Molecular Dynamic Studies of Interferon and Innate Immunity Resistance in MERS CoV Non-Structural Protein 3

Manal Alfuwaires, Abdallah Altaher, and Mahmoud Kandeel*

Department of Biology, Faculty of Science, King Faisal University; Alhofuf, 31982 Alahsa, Saudi Arabia:

Department of Physiology, Biochemistry and Pharmacology, Faculty of Veterinary Medicine, King Faisal University; Alhofuf, 31982 Alahsa, Saudi Arabia; and Department of Pharmacology, Faculty of Veterinary Medicine, Kafrelshikh University; Kafrelshikh 33516, Egypt.

Received November 7, 2016; accepted December 11, 2016

The new emerging Middle East Respiratory Syndrome Coronavirus (MERS CoV) encodes several resistance proteins against the innate immune response of the host, including interferon (IFN) resistance. Monitoring of the status of such proteins will be important to track viral pathogenicity. In this study, molecular dynamics approaches were used to investigate MERS CoV Non-Structural Protein 3 (NSP3) specific proteins that resist host innate immunity. MERS CoV papain-like protease (Plpro) was more conformationally flexible than Severe Acute Respiratory Syndrome CoV (SARS) CoV Plpro. This flexibility was evident in either the free form or when bound with ubiquitin. There were marked changes in the root-mean-square deviation (RMSD) in the ubiquitin like domain (Ubl) and the fingers subdomain of the catalytic domain of Plpro. An interesting feature is the dynamic change in Ubl, which shows a rigid conformation in the free form of Plpro but is fully flexible upon the binding of ubiquitin. This increased flexibility could be important for the downstream effects of the interaction with other proteins and the inhibition of the innate immune. Four major residues involved in deubiquitination, L106, P163, R168 and F265, were conserved in all MERS CoVs and differed from other Beta CoVs. These conserved CoV residues were associated with lower deubiquitinating activity and render MERS CoV Plpro with less potent deubiquitinating potential. The number of residues and total interactions with ubiquitin were lower for the MERS CoV Plpro than for the SARS CoV. These factors contribute to the lower deubiquitinating actions of MERS CoV NSP3 and its subsequently lower interaction with the host immune system.

Key words Coronavirus; molecular dynamics; papain like protease; innate immunity; molecular modeling

Coronavirus (CoV) was first discovered in Saudi Arabia in 2012 in a patient hospitalized with severe respiratory distress followed by pneumonia and renal failure. Early work indicated the relationship of the new virus with the Beta lineage of CoV. The study group of the CoV later termed this new infection Middle East Respiratory Syndrome Coronavirus (MERS CoV). Soon after its initial discovery, the virus was discovered in several countries, including the U.S.A., U.K., France, Jordan, Tunisia and other countries. Recently, MERS CoV took a dangerous turn by causing a serious outbreak in Korea and a marked number of cases in China. The case fatality rate of MERS CoV exceeds that of the known Severe Acute Respiratory Syndrome CoV (SARS) CoV. This might be due to the severe respiratory and renal failure found during MERS CoV infection. Both SARS and MERS CoVs are similar in many structural and epidemiological aspects. While SARS CoV infection is more acute in nature, and the number of cases decreased after the first episode, MERS CoV has become more endemic in the Arabian Peninsula, with the continuous discovery of new cases over several years from 2012 until the present.

The CoV genome is approximately 30 kilobases and is composed of 10 open reading frames (ORFs). At the 5’ end of the genome, there are two large ORFs, ORF1a and ORF1b, which are translated to polyprotein 1a and polyprotein 1b (in MERS CoV, one large polyprotein, AB, is expressed due to a frameshift mutation). The other 8 ORFs translate to structural and non-structural proteins. There are 4 structural proteins, including the spike protein (S), which directly follows ORF1b, and the other 3 proteins are encoded at the end of the genome and include the envelope (E), membrane (M) and nucleocapsid (N) proteins. Between the E and S proteins, there are 4 nonstructural proteins, including ORF3, ORF4a, ORF4b, and ORF5 (Fig. 1).

The MERS CoV polyprotein AB is processed by two virus-encoded proteases, a 3-C-like protease (3CLpro, main protease) and a papain-like protease (Plpro). Both MERS and SARS CoVs encode only one Plpro, while in some other CoVs, multiple copies are present. The MERS CoV Non-Structural Protein 3 (NSP3) is composed of several proteins, including the Plpro, Ubiquitin like (Ubl), ADP ribose (Rib) phosphatase and transmembrane domains. The main protease (3CLpro) cleaves the last 11 sites, while Plpro cleaves the N-terminal three sites, releasing 16 different nonstructural proteins of MERS CoV. The main protease of MERS CoV, 3CLpro, has strictly protease activity. In contrast, Plpro is a multifunctional protein that shows protease, deubiquitination, de-ISGylation and anti-host innate immunity activities. MERS CoV Plpro is composed of 2 independent domains, a C-terminal palm-shaped protease composed of the thumb, palm and finger subdomains, and the Ubiquitin like domain (Ubl), which comprises the N-terminal 62 residues. These two domains are functionally distinct, with various downstream activities and molecular interactions. The Ubl domain is conserved in most CoVs including MERS CoV and SARS CoV, while the transmissible gastroenteritis CoV lacks the Ubl domain.

In studies related to the SARS CoV, Ubl was found to interfere with the IRF3 and nuclear factor-kappaB (NF-κB) antiviral...
Therefore, Ubl is a viral protein that shares in antagonizing the host innate immune response.

Molecular modeling and molecular dynamics (MD) simulations are widely used techniques to understand the molecular motion, conformational changes and other molecular changes during protein–protein interactions. In this work, we performed MD analysis of the interaction of NSP3 Plpro with ubiquitin and Ubl among important viruses that pose serious health risks, such as SARS and MERS CoV. Investigation of the detailed interaction at the protein–protein interface will contribute to our knowledge about the interaction of viral proteins with host innate immunity proteins.

METHODS

Preparation of Structure Complexes Several crystal structures were obtained from the Protein Data Bank (PDB) database. The structures included SARS or MERS Plpro either in its free form or bound with ubiquitin (Table 1). After browsing of the PDB database, Plpro from the coronaviridae were only evident with SARS and MERS CoV and Infectious Bronchitis virus (IBV). 2X2Z, the structure of the IB virus Plpro, was also used for comparison with SARS and MERS CoV Plpro. The structures were inspected, and the missing lobes were modeled.

MD Simulation MD simulation was implemented for 10 ns in YASARA (version 14.12.2) using an AMBER99 force field. Simulations were applied to different crystal structures, including MERS CoV, SARS CoV and IB virus (Table 1). These structures represent the available coronaviridae structures in the database. For all of the used PDB files, the same protocol was repeated. In the first simulation, a cell was constructed with a real space cut-off of 7.9 Å. The pK_a values of ionizable groups were predicted, and protonation states were assigned at a default pH of 7.2. The simulated cell was filled with water to a density of 0.997 g/L. The initial structures were optimized by short steepest descent minimization to remove bumps and correct covalent geometry errors, followed by 500 steps of simulated annealing at 298 K as an equilibrium phase. Atom velocities were rescaled every 25 simulation steps to maintain the time average temperature at 298 K. The solvated structures were finally subjected to MD simulation for 10 ns. The final snapshots were selected from the lowest energy snapshots, which were saved every 250 ps. The root-mean-square deviation (RMSD) was calculated after least squares fitting using Cα atoms.

Sequence Alignment and Molecular Modeling The sequences of Plpro from SARS CoV, MERS CoV and IBV were imported from the gene databases by Geneious software ver. 7. The sequences were aligned by the ClustalW module using the BLOSUM cost matrix, the gap open cost was set to 10 and the gap extension cost was set to 0.1. The Ubl domain and fingers subdomain (selected after MD simulation results) were directly extracted from the alignment output file and realigned, and homology parameters were calculated based on rates of similarities or differences (Figs. 2A–E). The crystal structures of Plpros were visualized and annotated, and figures were generated by CLC drug discovery or Molsoft ICM browser software.

Analysis of the Protein–Protein Interaction of Plpro with Ubiquitin The protein complexes were submitted to the bioCOMplexes COntact MAPS server (https://www.molnac.unisa.it/BioTools/cocomaps/index.psp) using the default parameters. The output parameters included Plpro–
ubiquitin interface analysis, interactions map, distance range analysis, intermolecular hydrogen bonds and accessible surface area parameters (Table 2).

**Table 2. Comparative Interactions of SARS CoV and MERS CoV Plpro with Ubiquitin**

|                      | 4MM3 | 4WUR |
|----------------------|------|------|
| Number of interacting residues in ubiquitin | 40   | 30   |
| Number of interacting residues in Plpro     | 68   | 67   |
| Number of hydrophilic–hydrophobic interaction | 194  | 175  |
| Number of hydrophobic–hydrophobic interaction | 66   | 53   |
| Buried area upon the complex formation (Å²) | 1881.7 | 1634.6 |
| Buried area upon the complex formation (%)   | 9.12 | 8.2  |
| Interface area (Å²)                          | 940.85 | 817.3 |
| Interface area MOL1 (%)                      | 20.03 | 5.41 |

Genomics Analysis of NSP3-Specific Anti-innate Immunity  
Full genome sequences of the MERS CoV were imported into the CLC genomics workbench. The sequence of the MERS CoV Plpro, including Ubl, was used to blast either the local database of MERS CoV genomes or shorter sequences at the NCBI database. The markers for innate immunity resistance were compared among different MERS CoV and other coronavirus sequences. These markers were based on previous experimental evidence of the role of some residues or subdomains and domains in resistance to innate immunity or from our MD simulation results. Attention was directed to 1) the Ubl domain, 2) the fingers subdomain of the catalytic domain, and 3) the important residues for deubiquitinating activity, such as L106, P163, E168 and Y265.

**RESULTS AND DISCUSSION**

The sequences of MERS CoV, SARS CoV and IBV Plpros were aligned by the ClustalW method (Figs. 2A–E). Comparison of the amino acid composition of the three Plpros was dependent on the percent identity matrix. The % identity among all Plpros was less than 30% (Fig. 2C). Comparison of the MERS CoV Plpro with the SARS CoV Plpro showed 29.9% identity, which is higher than the 22.3% identity with IBV. During MD simulations, marked differences in RMSD residues were observed in two segments of Plpro. The first segment is the N-terminal Ubl domain, and the second segment
is the fingers subdomain of the catalytic domain of MERS CoV Plpro (highlighted in Fig. 2F). In agreement with the previous result, the fingers subdomain of Plpro showed low % identity among the examined Plpros. In contrast, the CoV Ubl domain showed the lowest % identity among the examined Plpros, which was in the range of 15.8–28.8%. The low identity among Plpros and its domains and subdomain components indicates subtle differences in the corresponding biological activities and interactions with other proteins. We compared the sequences of MERS CoV and SARS CoV Plpros. There was a putative weaker deubiquitinating activity of MERS CoV. In the catalytic domains of MERS CoV Plpro, leucine, proline, arginine and phenylalanine were present at position numbers 106, 163, 168 and 256, respectively. In contrast, SARS CoV has tryptophan, leucine, glutamate, and tyrosine in these respective sites. In this context, the L106W mutation of the

Fig. 3. Time Dependence of RMSD for Plpros from MERS CoV, SARS CoV and IBV (A)

MD simulation was run to 10ns and snapshots were taken every 250ps. The changes in RMSD per residue are represented in (B). Time dependence of RMSD for Plpros from MERS CoV and SARS CoV bound with ubiquitin (C). The changes in RMSD per residue in Plpro bound with ubiquitin (D). Comparison of the free form of Plpro in (C) with ubiquitin-bound Plpro reveals changes in the N-terminal (first 50 residues) Ubl domain and the fingers subdomain (residues 180–240).
MERS CoV Plpro resulted in a drastic increase in deubiquitinating activity. In addition, the mutation of SARS Plpro to residues similar to MERS CoV amino acids resulted in the profound loss of deubiquitinating activity. Therefore, MERS CoV Plpro naturally has lower deubiquitinating activity compared with SARS CoV Plpro.

The RMSD of the Ca atoms during MD simulation relative to the original structure is represented in Fig. 3. In the first 100 ps, there was a prompt increase in RMSD values due to the relaxation of the dissolved structure, and there were more changes in comparison with the original structure. It is also noticeable that the magnitude of changes in RMSD was higher in the MERS CoV Plpro than in the SARS or IBV Plpros (Figs. 3A, C). The RMSD average of MERS CoV was 2.5 Å compared with 1.5 Å in the SARS CoV and IBV Plpros (Fig. 3A). The dynamic behavior thus confirms that the SARS and IBV Plpros are more conformationally rigid, as indicated by the RMSD values. Analysis of the changes in individual amino acids of Apo Plpro reveals more generalized, higher RMSD in the structures of MERS Plpro (Fig. 3B), with a clear, high peak in the maximum changes in 2 segments of the MERS CoV Plpro. The range of the greatest changes in RMSD are A181-C195 and L218-C232. These 2 segments represent 3 of the 4 beta sheets of the fingers subdomain of the catalytic domain of the MERS CoV Plpro. These two segments are not in direct contact with ubiquitin, except for one residue, N227. The N-terminal region comprising the Ubl domain was almost stable and showed little or no change among the SARS, MERS and IBV structures (Figs. 2B, 3A). This may indicate that the conserved features and functions of the Ubl domain among all coronaviruses are rigid, with low rates of conformational change in the absence of ubiquitin.

The MD simulation of the CoVs Plpro bound with ubiquitin is shown in Figs. 3C and 4D. RMSD changes were more evident in the Ubl domain and fingers subdomain (Fig. 3D). Comparison of the free form of Plpro with those structures bound with ubiquitin reveals interesting features that are sum-
marized as follows: 1) The SARS Plpro was more rapidly stabilized than the MERS CoV Plpro (Fig. 3A). The SARS Plpro stabilized at 300 ps, while the MERS CoV Plpro stabilized after 1.2 ns. 2) After stabilization, the average RMSD values of SARS CoV Plpro and MERS CoV Plpro were 2 and 2.6, respectively. This indicates that more structural rearrangements are expected with the MERS CoV Plpro. 3) There were marked changes in the RMSD of the N-terminal Ubl domain, especially in the Apo MERS CoV Plpro (Figs. 4A, B). The Ubl domain therefore changes from a rigid to a flexible state in the presence of ubiquitin. Ubl is thought to share in protein–protein interactions and in the regulation of downstream interactions, especially in the modulation of the host innate immune response of the host.

Analysis of residues at the Plpro–ubiquitin interface was performed by submission of the corresponding structures to the bioCOmplexes COn tact MAPS server. The obtained results were imported into an Excel file and inspected for interface residues, energy factors and other analytical data. In the MERS CoV Plpro, there were 9 hydrogen bonds between the Plpro and ubiquitin, which was lower than the number in the SARS CoV Plpro complex (10 hydrogen bonds). In addition, the MERS CoV Plpro–ubiquitin interaction showed a lower number of hydrophobic interactions and total number of interacting residues (Table 2).

A representation of the MERS CoV Plpro–ubiquitin interface is shown in Fig. 5 and Table 2. The interface residues are distributed within the palm and finger subdomains of the catalytic domain of Plpro. MD analysis of the interface residues in both the MERS CoV and SARS CoV Plpros revealed variable degrees of changes among different residues (Fig. 6). While most of the residues showed RMSD values of approximately 1 Å, some residues showed marked changes. In the MERS CoV Plpro, Q227 was the only residue in the fingers subdomain that interacted with ubiquitin and showed marked increases in RMSD approaching 4.6 Å (Figs. 5, 6). In contrast, 5 residues in the SARS CoV Plpro fingers subdomain interacted with ubiquitin (Figs. 5, 6). These residues showed a moderate increase in RMSD up to 2 Å (Fig. 6). Therefore, the MERS CoV Plpro fingers subdomain interacts with ubiquitin with 1 residue, which showed drastic changes in RMSD. In comparison, the SARS CoV Plpro interacts with 5 residues with ubiquitin, with a moderate increase in RMSD during MD simulation.

Comparison of the Ubl domain from 300 different full genome sequences of the coronavirus indicates restricted conservation of Ubl among MERS CoV genomes and marked differences within the coronavirus. The matrix of pairwise comparison reveals that only 1 residue differs within the Ubl of MERS CoV genomes, constituting 98.4–100% of the identity. Comparison with other Ubl domains from SARS CoV or CoVs isolated from bats reveals an identity decreased by ca. 40%. Similarly, the fingers subdomain showed 0–1 differences among MERS CoV genomes, while the number of differences increased to 22 residues in comparison with other Beta CoVs (62%). For residues with proven importance in Plpro activity, among 300 full genomes of CoVs, L106 was conserved in all MERS CoV and bat CoVs, and R168 was conserved in all MERS CoV Plpros and was replaced with lysine or methionine in other beta CoVs. Similarly, the conserved residue F265 was replaced with tyrosine in the other viruses.

In conclusion, both previous experimental work and the present modeling studies revealed low deubiquitinating activity of MERS CoV. This is supported by the mutation of essential residues for Plpro actions. MERS CoV Plpro showed lower surface area of interaction and lower number of interface residues. Being more flexible and movable, the Ubl domain indicates different interaction styles with other host immune proteins that require more experimental evidence. Ubiquitin induces noticeable changes in the entire structure of the MERS CoV Plpro, especially at the Ubl domain. This finding indicates that the changes in Ubl domain is affected by the presence of ubiquitin.

Acknowledgment This work was supported by a research grant from King Faisal University, Deanship of Scientific Research under Grant number 160046.

Conflict of Interest The authors declare no conflict of interest.

REFERENCES

1) Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus AD, Fouchier RA. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. N. Engl. J. Med., 367, 1814–1820 (2012).

2) WHO MERS-Cov Research Group. State of knowledge and data gaps of Middle East respiratory syndrome coronavirus (MERS-
3) Su S, Wong G, Liu Y, Gao GF, Li S, Bi Y. MERS in South Korea and China: a potential outbreak threat? *Lancet*, **385**, 2349–2350 (2015).

4) Cowling BJ, Park M, Fang VJ, Wu P, Leung GM, Wu JT. Preliminary epidemiologic assessment of MERS-CoV outbreak in South Korea, May–June 2015. *Euro surveillance: bulletin European sur les maladies transmissibles=European communicable disease bulletin*, **25**, 7–13 (2015).

5) Hajjar SA, Memish ZA, McIntosh K. Middle East respiratory syndrome coronavirus (MERS-CoV): a perpetual challenge. *Ann. Saudi Med.*, **33**, 427–436 (2013).

6) Harcourt BH, Jukneliene D, Kanjanahaluethai A, Bechill J, Severson KM, Smith CM, Rota PA, Baker SC. Identification of severe acute respiratory syndrome coronavirus replicase products and characterization of papain-like protease activity. *J. Virol.*, **78**, 13600–13612 (2004).

7) Mielech AM, Kilianski A, Baez-Santos YM, Mesecar AD, Baker SC. MERS-CoV papain-like protease has delSGylating and deubiquitinating activities. *Virology*, **450–451**, 64–70 (2014).

8) Baez-Santos YM, Mielech AM, Deng X, Baker S, Mesecar AD. Catalytic function and substrate specificity of the papain-like protease domain of nsp3 from the Middle East respiratory syndrome coronavirus. *J. Virol.*, **88**, 12511–12527 (2014).

9) Bailey-Elkin BA, Knaap RC, Johnson GG, Dalebout TJ, Nihaner DK, van Kasteren PB, Bredenbeek PJ, Snijder EJ, Kikkert M, Mark BL. Crystal structure of the Middle East respiratory syndrome coronavirus (MERS-CoV) papain-like protease bound to ubiquitin facilitates targeted disruption of deubiquitinating activity to demonstrate its role in innate immune suppression. *J. Biol. Chem.*, **289**, 34667–34682 (2014).

10) Lei J, Mesters JR, Drosten C, Anemüller S, Ma Q, Hilgenfeld R. Crystal structure of the papain-like protease of MERS coronavirus reveals unusual, potentially druggable active-site features. *Antiviral Res.*, **109**, 72–82 (2014).

11) Yang X, Chen X, Bian G, Tu J, Xing Y, Wang Y, Chen Z. Proteolytic processing, deubiquitase and interferon antagonist activities of Middle East respiratory syndrome coronavirus papain-like protease. *J. Gen. Virol.*, **95**, 614–626 (2014).

12) Rattia K, Kilianski A, Baez-Santos Y, Baker S, Mesecar A, Rey FA. Structural Basis for the Ubiquitin-Linkage Specificity and delSGylating Activity of SARS. *PLoS Pathogens*, **10.5** (2014).

13) Kandeel M, Ando Y, Kitamura Y, Abdel-Aziz M, Kitade Y. Mutational, inhibitory and microcalorimetric analyses of Plasmodium falciparum TMP kinase. Implications for drug discovery. *Parasitology*, **136**, 11–25 (2009).

14) Kandeel M, Altaher Y, Abdelaziz M, Alnazawi M, Elshazli K. Evolution of camel CYP2E1 and its associated power of binding toxic industrial chemicals and drugs. *Comput. Biol. Chem.*, **64**, 271–280 (2016).

15) Altaher Y, Kandeel M. Molecular analysis of some camel cytochrome P450 enzymes reveals lower evolution and drug-binding properties. *J. Biomol. Struct. Dyn.*, **34**, 115–124 (2016).

16) Altaher Y, Nakanishi M, Kandeel M. Annotation of camel genome for estimation of drug binding power, evolution and adaption of cytochrome P450 enzymes. *International Journal of Pharmacology*, **11**, 243–247 (2015).

17) Kandeel M, Elaiziz M, Kandeel A, Altaher A, Kitade Y. Association of host tropism of Middle East syndrome coronavirus with the amino acid structure of host cell receptor dipeptidyl peptidase 4. *Acta Virol.*, **58**, 359–363 (2014).

18) Krieger E, Kortauhann G, Vriend G. Increasing the precision of comparative models with YASARA NOVA—a self-parameterizing force field. *Proteins*, **47**, 393–402 (2002).

19) Vangone A, Spinelli R, Scarano V, Cavallo L, Oliva R. COCOMAPS: a web application to analyze and visualize contacts at the interface of biomolecular complexes. *Bioinformatics*, **27**, 2915–2916 (2011).

20) Chou C-Y, Lai H-Y, Chen H-Y, Cheng S-C, Cheng K-W, Chou Y-W. Structural basis for catalysis and ubiquitin recognition by the severe acute respiratory syndrome coronavirus papain-like protease. *Acta Crystallogr. D Biol. Crystallogr.*, **70**, 572–581 (2014).