diagram of the generation of luciferase-based biosensors. The purple and the orange structures represent the recombined firefly luciferase. The green rectangle indicates the different recognition sequences that were used to evaluate the activity of PRRSV 3CLpro or EAV 3CLpro. (b and c) HEK293T cells in 24-well plates were transfected with the pRL-TK plasmid, the plasmid encoding PRRSV 3CLpro or EAV 3CLpro, and corresponding luciferase-based biosensors plasmids, or their control 233D. After 28 h transfection, cell lysates were prepared and analyzed by dual-luciferase assays (b) and western blotting (c). ns, not significant. ***p < 0.001. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
convert P1 substrate specificity of PRRSV 3CL\textsuperscript{pro} to that of EAV 3CL\textsuperscript{pro}. The only amino acid distinction in the binding site region between PRRSV 3CL\textsuperscript{pro} and EAV 3CL\textsuperscript{pro} is in position 138 at the loop near the active site, where Lys is found in PRRSV 3CL\textsuperscript{pro} and Thr in EAV 3CL\textsuperscript{pro}. We generated the T139K mutation in EAV 3CL\textsuperscript{pro}, which we predicted would form a salt bridge linking P1-Glu of the substrate. As shown in Fig. 5a, the T139K mutant of EAV 3CL\textsuperscript{pro} revealed a significant shift in substrate specificity, recognizing substrates containing Glu at P1 more efficiently. Gln-containing substrates were still hydrolyzed, but with much less relative efficiency as compared with the WT EAV 3CL\textsuperscript{pro} (Fig. 5b). Next, we tested the opposite mutations in PRRSV 3CL\textsuperscript{pro} to determine whether mutant PRRSV 3CL\textsuperscript{pro} would behave like EAV 3CL\textsuperscript{pro} and Thr in EAV 3CL\textsuperscript{pro}.

We generated the T139K mutation in EAV 3CL\textsuperscript{pro}, which we predicted would form a salt bridge linking P1-Glu of the substrate. As shown in Fig. 5a, the T139K mutant of EAV 3CL\textsuperscript{pro} revealed a significant shift in substrate specificity, recognizing substrates containing Glu at P1 more efficiently. Gln-containing substrates were still hydrolyzed, but with much less relative efficiency as compared with the WT EAV 3CL\textsuperscript{pro} (Fig. 5b). Next, we tested the opposite mutations in PRRSV 3CL\textsuperscript{pro} to determine whether mutant PRRSV 3CL\textsuperscript{pro} would behave like EAV 3CL\textsuperscript{pro} and Thr in EAV 3CL\textsuperscript{pro}.

The only amino acid distinction in the binding site region between PRRSV 3CL\textsuperscript{pro} and EAV 3CL\textsuperscript{pro} is in position 138 at the loop near the active site, where Lys is found in PRRSV 3CL\textsuperscript{pro} and Thr in EAV 3CL\textsuperscript{pro}. We generated the T139K mutation in EAV 3CL\textsuperscript{pro}, which we predicted would form a salt bridge linking P1-Glu of the substrate. As shown in Fig. 5a, the T139K mutant of EAV 3CL\textsuperscript{pro} revealed a significant shift in substrate specificity, recognizing substrates containing Glu at P1 more efficiently. Gln-containing substrates were still hydrolyzed, but with much less relative efficiency as compared with the WT EAV 3CL\textsuperscript{pro} (Fig. 5b). Next, we tested the opposite mutations in PRRSV 3CL\textsuperscript{pro} to determine whether mutant PRRSV 3CL\textsuperscript{pro} would behave like EAV 3CL\textsuperscript{pro} and Thr in EAV 3CL\textsuperscript{pro}.

We generated the T139K mutation in EAV 3CL\textsuperscript{pro}, which we predicted would form a salt bridge linking P1-Glu of the substrate. As shown in Fig. 5a, the T139K mutant of EAV 3CL\textsuperscript{pro} revealed a significant shift in substrate specificity, recognizing substrates containing Glu at P1 more efficiently. Gln-containing substrates were still hydrolyzed, but with much less relative efficiency as compared with the WT EAV 3CL\textsuperscript{pro} (Fig. 5b). Next, we tested the opposite mutations in PRRSV 3CL\textsuperscript{pro} to determine whether mutant PRRSV 3CL\textsuperscript{pro} would behave like EAV 3CL\textsuperscript{pro} and Thr in EAV 3CL\textsuperscript{pro}.

We generated the T139K mutation in EAV 3CL\textsuperscript{pro}, which we predicted would form a salt bridge linking P1-Glu of the substrate. As shown in Fig. 5a, the T139K mutant of EAV 3CL\textsuperscript{pro} revealed a significant shift in substrate specificity, recognizing substrates containing Glu at P1 more efficiently. Gln-containing substrates were still hydrolyzed, but with much less relative efficiency as compared with the WT EAV 3CL\textsuperscript{pro} (Fig. 5b). Next, we tested the opposite mutations in PRRSV 3CL\textsuperscript{pro} to determine whether mutant PRRSV 3CL\textsuperscript{pro} would behave like EAV 3CL\textsuperscript{pro} and Thr in EAV 3CL\textsuperscript{pro}.

We generated the T139K mutation in EAV 3CL\textsuperscript{pro}, which we predicted would form a salt bridge linking P1-Glu of the substrate. As shown in Fig. 5a, the T139K mutant of EAV 3CL\textsuperscript{pro} revealed a significant shift in substrate specificity, recognizing substrates containing Glu at P1 more efficiently. Gln-containing substrates were still hydrolyzed, but with much less relative efficiency as compared with the WT EAV 3CL\textsuperscript{pro} (Fig. 5b). Next, we tested the opposite mutations in PRRSV 3CL\textsuperscript{pro} to determine whether mutant PRRSV 3CL\textsuperscript{pro} would behave like EAV 3CL\textsuperscript{pro} and Thr in EAV 3CL\textsuperscript{pro}.

We generated the T139K mutation in EAV 3CL\textsuperscript{pro}, which we predicted would form a salt bridge linking P1-Glu of the substrate. As shown in Fig. 5a, the T139K mutant of EAV 3CL\textsuperscript{pro} revealed a significant shift in substrate specificity, recognizing substrates containing Glu at P1 more efficiently. Gln-containing substrates were still hydrolyzed, but with much less relative efficiency as compared with the WT EAV 3CL\textsuperscript{pro} (Fig. 5b). Next, we tested the opposite mutations in PRRSV 3CL\textsuperscript{pro} to determine whether mutant PRRSV 3CL\textsuperscript{pro} would behave like EAV 3CL\textsuperscript{pro} and Thr in EAV 3CL\textsuperscript{pro}.
3.6. Evolutionary mechanism of salt-bridge-triggered P1 substrate specificity switch of arterivirus 3CL\textsuperscript{pro}

To exclude the effect of different substrates on the above findings, two different substrates with Gln at P1 position were selected and tested whether the salt bridge disruption of PRRSV 3CL\textsuperscript{pro} leads to the P1 substrate specificity switch. First, we constructed a mutant of PRRSV-nsp3/4 biosensor in which the Glu at P1 position was replaced with Gln and examined their cleavage by PRRSV 3CL\textsuperscript{pro}. As expected, the luciferase activity of P1-E→Q substitution of PRRSV-nsp3/4 biosensor was markedly induced in the K138T mutant of PRRSV 3CL\textsuperscript{pro}-cotransfected cells, whereas no activity was detected in wild type of PRRSV 3CL\textsuperscript{pro}-cotransfected cells (Fig. 7\textsuperscript{a}).

Considering that arterivirus 3CL\textsuperscript{pro} cleaves not only viral polyproteins but also host proteins to facilitate the various steps of the virus infection cycle [30]. Our previous study found that EAV 3CL\textsuperscript{pro}, but not PRRSV 3CL\textsuperscript{pro}, cleaved host protein NEMO at Q205 in the P1 position [30]. We then chose NEMO as a representative host protein to study the effect of salt bridge disruption of PRRSV 3CL\textsuperscript{pro} on the cleavage of the host protein. In the NEMO-Q205 biosensor assay, the P6-P60 positions of EAV 3CL\textsuperscript{pro} recognition regions around NEMO-Q205 were introduced into the cyclized luciferase biosensor (Fig. 3\textsuperscript{a}). Consistent with our previous findings, wild type of PRRSV 3CL\textsuperscript{pro} did not induce the luciferase activity of NEMO-Q205 biosensor, whereas such activity was significantly induced in the K138T mutant of PRRSV 3CL\textsuperscript{pro} cotransfected cells (Fig. 7\textsuperscript{b}). As further evidence supporting this result, western blotting also revealed that the K138T mutant of PRRSV 3CL\textsuperscript{pro} was able to cleave NEMO at Q205, producing an expected faster migrating protein band that was detected by an anti-Flag antibody and had molecular masses of approximately 22 kDa. Presumably, this band is NEMO cleavage product, NEMO-(1–205), mediated by the K138T mutant. As we predicted, no cleavage products were observed for WT PRRSV 3CL\textsuperscript{pro} (Fig. 7\textsuperscript{c}). These data suggested that the salt bridge disruption of PRRSV 3CL\textsuperscript{pro} led to a substrate preference shift of arterivirus 3CL\textsuperscript{pro} from P1-Glu to P1-Gln.

To further investigate the importance of the residue K138 in the salt bridge, two additional mutants (K138L and K138M, mutants with hydrophobic side chain) of PRRSV 3CL\textsuperscript{pro} were constructed. As shown in Fig. 7\textsuperscript{d-e}, these two mutants (K138L and K138M) also gained the ability to cleave the substrate with P1-Gln, while significantly losing the ability to cleave the substrate with P1-Glu, as did the mutants K138A and K138T. Together, these findings identify an evolutionarily accessible mechanism for disrupting or reorganizing salt bridge with only a single mutation of arterivirus 3CL\textsuperscript{pro} to trigger a substrate specificity shift.
4. Discussion

In most positive-sense, single-stranded RNA viruses, viral 3CL\textsuperscript{pro} or 3C proteases (3C\textsuperscript{pro}) cleaved the viral polyprotein at multiple conserved sites and are engaged in translational processing of the viral non-structural proteins [54–59]. The hydrolysis sites of 3C\textsuperscript{pro}/3CL\textsuperscript{pro} are broadly resembling and usually include a Gln or Glu residue at the P1 position along with a small amino acid residue downstream [34,60,61]. For arterivirus 3CL\textsuperscript{pro}, EAV 3CL\textsuperscript{pro} cleave the bonds formed by both P1-Glu and P1-Gln like some of picornaviruses (e.g., Foot-and-mouth disease virus), while PRRSV 3CL\textsuperscript{pro} and Glu-specific protease from S. griseus (Glu-SGP, the first Glu-specific protease whose spatial structure was determined) prefer P1-Glu [60,62–64]. Here, we established a rapid, sensitive, and efficient luciferase-based biosensor to monitor the activity of arterivirus 3CL\textsuperscript{pro} in vivo and identified key amino acids for P1 substrate specificity in arterivirus 3CL\textsuperscript{pro} by site-directed mutagenesis and MD simulations. Specifically, a potential intermolecular salt-bridge in arterivirus 3CL\textsuperscript{pro} linked substrate binding and P1 substrate switch, this finding could augment the discovery of new specific inhibitors against arterivirus 3CL\textsuperscript{pro} [65–67].

Previous studies have shown that the structures of the S1 sub-sites of Glu-SGP, picornavirus 3C\textsuperscript{pro} and arterivirus 3CL\textsuperscript{pro} are very close (Fig. S6) [35,63,64,68]. The S1 subsite comprises the three main structural elements: the structures of the S1 subsite have three main structural elements: a conserved histidine residue (His213 in Glu-SGP, His182 in FMDV 3C\textsuperscript{pro}, His133 in PRRSV 3CL\textsuperscript{pro} and His134 in EAV 3CL\textsuperscript{pro}) at the base of the S1 subsite, a conserved Ser/Thr residue lining one "wall" of the S1 subsite (His213 in Glu-SGP, His182 in FMDV 3C\textsuperscript{pro}, His133 in PRRSV 3CL\textsuperscript{pro} and His134 in EAV 3CL\textsuperscript{pro}) at the base of the S1 subsite, a conserved Ser/Thr residue lining one "wall" of the S1 subsite (Ser190 in Glu-SGP, Thr159 in FMDV 3C\textsuperscript{pro}, Thr113 in PRRSV 3CL\textsuperscript{pro}, and His134 in EAV 3CL\textsuperscript{pro}) at the base of the S1 subsite, a conserved Ser/Thr residue lining one "wall" of the S1 subsite (Ser190 in Glu-SGP, Thr159 in FMDV 3C\textsuperscript{pro}, Thr113 in PRRSV 3CL\textsuperscript{pro}, and His134 in EAV 3CL\textsuperscript{pro}) at the base of the S1 subsite, a conserved Ser/Thr residue lining one "wall" of the S1 subsite (Ser190 in Glu-SGP, Thr159 in FMDV 3C\textsuperscript{pro}, Thr113 in PRRSV 3CL\textsuperscript{pro}, and His134 in EAV 3CL\textsuperscript{pro}). Previously, the significance of residues His134 and Thr115 for EAV 3CL\textsuperscript{pro} function was supported by site-directed mutagenesis experiments. The mutations His134 and Thr115 for EAV 3CL\textsuperscript{pro} function was supported by site-directed mutagenesis experiments. The mutations His134 and Thr115 for EAV 3CL\textsuperscript{pro} function was supported by site-directed mutagenesis experiments. The mutations His134 and Thr115 for EAV 3CL\textsuperscript{pro} function was supported by site-directed mutagenesis experiments. The mutations His134 and Thr115 for EAV 3CL\textsuperscript{pro} function was supported by site-directed mutagenesis experiments. The mutations His134 and Thr115 for EAV 3CL\textsuperscript{pro} function was supported by site-directed mutagenesis experiments. The mutations His134 and Thr115 for EAV 3CL\textsuperscript{pro} function was supported by site-directed mutagenesis experiments. The mutations His134 and Thr115 for EAV 3CL\textsuperscript{pro} function was supported by site-directed mutagenesis experiments. The mutations His134 and Thr115 for EAV 3CL\textsuperscript{pro} function was supported by site-directed mutagenesis experiments. The mutations His134 and Thr115 for EAV 3CL\textsuperscript{pro} function was supported by site-directed mutagenesis experiments. The mutations His134 and Thr115 for EAV 3CL\textsuperscript{pro} function was supported by site-directed mutagenesis experiments. The mutations His134 and Thr115 for EAV 3CL\textsuperscript{pro} function was supported by site-directed mutagenesis experiments. The mutations His134 and Thr115 for EAV 3CL\textsuperscript{pro} function was supported by site-directed mutagenesis experiments. The mutations His134 and Thr115 for EAV 3CL\textsuperscript{pro} function was supported by site-directed mutagenesis experiments. The mutations His134 and Thr115 for EAV 3CL\textsuperscript{pro} function was supported by site-directed mutagenesis experiments. The mutations His134 and Thr115 for EAV 3CL\textsuperscript{pro} function was supported by site-directed mutagenesis experiments.
decreased the processing efficiency. Moreover, our findings highlight the essential significance of His134 and the considerably smaller role of the residues Thr115 for the cleavage of substrates with both P1-Glu and P1-Gln by EAV 3CL\textsuperscript{pro}. Meanwhile, our MD analysis also showed that these two residues are essential for the formation of an appropriate geometry of the S1 subsite for recognizing both P1-Glu and P1-Gln (e.g., EAV 3CL\textsuperscript{pro}, FMDV 3C\textsuperscript{pro}).

The third element of the S1 site in Glu-SGP is Ser216, which is highly conserved in PRRSV 3CL\textsuperscript{pro} and EAV 3CL\textsuperscript{pro} but replaced with Gly or Ala residues in FMDV 3C\textsuperscript{pro}. It is worth noting that the mutagenesis in the Glu-SGP model indicates that the Ser216Ala/Gly replacement decreased the efficiency in their cleavage activity over P1-Glu substrate but did not make the substrates with P1-Gln the preferred ones\cite{63,69}. Consistent with these results, the mutant S136A of PRRSV 3CL\textsuperscript{pro} lost the ability to cleave substrates with P1-Glu (Fig. 4a) but did not lead to a P1 substrate specificity shift from P1-Glu to P1-Gln preference (data not shown). Since PRRSV 3CL\textsuperscript{pro} plays an indispensable role in the virus life cycle by cleaving the polyprotein precursors, we attempted to investigate whether the intermolecular salt bridge in PRRSV 3CL\textsuperscript{pro} is important for its replication by a PRRSV reverse genetics approach. However, the salt bridge-breaking mutant of PRRSV 3CL\textsuperscript{pro} (PRRSV 3CL\textsuperscript{pro}-K138T) could not be rescued, and one possible reason is that the salt bridge-breaking mutation almost completely lost the ability to cleave the substrate P1-E in the polyprotein of PRRSV (Fig. 5c).

5. Conclusion

Ensuring P1-Gln substrate specificity in PRRSV 3CL\textsuperscript{pro} required the substitution of Lys138 in the S1 subsite and disruption of a salt bridge between protease and substrate. In turn, the T139K mutant of EAV 3CL\textsuperscript{pro}, building a salt bridge, leads to a shift of EAV 3CL\textsuperscript{pro} from P1-Gln to P1-Glu preference. In conclusion, our findings identified an evolutionarily accessible mechanism for disrupting or reorganizing salt bridge with only a single mutation of viral proteases to trigger a substrate specificity switch.

Fig. 7. Evolutionary mechanism of salt-bridge-triggered P1 substrate specificity switch of PRRSV 3CL\textsuperscript{pro}. (a) HEK293T cells were cotransfected with PRRSV-nsp3/4-Q biosensor plasmid, pRL-TK and a plasmid encoding PRRSV 3CL\textsuperscript{pro} WT or the mutant K138T. After 28 h transfection, cell lysates were prepared and analyzed by dual-luciferase assays and western blotting. **p < 0.01. (b) HEK293T cells were cotransfected with NEMO-Q205 biosensor plasmid, pRL-TK and a plasmid encoding PRRSV 3CL\textsuperscript{pro} WT or the mutant K138T, along with Flag-tagged NEMO (E166A-E171A-E349A). Cell lysates were prepared 30 h post transfection and analyzed by western blotting. (d-e) HEK293T cells were cotransfected with plasmids encoding WT PRRSV 3CL\textsuperscript{pro} or its mutants, pRL-TK and corresponding biosensor plasmid. After 28 h transfection, cell lysates were prepared and analyzed by dual-luciferase assays and western blotting.

Q. Chen, J. Zhou, Z. Yang et al. Computational and Structural Biotechnology Journal 20 (2022) 3409–3421

3419
Author contributions
D.W., Q.C., L.F., S.X. and J.Z. conceptualization; Q.C., J.Z., Y.G., Z. L., X.S., and Q.J. methodology; J.Z. and Q.C. data curation; J.Z. and Q.C. software; J.Z. and Q.C. formal analysis; D.W. supervision; Q.C. and D.W. writing—original draft; L.F. and S.X. writing—review & editing; D.W., L.F. and S.X. funding acquisition.

CRediT authorship contribution statement
Qian Chen: Conceptualization, Methodology, Data curation, Software, Formal analysis, Writing – original draft. Junwei Zhou: Conceptualization, Methodology, Data curation, Software, Formal analysis. Zhixiang Yang: Methodology. Jiahui Guo: Methodology. Zimin Liu: Methodology. Xinyi Sun: Methodology. Qingsheng Jia: Methodology. Liurong Fang: Conceptualization, Writing – review & editing, Funding acquisition. Dang Wang: Conceptualization, Supervision, Writing – original draft, Funding acquisition. Shaobo Xiao: Conceptualization, Writing – review & editing, Funding acquisition.

Declaration of Competing Interest
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments
This study was supported by the National Natural Science Foundation of China (32120103, 31827485 and 31941005).

Appendix A. Supplementary data
Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2022.06.059.

References
[1] Brinton MA, Gulyaeva AA, Balasuriya UB, Dunowska M, Faaberg KS, Goldberg Q et al. Appearance of acute PRRS-like symptoms in sow herds after vaccination with a modified live PRRS vaccine. Vet Rec 1997;141:497–9. https://doi.org/10.1136/vr.141.13.497.
[2] Holyoak CR, Balasuriya UB, Broadds CC, Timoney PJ. Equine arteritis viruses: current status and prevention. Theriogenology 2008;70:403–14. https://doi.org/10.1016/j.theriogenology.2008.04.020.
[3] Snijder EJ, Kikkert M, Fang Y. Arterivirus molecular biology and pathogenesis. J Gen Virol 2013;94:2141–63. https://doi.org/10.1099/jgv.0.056341-0.
[4] van Dinten LC, Wassenaar AL, Gorbalyena AE, Spaan WJ, Snijder EJ. Procession of the equine arterivirus virus replicase ORF1b protein: identification of cleavage products containing the putative viral polymerase and helicase domains. J Virol 1996;70:6625–30. https://doi.org/10.1128/JVI.70.15.6625-6633.1996.
[5] Snijder EJ, Renssen S, Gorbalyena AE, Snijder EJ. Arterivirus molecular biology and pathogenesis. Current Biology 2013;23:1129–35. https://doi.org/10.1016/j.cub.2013.03.039.
[6] Qian Chen, J. Zhou, Z. Yang et al. Computational and Structural Biotechnology Journal 20 (2022) 3409–3421
