Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Journal of Molecular Graphics and Modelling 74 (2017) 125–134

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Topical Perspectives

Structural insights into type I and type II of nsp4 porcine reproductive and respiratory syndrome virus (nsp4 PRRSV) by molecular dynamics simulations

Niran Aeksiri[a,b,*], Tippawan Jantafong[c]

[a] Center for Agriculture Biotechnology, Faculty of Agriculture, Natural Resources, and Environment, Naresuan University, Phitsanulok 65000, Thailand
[b] Department of Agricultural Sciences, Faculty of Agriculture, Natural Resources, and Environment, Naresuan University, Phitsanulok 65000, Thailand
[c] Department of Virology and Immunology, Faculty of Veterinary Medicine, Mahanakorn University of Technology, Thailand

ARTICLE INFO

Article history:
Received 31 January 2017
Received in revised form 16 March 2017
Accepted 17 March 2017
Available online 24 March 2017

Keywords:
Porcine reproductive and respiratory virus
PRRSV
nsp4
Catalytic triad
Molecular dynamics
Principle component analysis
GROMACS

ABSTRACT

Porcine reproductive and respiratory virus (PRRSV) causes major economic concerns for the swine industry worldwide. We have performed molecular dynamics simulations (MD) and principle component analysis (PCA) to investigate the role of the catalytic triad and conformational dynamics of type I and type II of nsp4 PRRSV. The results showed that the RMSF of residues 136–142 near the active site of all models was highly flexible. Moreover, we identified the effect of single structural mutations of the catalytic triad. The percentage of residue with a 0.1 nm RMSF value and PCA results revealed that the mutations affected domain I and II suggesting the wild types were more stable than the mutants. At the catalytic triad, the distances between H39 and S118 were very flexible while the distances between H39 and D64 were very stable. H39, D64 and S118 showed high occupancy percentage of the hydrogen bond interaction with many residues that are conserved in PRRSVAS, PRRSVES, LDVC, LDVP and EAV. Moreover, S118 of wild-type protein formed H-bonds with T134 and G135 but these interactions were lost in PRRSVAV (S118A) and PRRSVES (S117A) indicating that the substitution of important H-bond interaction in PRRSVAS (S118A) and PRRSVES (S117A) affected the flexibility around the catalytic triad, conformation and proteolytic activity. Overall, our study may provide the structural basic of the catalytic triad and be useful for testing the protein activity in future experiments.

1. Introduction

Porcine reproductive and respiratory syndrome virus (PRRSV) is an important disease causing economic loss in the swine industry worldwide. PRRSV is the member of family *Arteriviridae* in the order *Nidovirales* which also includes equine arteritis virus (EAV), human coronavirus 229E, and severe acute respiratory syndrome coronavirus (SARS-CoV) which are similar in terms of structure and genome [1]. Nowadays, PRRSV is characterized into two group namely genotype I (European strain, PRRSVES) and genotype II (North American strain, PRRSVAS) due to the genetic difference which is 60% of nucleotide sequence similarity [2]. PRRSV causes severe reproductive failure in sows and respiratory symptoms in all ages of pigs [3,4]. In addition, this viral infection predisposes pigs to secondary bacterial and viral infections due to immunosuppression induced by this virus [5–7]. This disease was first discovered in the United States [8,9] and in Europe [10] and then reported in North American, Europe and Asia simultaneously [11–14]. At present, decreased pig production caused by PRRSV remains a critically important viral disease in many countries worldwide. Particularly, the emergence of highly virulent PRRSV strains in Asia and Europe threatens the global swine industry [15–17].

PRRSV, a small-enveloped RNA, is a positive-sense RNA virus. The genome is approximately 15.4 kb in size and encodes at least 10 open reading frames (ORFs) including ORF1a, ORF1b, and ORF2-7. ORF1a and ORF1b encode two long polypeptides (pp) namely pp1a and pp1b. After enzymatic cleavage, they produce at least 14 non-structural proteins (nsps) and two additional viral proteins called nap2TF and nsp2N. The cleaveage of nsps in the proteolytic process is performed by four viral proteases: nsp1α, nsp1β, nsp2 and nsp4. ORF2 to ORF6 code for eight structural proteins: GP2, E, GP3, GP4, GP5, GP5a, M and N proteins. As a member of the *Arteriviridae*, PRRSV is very similar to EAV in terms of genome organization. Espe-

* Corresponding author at: Department of Agricultural Science, Faculty of Agriculture, Natural Resources, and Environment, Naresuan University, 99 M. 1, T. Thapo, A. Muang, Phitsanulok 65000, Thailand.
E-mail address: nirana@nu.ac.th (N. Aeksiri).

http://dx.doi.org/10.1016/j.jmgm.2017.03.015
1093-3263/

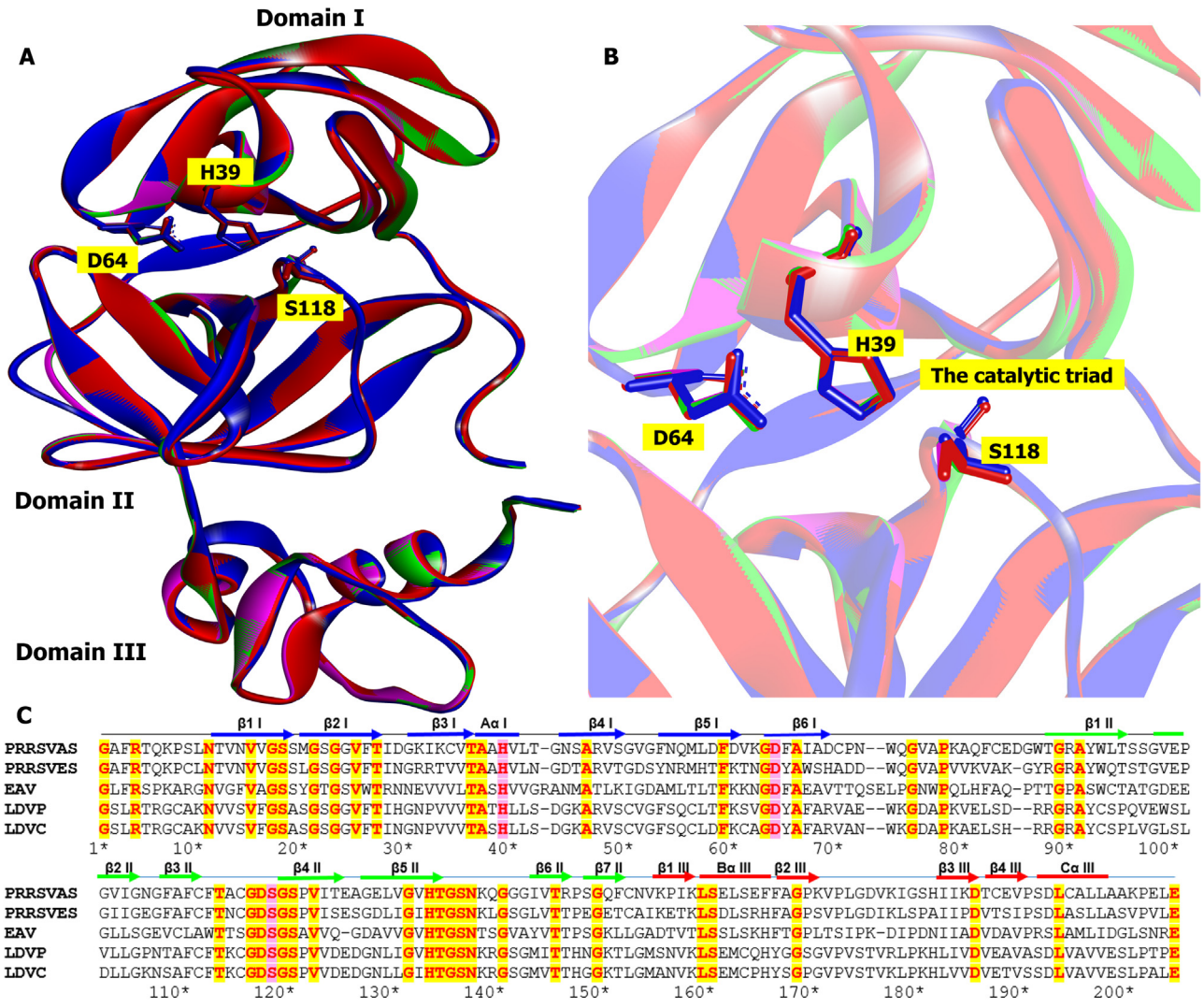

Fig. 1. The structures of PRRSV nsp4.
(A) The superimpose of 4 PRRSV nsp4 structures with the catalytic triad which represent by solid ribbon structure.
(B) The catalytic triad of PRRSVAS includes H39, D64 and S118 while PRRSVAS comprises H39, D64 and S117.
(C) Multiple sequence alignment of the chymotrypsin-like serine proteinase(3CLSPs) including PRRSVAS and PRRSVES (PRRSV American strain and European strain), LDVC and LDVP (lactate dehydrogenase-elevating virus neurovirulent type C and stain Plagemann). The pink colors indicate the catalytic triad.

cially, the nsp4 of PRRSV which is an orthologous protease of EAV nsp4 and regarded as the primary protease. The sequence identity between PRRSV and EAV nsp4 is more than 34%, and they share great structural similarity at the tertiary level. Nsp4, the 3C-like serine protease (3CLSP), cleaves nsp3 to nsp12, which is responsible for most of the nonstructural protein processing. The crystal structure of nsp4 [18] shows that nsp4 folds into three domains, in which the domains I and II are two chymotrypsin-like β-barrels domains and the domain III is an extra C-terminal α/β domain similar to EAV nsp4. The active site is situated between domains I and II and harbors a canonical catalytic triad which comprises H39, D64 and S188 [18,19]. The structure also reveals an atypical oxyanion hole and a partially collapsed S1 specificity pocket. H39, D64 and S118 are required for nsp4 to induce apoptosis, and nsp4-induced apoptosis is dependent on its serine proteinase activity. Moreover, the proteolytic activity decreased significantly, when the catalytic serine residue was mutated to alanine (S118A). It clearly confirmed the vital function of Ser118 as the nucleophile in catalysis [20,21]. The abilities of nsp4 not only include functioning as the main proteinase but also suppressing NF-kB-mediated IFN-β-production by interfering with the NF-kB signaling pathway in HeLa cells and CRL-2843 cells though the NF-kB essential modulator (NEMO) at the E349-S350 site [22,23].

Although comparative study of genetic similarity between PRRSV type I and type II ranges from 55% to 62% for non-structural protein and from 61% to 81% for structural protein, both are considered as the same virus and cause the same disease. Nevertheless, type I and type II isolates are biologically and genetically distinct. For example, type I viruses are preferentially propagated in porcine alveolar macrophages (PAM) and rarely can be grown on continuous cell lines, while type II viruses can routinely replicate in PAM and simian cell lines [24,25]. Type II viruses are more virulent than type I viruses in terms of their ability to cause respiratory diseases [26,27].

Molecular dynamics simulation (MD) is a tool for elucidating protein flexibility and rigidity. Moreover, it has been widely used to study the effect of single or multiple mutations. For example, it was applied to several proteins such as L-alanine dehydrogenase [28], protein tyrosine phosphatase 1B [29], HIV-1 RT [30] and RAS-relate C3 botulinum toxin substrate 1(RAC1) [31]. Therefore, MD simulations have been used for extensive analysis, and the results often help to explain the experimental data. Although the

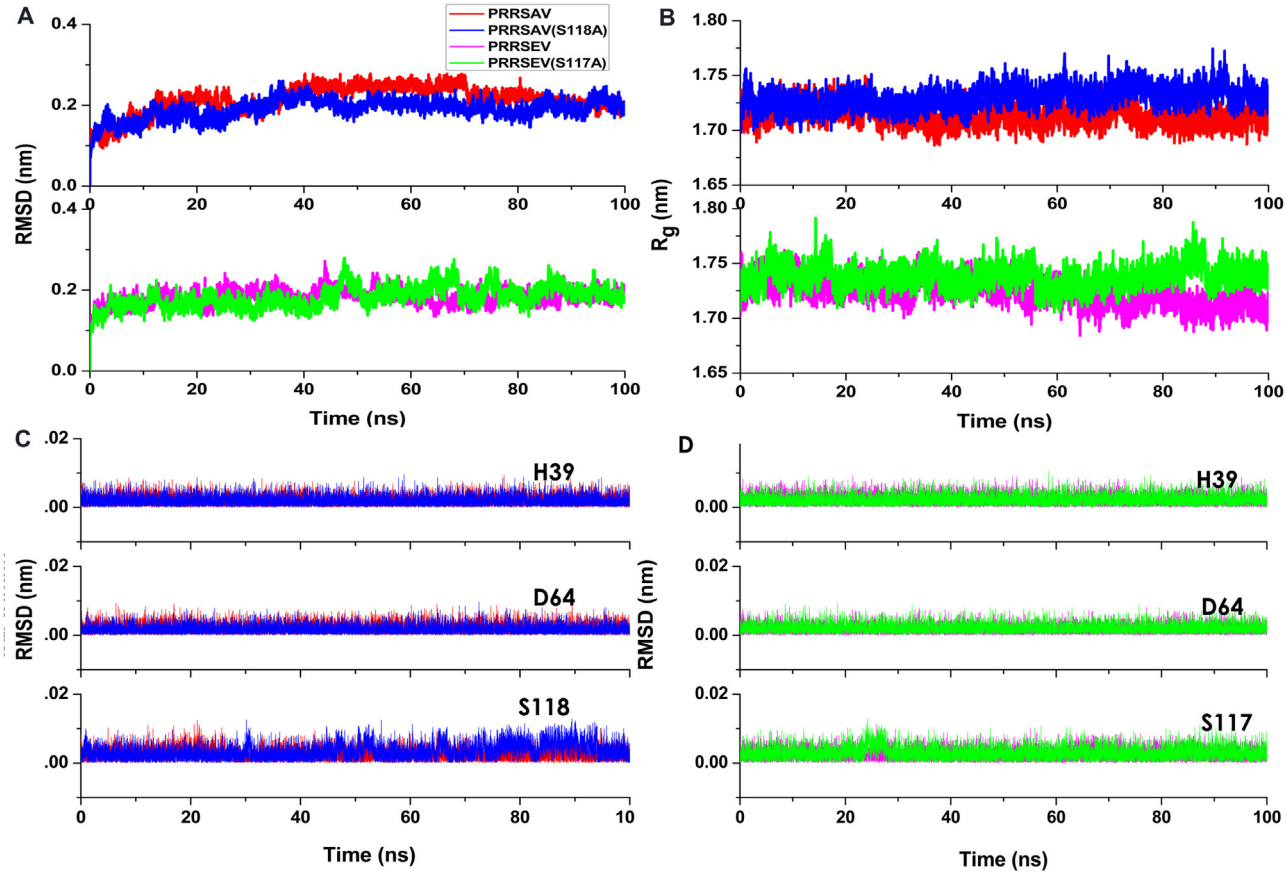

Fig. 2. The root mean square deviations (RMSD) of the backbone atoms from the appropriate starting structure for each simulation: PRRSVAS (red), PRRSVAS(S118A) (blue), PRRSVES (pink) and PRRSVES(S117A) (green) (A). The radius of gyration (Rg) was plotted as the function of stimulation time (B). RMSD of the catalytic triad residues of PRRSVAS (C) and PRRSVES (D). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

structural and functional features of nsp4 PRRSV have been provided experimentally, various important aspects have not yet been addressed. Particularly, there is a lack of data on the comparison of nsp4 between type I and type II isolates. Therefore, the objectives of this study were to study the structural conformations of whole protein and the canonical catalytic triad including the mutation of its pocket site. Mutation of the amino acid at the catalytic triad could influence the native conformation. Therefore, the computational three-dimensional model structures of the protein might help in understanding the dynamics of conformational changes.

2. Computational methods

2.1. Structural preparation

In this study, we evaluated four models including nsp4 of PRRSVAS(3FAN) [18], PRRSVAS (S118A; 3FAO) [18], PRRSVES and PRRSVES(S117A). The atomic coordinates of nsp4 of PRRSVAS were obtained from the Protein Data Bank (PDB). The models were generated using the PDB entries 3FAN and 3FAO, respectively. 3FAO had a substitution of Ala for Ser118 at the catalytic triad pocket. The SWISS-MODEL program, a protein structure homology-modeling server (http://swissmodel.expasy.org) [32], was used to enable the absence of electron density in the loop 136–140. The method is based on searching high identity protein template structures to build models for evolutionarily related target proteins. All missing atoms were reconstructed in Swiss-Pdb Viewer v4.1. After constructing nsp4 PRRSVAS, the three-dimensional models of nsp4 PRRSVES and PRRSVES (S117A) were modeled by a comparative modeling strategy using SWISS-MODEL based on evolutionarily related 3FAN to generate a structural model. The first model was selected and showed the best QMEAN score [33]. The programs Verify-3D [34], ERRAT [35] and PROCHECK [36] were used to perform quality analysis of the models at the SAVES server (https://services.mbi.ucla.edu/SAVES/).

Before performing MD simulations, the N- and C- termini of each structure were capped with acetyle (ACE) and methyle amino (NME) groups.

2.2. Molecular dynamics simulations

MD simulations for all models were performed using explicit-solvent periodic boundary conditions by GROMACS v5.1 [37,38]. Each model was solvated by SPC216 water molecules in a cubic box keeping a distance of 1.0 nm between each side of the solvated box and the protein. We added three sodium ions (Na^+) to make the simulation system electrically neutral using the default program. The minimization was performed using the steepest descent algorithm until the maximum force was less than 1000 kJ/mol/nm on every atom and was implemented with the Amber ff99SB-ILDN force field [39]. The equilibration for dynamic simulations was performed under NVT (constant number of particles, volume, and temperature) for 1000 ps and then was followed by NPT (constant number of particles, pressure and temperature) for 5000 ps. The first step was conducted under the NVT condition at 300 K using the modified Berendsen thermostat [40,41]. The NPT ensemble followed at 1 bar of pressure using Parrinello-Rahman barostat [42,43]. After equilibration, the full dynamic production was performed for a period of 100 ns under the NPT condition at 300 K and 1 bar of the system. The LINCS algorithm was used to constrain the bond lengths

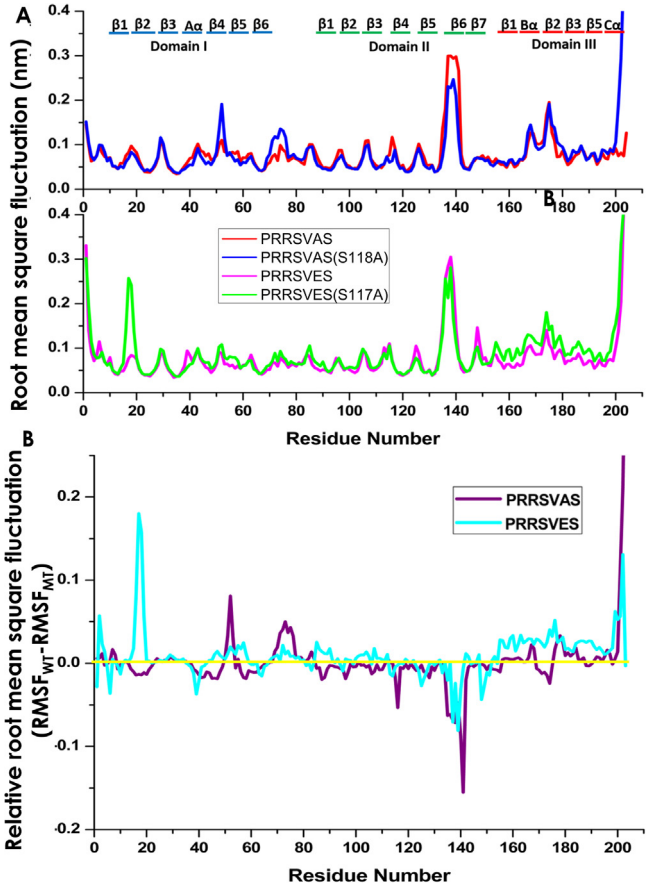

Fig. 3. Root mean square fluctuations (RMSF) by residue: PRRSVAS(red), PRRSVAS(S118A) (blue), PRRSVES (pink) and PRRSVES(S117A) (green). RMSF indicates conformational changes from the initial structure (A). Relative root mean square fluctuations (RMSF) by residue of PRRSVAS (violet) and PRRSVES (cyan). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

by performing a 2-fs time step [44]. The cut-off distance were set to 1.0 for the short rang interactions. The Particle Mesh Ewald (PME) method with a Fourier grid spacing of 0.12 nm accounted for long-range electrostatic interactions [45,46]. The atomic coordinates were recorded every 10 ps for further analysis.

2.3. Trajectory analysis

The atomic coordinates were thoroughly analyzed using the built-in analysis package of GROMACS, XMGRACE [47]. We determined the root mean square deviations (RMSD), root mean square fluctuations (RMSF), the radius of gyration (Rg), the solvent accessible surface area (SASA) and the number of hydrogen bonds (NH-bonds). The hydrogen bonds were considered following the default values which with a donor-acceptor distance ≤0.35 nm and H-donor-acceptor angle ≤30.

2.4. Principle component analysis

Principle component analysis (PCA) [48] was used to calculate the eigenvectors and eigenvalues and their projection by comparing the first two principal components. PCA was performed in order to reduce the complexity of the data and differentiate the dominant modes of motion in the models. The protein conformation was derived from the last 50 ns of simulations. After the elimination of translational and rotational movements, a covariance matrix was generated from the trajectories. Diagonalization of the covariance matrix was then performed to yield a set of eigenvectors and their corresponding eigenvalues representing the direction and amplitude of a motion. Ranking the eigenvector was subsequently conducted by decreasing eigenvalues. The first principle component analysis or the first eigenvector represented the largest contribution to the total fluctuation of the protein. The motion of each PC was visualized by projecting the protein conformations of trajectories onto each eigenvector of interest and then they were transformed back to atomic coordinates.

3. Results and discussion

3.1. Conformational analysis

In Fig. 1, nsp4 PRRSVAS shared 63% sequence identity with nsp4 PRRSVES. We have investigated molecular modeling and MD simulations studies on nsp4 PRRSVAS and PRRSVES for a detailed understanding of structural conformations especially the canonical catalytic triad. The crystal structures of nsp4 of PRRSVAS, and PRRSVAS (S118A) were generated using the PDB entries 3FAN and 3FAO, respectively. In the case of PRRSVES, three-dimensional models were built by SWISS-MODEL, a protein structure homology-modeling server (http://swissmodel.expasy.org). All initial structures were explored in depth. nsp4 PRRSVAS folds into three domains, in which the domains I and II are two chymotrypsin-like β-barrel domains and the domain III is an extra C-terminal α/β domain similar to EAV nsp4. The RMSD between PRRSVAS and PRRSVES was 0.35 nm indicating that the structures were similar. The active site is situated between domains I and II and harbors a canonical catalytic triad which is comprised of H39, D64 and S118 [18].

3.2. Flexibility of nsp4

The general properties from all simulations were shown in Table 1, which was represents 100 ns of each simulation. The root mean square deviations (RMSD), and root mean square fluctuations (RMSF) were calculated to estimate the MD trajectory quality and convergence. The 100 ns simulations first analyzed the RMSD values (Fig. 2). The RMSD values ranged from 0.06–0.27, 0.07–0.26, 0.09–0.27, 0.07–0.27 nm for PRRSVAS, PRRSVAS (S118A), PRRSVES, and PRRSVES (S117A), respectively. Moreover, the average RMSD values of the protein C-alpha over the simulations were 0.21 ± 0.03, 0.19 ± 0.02, 0.18 ± 0.02, and 0.19 ± 0.03 for PRRSVAS, PRRSVAS (S118A), PRRSVES, and PRRSVES (S117A), respectively. The RMSD values did not exceed 0.35 nm indicating that all MD trajectories reached equilibrium. The time fluctuations of the radius of gyration (Rg) were monitored to equilibration in each system. The Rg values showed similar trends indicating that the protein structure equilibrated and remained compact though the simulations. Furthermore, SASA results were analyzed to support the Rg data. The result demonstrated that SASA values followed a similar trend (Supplementary Fig. 1). It is clear that the flexibility of PRRSVAS, PRRSVAS (S118A), PRRSVES, and PRRSVES (S117A) remained stable and was similar in trend to the mutant as well. Moreover, the RMDS of the catalytic triad was also calculated. The results revealed that the S118A of PRRSVAS and S117A of PRRSVES residues were less stable than the residues of the wild-type protein (Fig. 2C,D). These results are related to biochemical studies indicating that when the catalytic serine residue mutated to alanine (S118A), the proteolytic activity decreased significantly [20,21].

The RMSF is another tool for evaluating the dynamics stability of the system. Therefore, we calculated the root mean square fluctuations (RMSF) as a function of the residue number to understand the overall flexibility of the protein. The reference structure was aver-

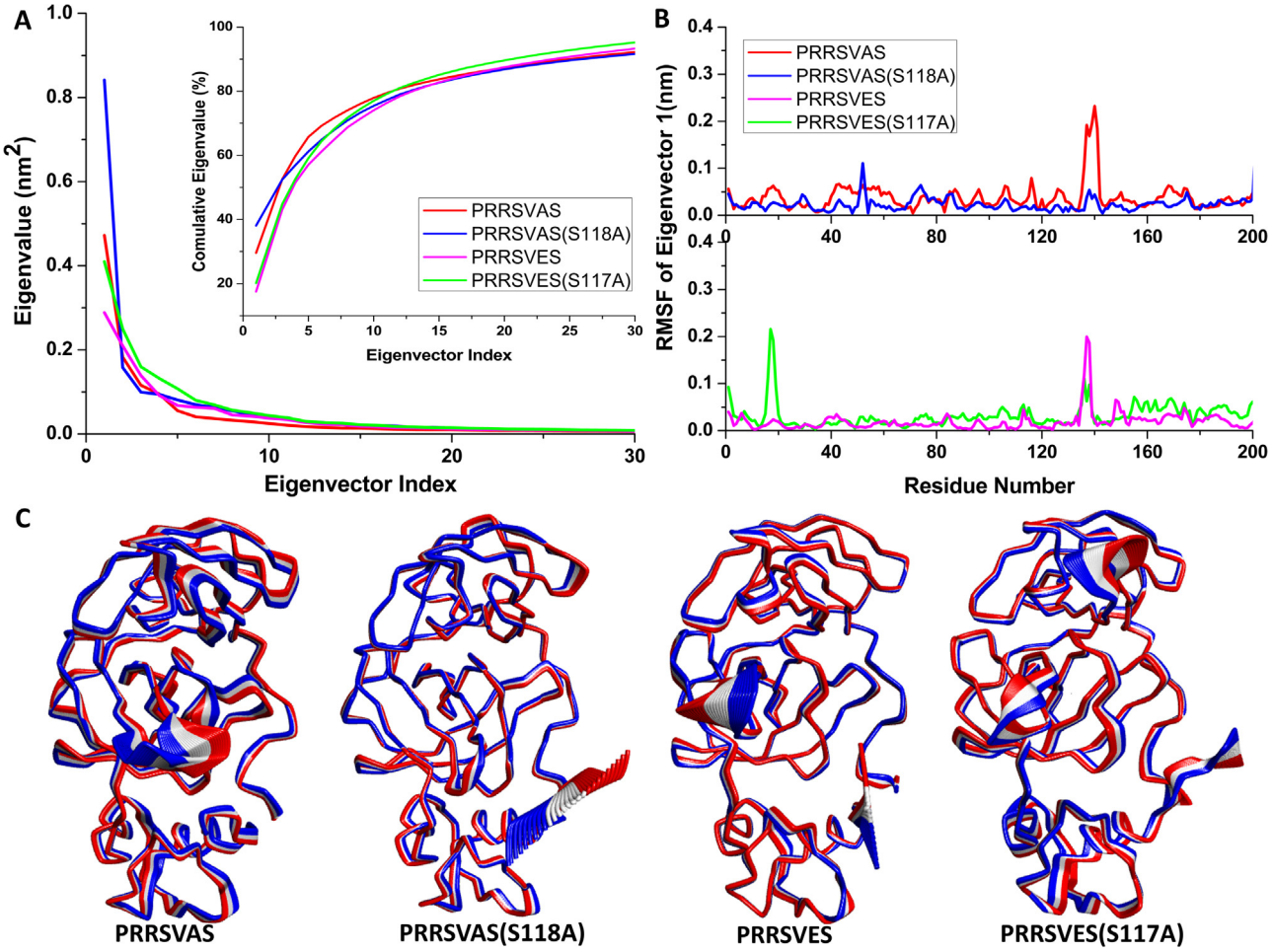

Fig. 4. Essential dynamic analysis. The cumulative proportion of the total variance calculated from a PCA of each system. Inset showed the cumulative eigenvalue: PRRSVAS (red), PRRSVAS(S118A) (blue), PRRSVES (pink) and PRRSVES(S117A) (green) (A). RMSF of eigenvector (B), colored as the same as A. Comparison of eigenvector 1 motion from each system (C). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
The general properties of four simulations including PRRSVAS, PRRSVAS(S118A), PRRSVEV and PRRSVEV(S117A).

Molecule	Cα-RMSD (nm)	Cα-RMSF (nm)	Total SASA (nm^2)	Total number of hydrogen bonds
PRRSVAS	0.23 ± 0.02	0.07 ± 0.04	114.33 ± 1.04	134.52 ± 5.24
PRRSVAS(S118A)	0.20 ± 0.02	0.08 ± 0.06	116.16 ± 0.99	129.72 ± 4.77
PRRSVES	0.18 ± 0.02	0.08 ± 0.05	115.18 ± 1.03	143.15 ± 5.68
PRRSVES(S117A)	0.20 ± 0.02	0.09 ± 0.05	116.49 ± 1.13	139.91 ± 5.42

Cα-RMSD-Root mean square deviation from the starting structure.
Cα-RMSF- Root mean square fluctuations.
SASA-Solvent accessible surface area.

aged over the last 50 ns of the simulation time. As shown in Fig. 3, the comparative RMSF values were almost of the same pattern for each system. Residues 136–142 near the active site were highly flexible as confirmed by the total lack of electron density from the crystal structure and may play an important role in enzymatic activity [18]. A critical analysis of the relative RMSF plots showed that the fluctuations were concentrated mainly in certain zone of protein. In this study we focused on the domain I and II that are the catalytic triad. We therefore compared the effect of mutation on the RMSF in each residue. We found that the mutation PRRSVAS (S118A) had an effect on residues 23–36 and 67–84, while PRRSVES (S117A) impacted the RMSF of residues 11–35, 43k59, 67–80 and 84–93. When we compared the RMSF results between PRRSVAS (S118A) and PRRSVES (S117A), the result showed that the PRRSVES (S117A) had the highest RMSF in residues 15–19, while other residues had the same lower pattern of the fluctuation (Supplementary Fig. 2).

Furthermore, we calculated the percentage of residues in each domain with a > 0.1 nm RMSF value. This calculation produced a quantitative view of the fluctuation (Table 2). The result revealed that Domain I and II had a lower percentage of residues indicating they were more rigid. When comparing wild type and mutant proteins, the mutations increased the percentage of residues in domain I and III. The domain II of mutant proteins exhibited a lower percentage of residues which had an impact on the flexibility of the protein.

3.3. *Essential dynamics of nsp4 domains*

Principle component analysis (PCA) can be used to elucidate functional dynamics. We therefore performed PCA to investigate the similarities and differences in the dynamics of each system. Additionally, the quality and direction were represented by the

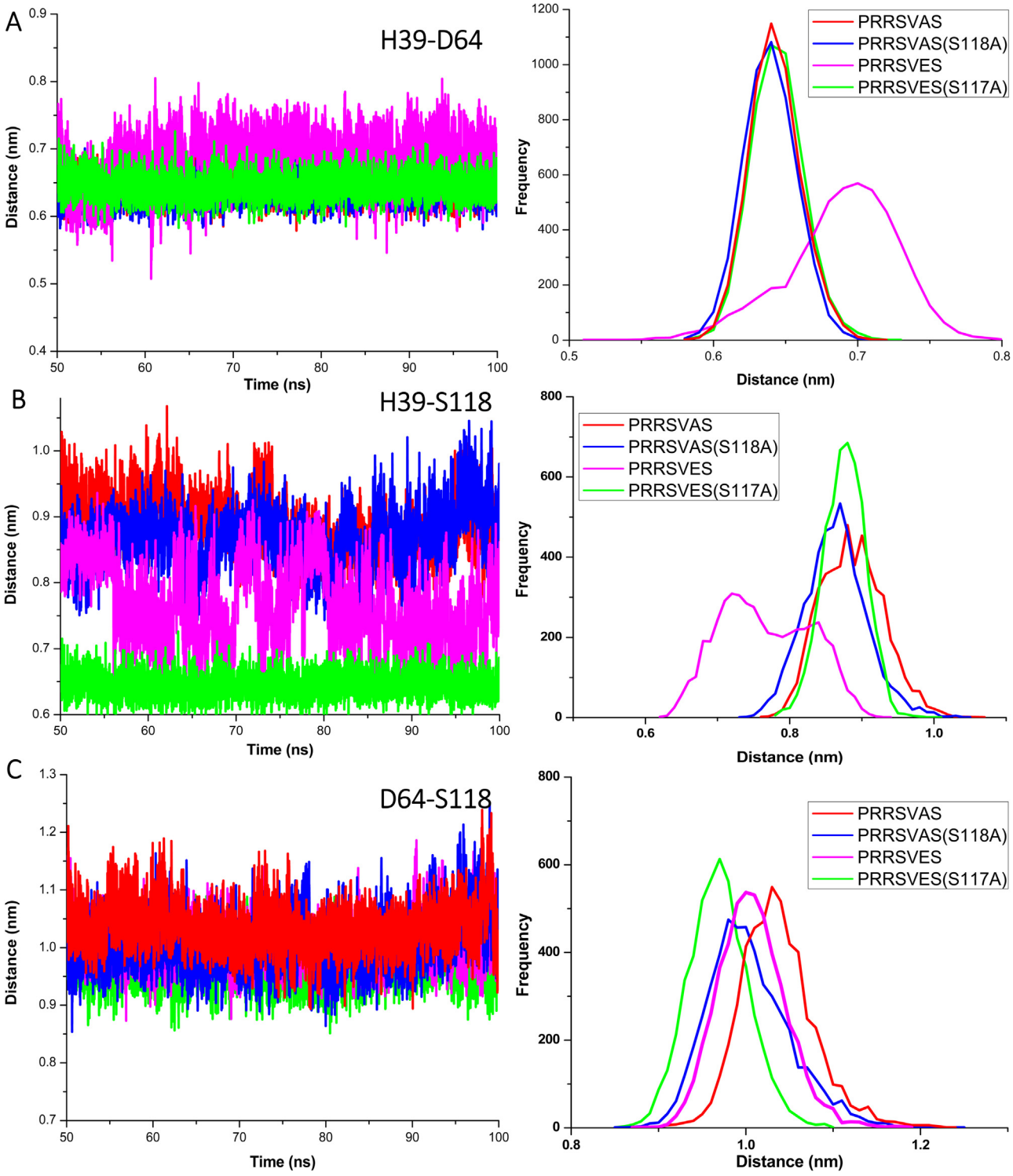

Fig. 5. The distance between C-alpha atom of the catalytic triad as the function of time(A). The distance distribution profiles (B).

Table 2
The percentage of residues in each domain with >0.1 nm RMSF value.

Molecule	DI (Residue 1–88)	DII (Residue 89–157)	DIII(Residue 158–204)	All (Residue 1–204)
PRRSVAS	4.5	14.7	24.5	12.6
PRRSVAS (S118A)	15.7	8.8	30.6	17
Molecule	DI (Residue 1–87)	D2 (Residue 88–155)	D3(Residue 156–203)	All (Residue 1–203)
PRRSVES	4.5	13.2	16.3	16.9
PRRSVES (S117A)	10.2	10.3	49	27.7

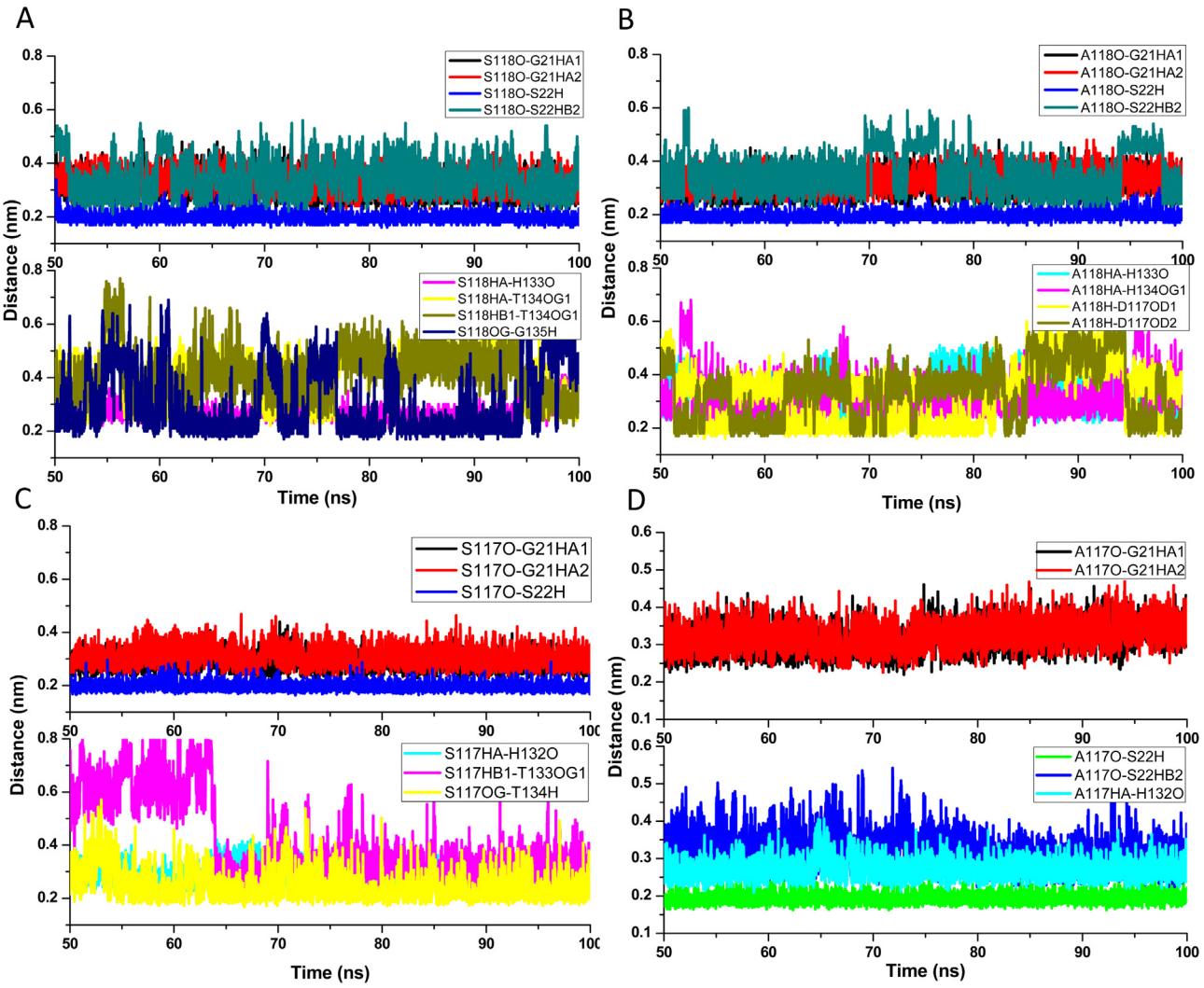

Fig. 6. The distance between 2 atoms which formed Hydrogen bond as the function of time (A). Snapshots of the interaction at the catalytic triad (B).

first principle component (PC1) which were individually generated from the MD simulations. The relative contribution of the different eigenvectors to the overall motion of the protein clearly demonstrated that the bulk of protein dynamics could be described by a small number of eigenvectors. Furthermore, the insert revealed the cumulative proportion of the total variance captured by each eigenvector from the PCA of the trajectories. The results showed that the first few eigenvectors of the four systems shared the common trend that a largest minimum plateau eigenvector is the first eigenvector and then the eigenvalues decreased with increased eigenvector index until they reached a plateau. Nevertheless, the degree of the steepness exhibited by PRRSVES, PRRSVES (S117A), PRRSVAS and PRRSVAS (S118A) suggested that the wild type protein is more stable than the mutant protein (Fig. 4A). We also calculated the displacement of PC1 of the four systems in order to reveal that the point mutations at the catalytic triad had an effect on the motions. In Fig. 4B, the result suggests that PRRSVAS (S118A) had an impact on the motions of residues 3, 5, 12, 23–37, 52, 69–79, 83–84, 121, and 133. PRRSVES (S117A) affected mainly the motion of domain I and domain II, which was consistent with the result of RMSF analysis. However, PRRSVES (S117A) obviously affected the motions of residues 15–20. Moreover, we showed in Fig. 4C that the motion of eigenvector 1 in Fig. 4C demonstrated that residues 136–142 near the active site was highly flexible.

Furthermore, the 2D projection of the MD structure created to compare regions sampled by the essential motions from different systems indicted that the results cannot be clustered (Supplementary Fig. 3). However, the 2D PC graphs of eigenvector 1 and eigenvector 2 characterizes the essential subspace of the protein. The wild type protein exhibited lower values than the mutant protein indicating that the mutation had a destabilizing effect on the protein.

The cross correlation analysis was explored between correlated and anti-correlated motions in proteins. For the cross-correlation coefficient between the atomic displacements, a strong correlation in the movements of specific residues is displayed by the highly positive regions (red); whereas strong anti-correlated motion of the residues is represented by negative regions (blue). The dynamic cross-correlation maps of PRRSVES, PRRSVES (S117A), PRRSVAS and PRRSVAS (S118A) revealed that there were very little correlated motions except the diagonal square indicating the correlated motion of the residues with itself (Supplementary Fig. 4).

3.4. Role of the canonical catalytic triad and the effects of single mutations on the structure

In 3CL protease, the catalytic triad (the S1 specificity pocket) plays an important role in substrate recognition and catalysis. It comprises three residues namely His39, Aap64 and Ser118 lying

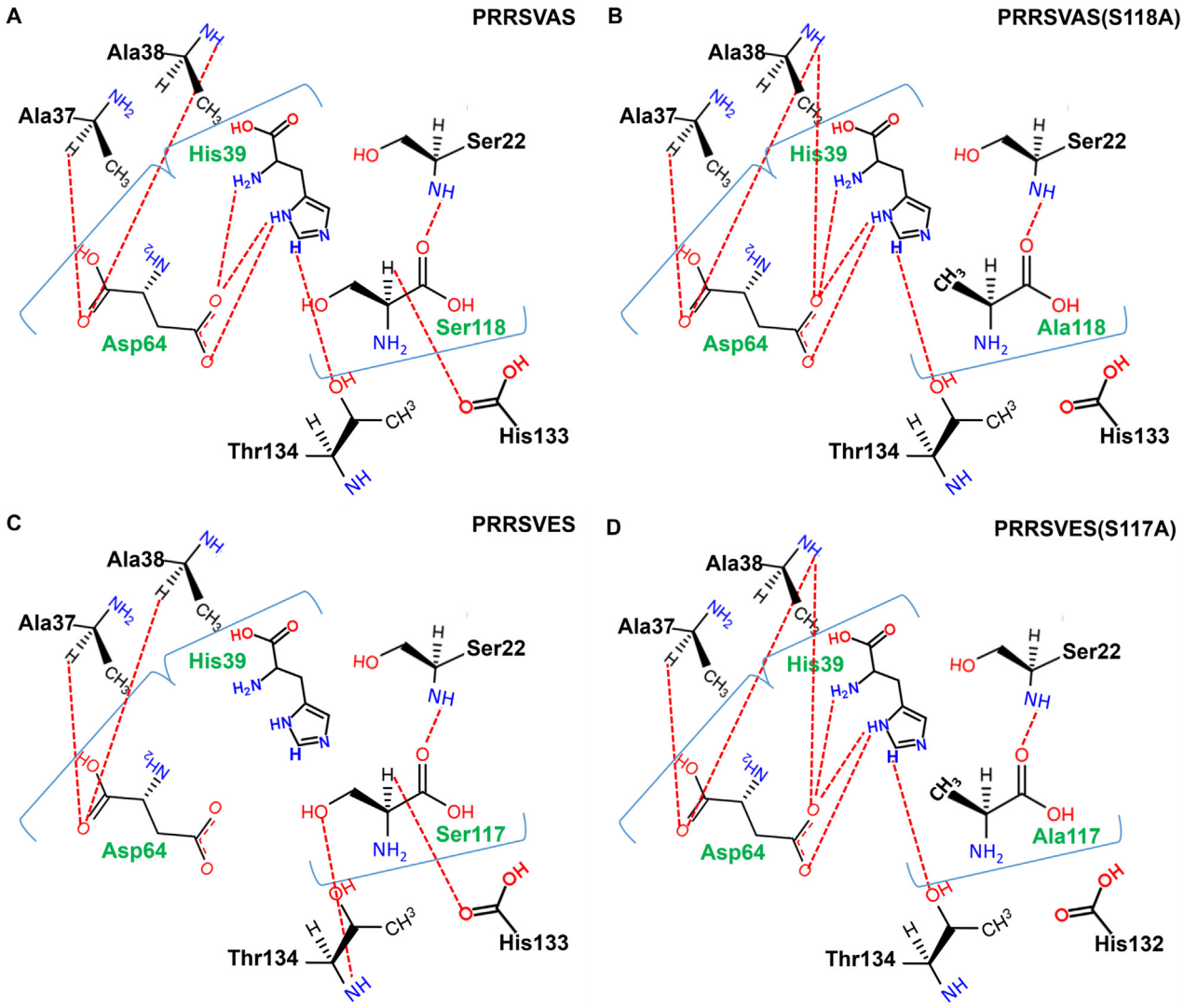

Fig. 7. The main Hydrogen bond interaction at the catalytic triad. PRRSVAS (A), PRRSVAS(S118A) (B), PRRSVES (C) and PRRSVES(S117A) (D).

at the pocket between domain I and II. EAV nsp4 is composed of His134 lying at the bottom of the pocket, Ser137 located at one side and Thr115 situated on the opposite side. We calculated the distance between C-alpha atoms of the catalytic triad as shown in Fig. 5 to elucidate the dynamic behavior as a function of time. The results demonstrated that the distances between His39 and Ser118 were very flexible. Moreover, the distance distribution profiles were normalized to confirm the dynamics behavior. The distance distribution frequency between His39 and Ser118 showed wider distribution than other distributions related to the distance results (Fig. 5).

Moreover, the conserved residues of the catalytic triad were mutated to evaluate the effect of the structure and activity of the protein and to understand how the mutations induce conformational changes in the catalytic triad. The selected distance distributions were calculated as the function of time and were shown in Fig. 6. From this result, the distances between His39 and Asp64 were very stable. In the PRRSVAS (S118A) mutation, the distance between S118HA:H133O and S118OG:G135H was lost while PRRVSEV(S117A) showed S118OG:G134H disappeared as well. Moreover, we observed in a Ramachandran plot of the wild type and mutant after simulations (Supplementary Fig. 5) that mutation caused the catalytic triad residues to shift a little.

3.5. Dynamics of hydrogen bond interactions

The evolution of the hydrogen bond interactions at the catalytic triad was measured over the last 50 ns of the simulation time to gain an insight into the main interaction at the pocket site. The amino acids at the catalytic triad are expected to maintain H-bond interaction around the site. Therefore, substitution of the amino acid at the active site results in changing H-bond interactions. In Table 3, we calculated occupancy percentage of the hydrogen bond interactions at the catalytic triad. Each interaction pair maintained at least a 10% interaction. We found that H39 showed a high occupancy percentage of the hydrogen bond interactions with D64, S118, and T134. D64 also had H-bond interactions with A37, A38, and T134, while S118 demonstrated interactions with G21, S22, H134, T134, and G135. More interestingly all residues are conserved in PRRSVAS, PRRSVES, LDVC, LDVP and EAV. Additionally, the results also revealed that mutation at the catalytic triad had an impact on changing H-bond interactions; S118 of the wild-type protein formed H-bonds with T134 and G135 but these interactions were lost in PRRSVAS (S118A) and PRRSVES (S117A). Therefore, the substitution of important H-bond interactions in PRRSVAS (S118A) and PRRSVES (S117A) had an effect on the flexibility around the catalytic triad, conformation and the proteolytic activity. The main

Table 3
Occupancy percentage of the hydrogen bond interactions at the catalytic triad.

Residue	Hydrogen bond	PRRSVAS	PRRSVAS(S118A)	PRRSVES	PRRSVES(S117A)
H39	H39-H···D64-OD1	26.9			24.12
	H39-H···D64-OD2	71.86	99.52		74.32
	H39-HD1···D64-OD1	99.8	99.96		99.84
	H39-HD1···D64-OD2	99.86	99.74		99.28
	H39-HD2···S117-OG			15.7	
	H39-HE1···T134-OG1	70.8	90.96		
	H39-NE2···S118-HG	27.52			
	H39-NE2···G134-H			12.8	
	H39-O···S18-HG			10.18	
H64	D64-O···A37-HA	99.56	99.16	94.44	99.92
	D64-O···A38-H	97.92	95.18		99.92
	D64-O···A38-HA			99.96	
	D64-OD1···A38-H	21.28			18.86
	D64-OD1···T133-HG1				74.64
	D64-OD1···T134-HG1		10.3		
	D64-OD2···A38-H	55.08	85.66		
	D64-OD2···T133-HG1				24.58
	D64-H···N62-O			21.64	
S118	S118-O···G21-HA1	43.14	45.62	69	37.36
	S118-O···G21-HA2	25.8	19.48	45.86	25.98
	S118-O···S22-H	99.96	99.96	100	99.98
	S118-O···S22-HB2	29.18	25.26		39.8
	S118-HA···H132-O			68.08	86.26
	S118-HA···H133-O	81.54	40.22		
	S118-HA···H134-OG1	14.58	54.72		
	S118-HB1···T133-OG1			30.4	
	S118-HB1···T134-OG1	12.04			
	S118-OG···G134-H			83.74	
	S118-OG···G135-H	63.94			
	A118-H···D117-OD1		41.34		
	A118-H···D117-OD2		33.94		

hydrogen bond interactions were represented in Fig. 7. It illustrates that the main binding interactions showed 70% occupancy of the hydrogen bond interactions.

4. Conclusions

Type I and type II nsp4 PRRSV play a vital role as the main proteinases. We have therefore performed molecular dynamics simulations (MD) and principle component analysis (PCA) to investigate the functional roles of the catalytic triad and conformational dynamics of type I and type II of nsp4 PRRSV. The results of MD simulations showed that the RMSF of residues 136–142 near the active site of all models were highly flexible. Moreover, we identified the effect of single mutations of the structure at the catalytic triad. The results revealed that the mutations increased the percentage of residues with a 0.1 nm RMSF value in domains I and II. The PCA also demonstrated that the wild types were more stable than the mutants. The distances between His39 and Ser118 were very flexible, while the distances between His39 and Asp64 were very stable. In PRRSVAS (S118A) mutation, the distance between A118HA:H133O and A118OG:G135H was lost while PRRSVEV (S117A) showed thst A118OG:G134H disappeared as well. In addition, at the catalytic triad H39 there was a high occupancy percentage of the hydrogen bond interaction with D64, S118, and T134. D64 also had H-bond interactions with A37, A38, and T134, while S118 demonstrated interactions with G21, S22, H134, T134, and G135. More interestingly all residues were conserved in PRRSVAS, PRRSVES, LDVC, LDVP and EAV. Moreover, S118 of wild-type protein formed H-bonds with T134 and G135 but these interactions were lost in PRRSVAS (S118A) and PRRSVES (S117A). Therefore, the substitution of important H-bond interactions in PRRSVAS (S118A) and PRRSVES (S117A) had an effect on the flexibility around the catalytic triad, conformation and the proteolytic activity. Overall, our research may provide the structural basis of the catalytic triad that cannot be ascertained experimentally, and be useful for testing protein activity in future experiments. Moreover, these findings may help us to understand the functional mechanisms of type I and type II nsp4 PRRSV and develop a new approach to control PRRS infection in the future.

Acknowledgments

This work was supported by Naresuan University (R2558C162), Faculty of Agriculture, Nature Resources and Environment, Naresuan University. We wish to express our gratitude for the use of high-performance computer cluster from The National Center for Genetic Engineering and Biotechnology (BIOTEC) and The National Science and Technology Development Agency (NSTDA).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jmgm.2017.03.015.

References

[1] D. Cavanagh, Nidovirales a new order comprising Coronaviridae and Arteriviridae, Arch. Virol. 142 (1997) 629–633.
[2] C.J. Nelsen, M.P. Murtaugh, K.S. Faaberg, Porcine reproductive and respiratory syndrome virus comparison: divergent evolution on two continents, J. Virol. 73 (1999) 270–280.
[3] E. Albina, Epidemiology of porcine reproductive and respiratory syndrome (PRRS): an overview, Vet. Microbiol. 55 (1997) 309–316.
[4] E.J. Snijder, M. Kikkert, Y. Fang, Arterivirus molecular biology and pathogenesis, J. Gen. Virol. 94 (2013) 2141–2163.
[5] S.H. Done, D.J. Paton, Porcine reproductive and respiratory syndrome: clinical disease, pathology and immunosuppression, Vet. Rec. 136 (1995) 32–35.
[6] W.H. Feng, S.M. Laster, M. Tompkins, T. Brown, J.S. Xu, C. Altier, et al., In utero infection by porcine reproductive and respiratory syndrome virus is sufficient to increase susceptibility of piglets to challenge by Streptococcus suis type II, J. Virol. 75 (2001) 4889–4895.
[7] E. Mateu, I. Diaz, The challenge of PRRS immunology, Vet. J. 177 (2008) 345–351.

[8] D.J. Paton, I.H. Brown, S. Edwards, G. Wensvoort, 'Blue ear' disease of pigs, Vet. Rec. 128 (1991) 617.
[9] H.S. Nielsen, G. Liu, J. Nielsen, M.B. Oleksiewicz, A. Botner, T. Storgaard, et al., Generation of an infectious clone of VR-2332, a highly virulent North American-type isolate of porcine reproductive and respiratory syndrome virus, J. Virol. 77 (2003) 3702–3711.
[10] G. Wensvoort, C. Terpstra, J.M. Pol, E.A. ter Laak, M. Bloemraad, E.P. de Kluyver, et al., Mystery swine disease in The Netherlands: the isolation of Lelystad virus, Vet. Q. 13 (1991) 121–130.
[11] D.A. Benfield, E. Nelson, J.E. Collins, L. Harris, S.M. Goyal, D. Robison, et al., Characterization of swine infertility and respiratory syndrome (SIRS) virus (isolate ATCC VR-2332), J. Vet. Diagn. Invest. 4 (1992) 127–133.
[12] S. Dea, R. Bilodeau, R. Athanassious, R.A. Sauvageau, G.P. Martineau, Quebec Isolation of the porcine reproductive and respiratory syndrome virus in Quebec, Can. Vet. J. 33 (1992) 552–553.
[13] A. Botner, J. Nielsen, V. Bille-Hansen, Isolation of porcine reproductive and respiratory syndrome (PRRS) virus in a Danish swine herd and experimental infection of pregnant gilts with the virus, Vet. Microbiol. 40 (1994) 351–360.
[14] H. Kuwahara, T. Nunoya, M. Tajima, A. Kato, T. Samejima, An outbreak of porcine reproductive and respiratory syndrome in Japan, J. Vet. Med. Sci. 56 (1994) 901–909.
[15] K. Tian, X. Yu, T. Zhao, Y. Feng, Z. Cao, C. Wang, et al., Emergence of fatal PRRSV variants: unparalleled outbreaks of atypical PRRS in China and molecular dissection of the unique hallmark, PLoS One 2 (2007) e526.
[16] U.U. Karniychuk, M. Geldhof, M. Vanhee, J. Van Doorsselaere, T.A. Saveleva, H.J. Nauwynck, Pathogenesis and antigenic characterization of a new East European subtype 3 porcine reproductive and respiratory syndrome virus isolate, BMC Vet. Res. 6 (2010) 30.
[17] C. Huang, Y.P. Du, Z.B. Yu, Q. Zhang, Y.H. Liu, J. Tang, et al., Highly pathogenic porcine reproductive and respiratory syndrome virus Nsp4 cleaves VISA to impair antiviral responses mediated by RIG-I-like receptors, Sci. Rep. (2016) 6.
[18] X. Tian, G. Lu, F. Gao, H. Peng, Y. Feng, G. Ma, et al., Structure and cleavage specificity of the chymotrypsin-like serine protease (3CLSP/nsp4) of porcine reproductive and respiratory syndrome virus (PRRSV), J. Mol. Biol. 392 (2009) 977–993.
[19] A.T. Xu, Y.J. Zhou, G.X. Li, H. Yu, L.P. Yan, G.Z. Tong, Characterization of the biochemical properties and identification of amino acids forming the catalytic center of 3C-like proteinase of porcine reproductive and respiratory syndrome virus, Biotechnol. Lett. 32 (2010) 1905–1910.
[20] Z.T. Ma, Y.L. Wang, H.Y. Zhao, A.T. Xu, Y.Q. Wang, J. Tang, et al., Porcine reproductive and respiratory syndrome virus nonstructural protein 4 induces apoptosis dependent on its 3C-like serine protease activity, PLoS One 8 (2013).
[21] Z. Ma, Y. Wang, H. Zhao, A.-T. Xu, Y. Wang, J. Tang, et al., Porcine reproductive and respiratory syndrome virus nonstructural protein 4 induces apoptosis dependent on its 3C-like serine protease activity, PLoS One 8 (2013) e69387.
[22] Z. Chen, M. Li, Q. He, J. Du, L. Zhou, X.N. Ge, et al., The amino acid at residue 155 in nonstructural protein 4 of porcine reproductive and respiratory syndrome virus contributes to its inhibitory effect for interferon-beta transcription in vitro, Virus Res. 189 (2014) 226–234.
[23] C. Huang, Q. Zhang, X.K. Guo, Z.B. Yu, A.T. Xu, J. Tang, et al., Porcine reproductive and respiratory syndrome virus nonstructural protein 4 antagonizes beta interferon expression by targeting the NF-kappaB essential modulator, J. Virol. 88 (2014) 10934–10945.
[24] M.F. de Abin, G. Spronk, M. Wagner, M. Fitzsimmons, J.E. Abrahante, M.P. Murtaugh, Comparative infection efficiency of Porcine reproductive and respiratory syndrome virus field isolates on MA104 cells and porcine alveolar macrophages, Can. J. Vet. Res. 73 (2009) 200–204.
[25] M.P. Murtaugh, T. Stadejek, J.E. Abrahante, T.T. Lam, F.C. Leung, The ever-expanding diversity of porcine reproductive and respiratory syndrome virus, Virus Res. 154 (2010) 18–30.
[26] P.G. Halbur, P.S. Paul, M.L. Frey, J. Landgraf, K. Eernisse, X.J. Meng, et al., Comparison of the pathogenicity of two US porcine reproductive and respiratory syndrome virus isolates with that of the Lelystad virus, Vet. Pathol. 32 (1995) 648–660.
[27] F.J. Martinez-Lobo, F. Diez-Fuertes, J. Segales, C. Garcia-Artiga, I. Simarro, J.M. Castro, et al., Comparative pathogenicity of type 1 and type 2 isolates of porcine reproductive and respiratory syndrome virus (PRRSV) in a young pig infection model, Vet. Microbiol. 154 (2011) 58–68.
[28] B. Ling, S. Bi, M. Sun, Z. Jing, X. Li, R. Zhang, Molecular dynamics simulations of mutated Mycobacterium tuberculosis L-alanine dehydrogenase to illuminate the role of key residues, J. Mol. Graph. Model. 50 (2014) 61–70.
[29] M.Y. Liu, L.S. Wang, X. Sun, X. Zhao, Investigating the impact of Asp181 point mutations on interactions between PTP1B and phosphotyrosine substrate, Sci. Rep. (2014) 4.
[30] N. Aeksiri, N. Songtawee, M.P. Gleeson, S. Hannongbua, K. Choowongkomon, Insight into HIV-1 reverse transcriptase–aptamer interaction from molecular dynamics simulations, J. Mol. Model. 20 (2014) 2380.
[31] V. Rajendran, C. Gopalakrishnan, R. Purohit, Impact of point mutation P29S in RAC1 on tumorigenesis, Tumor Biol. 37 (2016) 15293–15304.
[32] M. Biasini, S. Bienert, A. Waterhouse, K. Arnold, G. Studer, T. Schmidt, et al., SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information, Nucleic Acids Res. 42 (2014) W252–W258.
[33] P. Benkert, M. Kunzli, T. Schwede, QMEAN server for protein model quality estimation, Nucleic Acids Res. 37 (2009) W510–W514.
[34] D. Eisenberg, R. Lüthy, J.U. Bowie, 20] VERIFY3D: Assessment of protein models with three-dimensional profiles Methods in Enzymology, vol. 277, Academic Press, 1997, pp. 396–404.
[35] C. Colovos, T.O. Yeates, Verification of protein structures: patterns of nonbonded atomic interactions, Protein Sci. 2 (1993) 1511–1519.
[36] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, J. Appl. Crystallogr. 26 (1993) 283–291.
[37] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J. Berendsen, GROMACS: fast, flexible, and free, J. Comput. Chem. 26 (2005) 1701–1718.
[38] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, et al., GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX 1 (-2) (2015) 19–25.
[39] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J.L. Klepeis, R.O. Dror, et al., Improved side-chain torsion potentials for the Amber ff99SB protein force field, Proteins 78 (2010) 1950–1958.
[40] H.J.C. Berendsen, J.P.M. Postma, W.F.V. Gunsteren, A. DiNola, J.R. Haak, Molecular dynamics with coupling to an external bath, J. Chem. Phys. 81 (1984) 3684–3690.
[41] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling, The J. Chem. Phys. 126 (2007) 014101.
[42] S. Nosé, M.L. Klein, Constant pressure molecular dynamics for molecular systems, Mol. Phys. 50 (1983) 1055–1076.
[43] M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: a new molecular dynamics method, J. Appl. Phys. 52 (1981) 7182–7190.
[44] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraaije, LINCS: a linear constraint solver for molecular simulations, J. Comput. Chem. 18 (1997) 1463–1472.
[45] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems, J. Chem. Phys. 98 (1993) 10089–10092.
[46] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, A smooth particle mesh Ewald method, J. Chem. Phys. 103 (1995) 8577–8593.
[47] P. Turner, XMGRACE, Version 5.1.19, Center For Coastal and Land-Margin Research, Oregon Graduate Institute of Science and Technology, Beaverton, Ore, USA, 2005.
[48] A. Amadei, A.B.M. Linssen, H.J.C. Berendsen, Essential dynamics of proteins, Proteins 17 (1993) 412–425.