Mechanism of a Mutation in Non-Structural Protein 1 Including High Pathogenicity of Avian Influenza Virus H5N1

Running Title: NS1 Mechanism of High Pathogenicity of H5N1

Yusuke S. Kato *a,b, Kiyoshi Fukui a and Kazuo Suzuki #c

*a The Institute for Enzyme Research, Tokushima University, Tokushima, Japan; b Institute for Health Sciences, Tokushima Bunri University, Tokushima, Japan; c Asia International Institute of Infectious Disease Control, and Department of Health Protection, Graduate School of Medicine, Teikyo University, Itabashi-ku, Japan

Abstract: Avian influenza H5N1 has shown high mortality rate in human. Non-structural protein 1 (NS1) is a virulence factor of H5N1. Mutation at the 42nd residue within the RNA-binding domain (RBD) of NS1 dramatically changes the degree of pathogenicity of H5N1 in mice. We here studied the impact of this mutation on the function of RBD, and found that RBD with serine at the 42th residue binds double-stranded RNA (dsRNA), whereas that with proline at the 42th residue does not. Analysis of structural models of the RBD proteins with S42 and P42 suggested remarkable difference in the structure of the dsRNA-binding interface, whereas structural analysis by analytical gel filtration and CD measurements did not indicate difference between those RBD proteins. Our results suggest that the single amino acid replacement induces a minor, but global structural change leading to the loss of function of NS1 thereby the change in the degree of pathogenicity.

Keywords: avian Influenza, non-structural protein 1, structure modeling, cytokine storm, double-stranded RNA, interaction

1. INTRODUCTION

Since the outbreak of highly pathogenic avian H5N1 influenza virus infection in humans in 2003, the World Health Organization (WHO) has reported a mortality rate of about 60% [1]. It is remarkable that most of the victims with high pathogenic influenza die within two weeks with rapid progress of pathology. In the case of the influenza pandemic that occurred in 2009, some patients infected with influenza showed severe acute respiratory distress syndrome (ARDS). It is presumed that the direct cause of death is respiratory failure caused by severe ARDS. The final pathological stage of severe ARDS by highly pathogenic influenza infection shows diffuse alveolar damage. In addition, several studies have reported evidences of cytokine storms in specimen of the patients with ARDS caused by H5N1 infection [4-6]. Thus, intense inflammatory responses with hypersecretion of cytokines are directly linked to the severe ARDS and subsequent death observed in patients infected with influenza virus H5N1.

The genome of influenza virus is composed of eight single-stranded RNA segments, so that the RNA segments can be exchanged between different strains of influenza virus in a host cell infected by multiple strains. A novel strain of influenza can be easily created by such mechanism. Through this exchanging mechanism, it is possible that a high pathogenic gene from avian influenza virus (H5N1) is transferred to mammalian influenza viruses.

One of proteins from the viral genes expressed in a host, NS1, is implicated in the virulence and determination of host range [7-9]. NS1 suppresses the dsRNA-mediated interferon response to attenuate the innate immune system [10, 11]. On the other hand, it has been reported that NS1 induces secretion of huge amount of cytokines from A549 cells, which are human alveolar basal epithelial cell line, under oxidative stress [12]. Thus, NS1 may be a viral factor to induce cytokine storms in the patients infected by H5N1.

A/Duck/Guangxi/27/03 (DK/27), a subtype of H5N1 virus, is reported to cause highly pathogenic infection to mammals, and contains a Ser residue at the 42nd residue of NS1 (S42 type) [13]. By contrast, another subtype, A/Duck/Guangxi/12/03 (DK/12), contains NS1 with a Pro residue at the 42nd residue (P42 type), and exhibits lower pathogenicity than DK/27. This mutation causes difference in the degree of suppression of the dsRNA-mediated interferon response between these strains. These together suggest that NS1 is a virulence factor for high pathogenicity of the H5N1 strain.

A polypeptide of NS1 is composed of two structural domains such as the RNA-binding domain (RBD) and the effector domain (ED) [14]. The crystal structure of the complex of RBD of the S42 type and dsRNA [15] shows that...
S42 directly interacts with dsRNA through a hydrogen bond. In addition, the side chains of the R35 and R38 residues in the complex directly interact with dsRNA. Replacement of R35, R38 as well as R46 has negative impacts on the binding to dsRNA [16, 17]. Some of these basic residues should interact with dsRNA through electrostatic interactions in addition to hydrogen bonding. Thus, those residues are likely to interact with dsRNA more tightly than S42 [15]. However, replacement of the 42nd residue surprisingly has a greater impact on the degree of pathogenicity of H5N1 than that of R38 [13]. The mechanism underlying this observation is enigmatic.

Here we analyze differences in RBD between the S42 and P42 types by means of biochemistry and molecular modeling in order to elucidate the molecular mechanism.

2. MATERIALS AND METHOD

2.1. Plasmid construction and purification of RBD

RBD(S) from PR8 was subcloned into pGEX-4T1 (GE Healthcare, Little Chalfont, UK) with a TEV protease cleavage site instead of a thrombin site [18]. The 42nd residue of RBD(S) was replaced with a Pro residue by site-directed mutagenesis to produce RBD(P). Protein expression experiments were carried out using Rosetta 2 (DE3) (Merck Millipore, Billerica, MA). After centrifugation of disrupted E. coli cells by sonication, soluble extracts were loaded onto Glutathione Sepharose 4B equilibrated with 20 mM Tris/HCl pH 7.7, 150 mM NaCl, 1 mM EDTA and 1 mM DTT. GST-fusion proteins were digested in the column by AcTEV protease (Life Technologies, Waltham, MA). After 0.3% polyethylenimine treatment, sample solutions were loaded onto Source S 4.6/100 PE (GE Healthcare).

2.2. Analytical gel filtration experiments

Gel filtration experiments were performed using an ÄKTA FPLC (GE Healthcare) and a Superdex 75 10/300 column (GE Healthcare) equilibrated with 50 mM Na-Pi, 100 mM NaCl, ~50 μM RBD proteins were loaded. Gel Filtration Calibration Kit LMW (GE Healthcare) was used for the calibration.

2.3. CD measurements

CD spectra were recorded on a J-820 spectropolarimeter (Jasco, Tokyo, Japan) at room temperature. The solution compositions were ~5 μM RBD, 100 mM NaCl and 50 mM Na-Pi buffer (pH 7.0).

2.4. EMSA

16nt dsRNA was designed following Chien et al. [19] and purchased from Sigma-Aldrich (St. Louis, MO). The RBD dimers (1.0 μM) and dsRNA (2.5 μM) were incubated on ice in 8% glycerol, 2 unit/µl RNasin, 50 mM Tris/HCl pH 8.0, 1 mM DTT, 100 mM NaCl and 0.1 mg/ml heparin for 20 minutes. 6x dsRNA loading buffer (BioDynamics Laboratory Inc., Tokyo, Japan) was added to the mixture to load onto a 15% polyacrylamide gel. Electrophoresis was carried out at 4 ºC in 44.5 mM Tris/Borate pH 8.3 and 1 mM EDTA.

2.5. Modeling Structures

Modeling of monomers of RBD(P) and RBD(S) was carried out with I-TASSER [20]. Dimerization of RBD(P) and RBD(S) was carried out with MZDOCK [21]. Docking between dsRNA and the dimer of either RBD(S) or RBD(P) was carried out with ZDOCK [22]. Energy minimization was carried out using Swiss PDB Viewer [23]. Buried surface
area between the RBD proteins and dsRNA was calculated with PRinse [24]. Evaluation of the model structures was carried out with Procheck, Verify3D and ERRAT [25] [26] [27].

3. RESULTS

3.1. Sample preparation of the RBD proteins

To compare characteristics of RBD with P42 (RBD(P)) and that with S42 (RBD(S)), we first produced the recombinant proteins. Sequences of RBDs are listed in Fig. 1A. Both RBD(P) and RBD(S) were expressed stably as soluble fraction and successfully purified (Fig. 1B). Behavior of these proteins in the course of purification was stable and showed no difference. This implies that both RBD(P) and RBD(S) possess stable folds in aqueous solution.

3.2. Biochemical structural analysis of the RBD proteins

Amino acid substitution in a protein can occasionally cause a change in the state of multimerization or formation of aggregation, leading to a change in the function. To examine the possibility of such a change between RBD(P) and RBD(S), we analyzed molecular weights of the RBDs in aqueous solution by using analytical gel filtration chromatography (Fig. 1C). Elution volumes of these proteins were almost the identical so that this indicates that molecular weights of RBD(P) and RBD(S) are almost the identical. Results of calibration of the gel filtration column suggested that molecular weights of RBD(P) and RBD(S) in aqueous solution are 19.4 ± 0.2 and 19.4 ± 0.1 kDa, respectively. Because the molecular weight of a monomer of RBD calculated based on the amino acid sequence is ~8.5 kDa, both RBD(P) and RBD(S) are likely to be dimers in aqueous solution.

We measured CD spectra to investigate whether there is difference in the secondary structure contents of RBD(P) and RBD(S), and observed marginal difference between the two (Fig. 1D). It was suggested that both RBD(P) and RBD(S) are rich in α-helices, because each of the spectra shows two negative peaks at 208 and 222 nm. Thus, RBD(P) and RBD(S) have similar contents of the secondary structure.

We did not observe remarkable structural difference between RBD(P) and RBD(S) from the results of gel filtration chromatography and CD spectra. We next investigated whether there is difference or not in the function.

3.3. Binding analysis to dsRNA

In order to examine the binding ability of the RBD proteins to dsRNA, we carried out EMSA and found a shift of a band in the lane of RBD(S) but did not in the lane of RBD(P) (Fig. 2). These results indicated that RBD with Ser at the position 42 bound dsRNA, whereas that with Pro did not. In the complex structure between an S42 type of RBD and dsRNA (PDB code: 2ZKO), the hydroxyl group of S42 binds to a ribosyl group of dsRNA [15]. Thus, S42 of RBD(S) from our construct was likely to bind to a ribosyl group of dsRNA. RBD(P) should not be able to form a corresponding hydrogen bond because RBD(P) replaces this Ser with Pro. RBD(P) would therefore lose at least one hydrogen bond compared with RBD(S).

RBD has basic residues such as R35 and R38 on the dsRNA-binding interface [15]. Replacement of these residues has negative impacts on the binding to dsRNA [16, 17]. R35 and R38 are conserved in both RBD(S) and RBD(P). These Arg residues should play more important roles in the binding to dsRNA than S42 because of the charges of these residues, so we had expected that the binding ability of RBD should be more or less preserved with the replacement of S42 with P42. It was therefore unexpected that RBD(P) showed no binding to dsRNA, given RBD(P) lost only the sole hydrogen bond. Thus, we speculated that RBD(P) differs from RBD(S) in higher order structure in addition to the difference in the 42nd side chain. However, the biochemical data above showed that RBD(P) and RBD(S) are almost the same as to molecular weight and secondary structure contents in aqueous solution. As a consequence, it is still obscure by what mechanism the dsRNA-binding ability of RBD(P) and RBD(S) are different.

3.4. Molecular modeling

Table 1 Statistics of model structures

<table>
<thead>
<tr>
<th>Procheck</th>
<th>RBD(P)</th>
<th>RBD(S)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Favored</td>
<td>93.7%</td>
<td>93.5%</td>
</tr>
<tr>
<td>Additionally</td>
<td>4.8%</td>
<td>6.5%</td>
</tr>
<tr>
<td>Generously</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Allowed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disallowed</td>
<td>1.6%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Verify3D</td>
<td>79.17%</td>
<td>100%</td>
</tr>
<tr>
<td>ERRAT</td>
<td>90.698</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 2 Statistics of the binding interfaces between the RBD proteins and dsRNA

<table>
<thead>
<tr>
<th>Interface Area (Å²)</th>
<th>RBD(P)</th>
<th>RBD(S)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2021.9</td>
<td></td>
<td>2161.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of H-bonds</th>
<th>RBD(P)</th>
<th>RBD(S)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td></td>
<td>11</td>
</tr>
</tbody>
</table>
Overall structures of the RBD(P) and RBD(S) models are also similar (Fig. 3B). RMSD between these models are 2.18 Å for backbone Ca atoms. The RMSD values of the RBD(P) and RBD(S) models to the crystal structure (2ZKO.pdb) are 1.99 and 1.02 Å, respectively. Monomers of RBD(P) and RBD(S) are composed of three α-helices called Helices 1, 2 and 3. Both of the models of RBD(P) and RBD(S) are dimers and rich in α-helices, which is consistent with the results from the gel filtration chromatography and CD measurements.

A remarkable difference between RBD(P) and RBD(S) is relative angles of Helices 2. Helix 2 of a monomer is tilted to that of the other monomer in RBD(P) (Fig. 3B). By contrast, Helix 2 of RBD(S) are parallel to each other. As a consequence, relative locations of the residues composing the dsRNA-binding interface on Helices 2 are different between RBD(P) and RBD(S). In addition, the structures of Helices 2 of the RBD models are different. Helix 2 of RBD(S) is composed of a typical α-helix, whereas that of RBD(P) has a disruption of the helix at the position of P42 (Fig. 3B). Because the dsRNA-binding interface of the RBD proteins is composed of Helices 2, such differences in the structure should have impact on the binding capability of the RBD proteins to dsRNA.

Calculated binding interface areas between RBD and dsRNA are 2021.9 and 2161.7 Å² for the RBD(P)-dsRNA and RBD(S)-dsRNA complexes, respectively (Table 2). Thus, affinities by Van der Waals interactions between RBD and dsRNA are expected to be similar for those complexes. Numbers of hydrogen bonds between RBD and dsRNA are 6 and 11 for the RBD(P)-dsRNA and RBD(S)-dsRNA complexes, respectively. This suggests that the stability of complex formation should be higher in the RBD(S)-dsRNA complex than in the RBD(P)-dsRNA complex, and is consistent with the result from the EMSA experiment in which only RBD(S) bound to dsRNA.

The guanidium group of R38 of RBD(S) is sandwiched by two phosphate groups of dsRNA and interacts with two phosphate groups of dsRNA in the complex model of RBD(S)-dsRNA (Fig. 4A). R38 of RBD(P) and RBD(S)-dsRNA are similar to each other (Fig. 3A).
is also sandwiched by two phosphate groups of dsRNA, but interacts with only one of those phosphate groups (Fig. 4B). This phosphate groups is sufficiently close to R38 to form a hydrogen bond although the other one is not. The hydroxyl group of S42 of RBD(S) interacts with a hydroxyl group of a ribosyl group of dsRNA (Fig. 4C), whereas P42 of the RBD(P) model lacks a hydrogen bond to dsRNA (Fig. 4D). This is because a side chain of Pro lacks a hydroxyl group. R38 and T49 of the RBD(S) model also form hydrogen bonds to dsRNA, whereas those of RBD(P) lack these bonds. Taken together, several elements of the local structure implicated in the interaction between RBD and dsRNA are different between the RBD(P)-dsRNA and RBD(S)-dsRNA complex models.

4. DISCUSSIONS

The mortality rate of patients infected with H5N1 influenza virus in 2003 was reported to be approximately 60% [1]. NS1 is a virulence factor contributing to high pathogenicity of H5N1, because the replacement of the 42nd residue on this protein causes critical change in the degree of pathogenicity [13]. NS1 is implicated in the occurrence of cytokine storms under oxidatively stressed conditions [12], which supports the proposal that NS1 is responsible for the high pathogenicity of H5N1 virus.

We carried out combined studies of biochemistry and molecular modeling to analyze the mechanical basis of high pathogenicity by NS1, and demonstrated that RBD(P) and RBD(S) share several common features as to their structural facets despite critical difference in the binding ability to dsRNA. Replacement of Ser at the 42nd residue with Pro is technically expected to lead to loss of a hydrogen bond because a side chain of Pro lacks a hydroxyl group. However, it is skeptical whether loss of only a hydrogen bond caused by the replacement of a Ser residue could lead to severe change in the degree of pathogenicity. This is because replacement of R38 to Ala shows much smaller impact on the pathogenicity [13].

Taken together, it appeared difficult to analyze the mechanical basis of the difference in the degree of pathogenicity between RBD(P) and RBD(S) only through the results from our biochemical experiments and discussion as to the change of only the 42nd side chain. To address this problem, we modeled the structures of the complexes of RBD(P)-dsRNA and RBD(S)-dsRNA to observe global conformational changes of the RBD proteins in binding to dsRNA. Although the overall structures of RBD(P) and RBD(S) appear similar to each other, the relative angles of Helices 2 of the two monomers are different between these RBD models. In addition, Helix 2 of RBD(P) shows a disruption at P42. A disruption in an α-helix at a Pro residue is observed in other proteins [28]. By these changes in relation to the structure of Helix 2, the relative locations of the dsRNA-binding residues such as R35, R38 and T49 were predicted to be changed. Thus, it was suggested that the binding affinity at those dsRNA-binding residues as well as the 42nd residue becomes lower by such structural changes, leading to the entire loss of the affinity of RBD(P). These results collectively suggested that such global conformational change of RBD by the replacement of the 42nd residue leads to the change in the degree of the pathogenicity of H5N1 influenza virus.

Numerous NS1 monomers form a huge polymer through dimerization mechanism at RBD and ED [29]. Our modeling results suggested a change in the dimer arrangement between RBD(P) and RBD(S), which might have an impact on the ability of NS1 to polymerize. It has been reported that RBD interacts with a variety of host proteins [14]. So, it is speculated that the structural change between RBD(S) and RBD(P) may cause changes in affinities to host proteins. These changes might also have impact on the degree of pathogenicity of H5N1.

5. CONCLUSION

The present studies showed that the replacement of the 42nd residue of NS1 has a direct impact on the binding to dsRNA, so that the 42nd residue should be responsible for the high pathogenicity of H5N1. Our structural models also suggested that the replacement causes structural change of RBD, leading to the change in the dsRNA-binding affinity. In addition, such structural change might alter the other characteristics of NS1, such as those of polymer formation and binding to other host proteins. The results from our analysis of the RBD mechanism suggested that alteration of the structure of RBD causes change in the degree of pathogenicity of H5N1. These findings may contribute to development of a new strategy for designing drugs with a novel mechanism of action, for example, a specific chemical chaperone that alters the structure of RBD.

LIST OF ABBREVIATIONS

ARDS, acute respiratory distress syndrome; DK/12, A/Duck/Guangxi/12/03; DK/27, A/Duck/Guangxi/27/03; dsRNA, double-stranded RNA; ED, effector domain; EMSA, Electrophoresis mobility shift assay; NS1, Non-structural protein 1; PR8, A/Puerto Rico/8/34; RBD, RNA-binding domain; RBD(P), RBD of NS1 with P42; RBD(S), RBD of NS1 with S42

CONFLICT OF INTEREST

The authors declare no competing financial interests.

ACKNOWLEDGEMENTS

This work was supported in part by grants-in-aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan and supported in part by Japan Science and Technology Agency (JST), e-ASIA Joint Research Program, e-ASIA JRP. YSK and KS designed study and performed experiments. YSK, KF and KS analyzed data and wrote manuscript.

REFERENCES


