Alanine-scanning mutagenesis for revealing the role of highly conserved regions of influenza A virus neuraminidase

Chee-Keng Mok 1, Guang-Wu Chen 2,3 and Shin-Ru Shih 1,3,4

Graduate Institute of Bio-Medical Science, College of Medicine 1, Department of Computer Science and Information Engineering, College of Engineering 2, Research Center for Emerging Viral Infections 3 and Department of Medical Biotechnology & Laboratory Science, College of Medicine 4, Chang Gung University, Tao-Yuan, Taiwan, R.O.C.

Abstract

Background: Neuraminidase (NA), a surface glycoprotein of influenza A virus, is an important molecular target for antiviral drugs. Recent reported NA-inhibition drug resistant cases, such as mutations in H274Y in N1 and R292K and E119G/A/D in N9 and N2, however, has raised concerns about the needs for better understanding NA functional role towards developing a new anti-influenza drug.

Methods: Twenty-eight highly conserved amino acid residues in influenza A viral NA protein were identified through in silico analysis based on 2,827 NA sequences deposited in GenBank, NCBI. To understand the role of conserved residues on viral viability, we have introduced series of mutations (alanine substitutions) into NA by reverse genetics using A/WSN/33(H1N1) as a backbone.

Results: Seven out of 28 mutants were rescued, indicating that the other 21 positions in NA are essential to viral viability. Among those 21 lethal mutants, 5 were rescued by exogenously adding NA from C. perfringens, suggesting that these 5 positions may reside on the NA active site. This assumption was reinforced by structural modeling by SWISS-MODEL server. By simulation, we also found 9/21 mutants located on the side-surface of NA protein. When being substituted, they remarkably reduced the virus survival by losing its biological function, possibly due to the associated structural alteration.

Conclusion: This study identified several amino acid residues important for viral viability. Via structural simulation they are found located at NA active sites or oligomerization sites. We believe the results obtained herein provide valuable information in antiviral drug design targeting influenza NA protein.
Table of Contents

ABSTRACT

CHAPTER 1 INTRODUCTION
  1.1 Background
  1.2 Review of Related Study
  1.3 Objective of the Study

CHAPTER 2 MATERIALS AND METHODS
  2.1 Viruses and cells
  2.2 Generation of A/WSN/33 wild type and NA mutated viruses by reverse genetics
  2.3 Transfection of 293T cells
  2.4 Plaque assay in MDCK cells
  2.5 Sequence analysis
  2.6 In silico based NA structure modification analysis

CHAPTER 3 RESULT
  3.1 In silico based surveillance assay
  3.2 Generation of recombinant influenza A virus with mutation on NA
  3.3 Computational prediction for NA-mutants Structure

CHAPTER 4 DISCUSSION

REFERENCE

APPENDIX
  Table 1: Viability of rescued NA mutant viruses by reverse genetics system.
  Figure 1: Alignment of NA from consensus sequences with various subtypes.
  Figure 2: Plaque morphology of successfully rescued viruses.
  Figure 3: NA structure labeled with 28 amino acid residues.
  Figure 4: Three viability phenomena of NA mutant virus found.
  Figure 5: The interrupted mutagenesis may affect NA oligomerization.

VITA
CHAPTER 1  Introduction

1.1 Background

Influenza A viruses are negative-strand enveloped RNA viruses of the family Orthomyxoviridae that infect a wide range of warm-blooded animals, including domestic and wild birds and mammals (e.g., humans, pigs, and horses) (1). The natural reservoir for influenza virus is thought to be wild waterfowl, and genetic material from avian strains episodically emerges in strains infectious to humans. These human viruses continually circulate in yearly epidemics (mainly during the winter months in temperate climates), and antigenically novel strains emerge sporadically as pandemic viruses (2, 3).

Influenza A is comprises a protein coat or capsid that houses the viral genome, a single strand of RNA split into eight segments (Influenza A and B, influenza C only split into seven segments), each carrying a single gene. One of these encodes hemagglutinin (HA), a surface antigen that is used by the virus to bind to and break into host cells. Another gene produces a second surface antigen, neuraminidase (NA), which helps newly formed virus release to infect other cells (4).

The host immune system targets HA and NA, subjecting them to strong selective pressure. Genetic mutation, subtly altering the antigenic properties of these two target proteins, is sometimes referred to as antigenic drift. Antigenic shift, by contrast, is a major change in the antigenic properties of these two proteins caused by reassortment of one or more of the 16 known HA subtypes (H1 to H16) or the nine NA subtypes (N1 to N9) that circulate in wild birds (5).

Humans have been exposed to only H1, H2, and H3 viruses in the recent past. Consequently, a virus with an unfamiliar subtype, such as H5N1, will go undetected by the immune system of everyone alive today (6). The six other genes that make up the influenza genome have received less scrutiny, but
there is increasing evidence that they play an important role in adaptation of virus to host (7).

Pandemic influenza A virus infections have occurred three times during the last century, among which the 1957 (H2N2) and 1968 (H3N2) pandemic strains emerged from a reassortment of human and avian viruses (8, 9). Recently, all eight genome segments from the 1918 (H1N1) influenza A virus have been completely sequenced. The results indicate that the 1918 pandemic virus may not have emerged by a reassortment from avian and human virus as did the two other pandemic strains. While the 1918 H1N1 is not considered a bird virus, it is the most bird-like one among all mammalian influenza viruses (10, 11). The recent circulation of highly pathogenic avian H5N1 viruses in Asia from 2003-2006 has caused at least 90 human deaths and has raised concern about the development of a new pandemic (12).

1.2 Review of Related Study

Two classes of anti-influenza virus antiviral agents targeting either the M2 ion channel or the NA are currently available for influenza management and under consideration for stockpiling in the event of an influenza pandemic. However, use of the M2 blockers, amantadine and rimantadine, is limited by a lack of inhibitory effect against influenza B viruses (influenza B does not possess an M2 protein), side effects, and a rapid emergence of antiviral resistance. M2 inhibitor-resistant variants are transmissible from person to person, pathogenic, and can be recovered occasionally from untreated individuals. Importantly, recent human isolates of highly virulent A/H5N1 influenza viruses are naturally resistant to these drugs (13).

The two NA inhibitors, zanamivir and oseltamivir, are the only drugs, apart from M2 inhibitors, which have been approved for treating influenza virus infections in humans (14). Zanamivir has been approved for treatment the uncomplicated influenza infections in patients of 7 years of age or older. The delivery of zanamivir to the respiratory tract requires an inhaling device and, the
drug is not recommended for treatment of people with chronic respiratory disease. Because the limitation, an orally bioavailable drug, Oseltamivir is imperative. Following the increased cases of amantadine resistant, cause the M2 mutation, the emergence of oseltamivir resistance in clinical isolates has been found and associated with substitutions at residue 119, 198, 274, 292, or 294 in the NA active site (15-17). In contrast, the mutations that offer resistance to the NA inhibitors are at the highly conserved regions of active site of the NA enzyme. These mutations result in compromised NA activity and therefore a lower potential for the emergence of resistance.

1.3 Objective of the Study

NA, termed sialidase, possessed a rather unique enzyme that cleaved sialic acid from sialylated glycoproteins on influenza virus. Sialic acid is the receptor for the influenza virus on host cells and red blood cells. Only after painstaking work by biochemists and electron microscopists had defined two surface spikes on the influenza virus, was one of them, a strikingly long stalked mushroom, identified as the NA.

NA has nine subtypes with different functions and antigenicities that circulate in the host. Their genomic sequences differ significantly from one to another, with only 52% identity (data not shown). Additionally, in each subtypes, with the combination of NA with different HA is responsible for a difference of nearly 50% among them. The statistical results indicate that the NA entropy exceeds that of other genes.

However, in functional structure screening, the NA proteins are significantly conserved. These results demonstrate that some of the regions of NA protein play an important role for maintaining function. Through the in silico prediction and site-direct mutagenesis reverse genetics, the NA-mutant viruses were produced.
CHAPTER 2  Materials and Methods

2.1 Viruses and cells

All of the viruses were generated from A/WSN/33 reverse genetic system. The system plasmids was provided by Fodor E. et al. (18) from Netherlands. 293T cell line from our laboratory was used in transfection. MDCK cell line from Chang Gung Memorial Hospital Dept. Clinical Virology Laboratory was used in maintaining and amplifying the viruses.

2.2 Generation of A/WSN/33 wild type and NA mutated viruses by reverse genetics

The NA genomes of the wild-type and NA mutated A/WSN/33 viruses passaged in vitro in our laboratory were amplified by reverse transcription-PCR (19), and their sequences were determined by the DNA Sequencer. The wild-type virus was used as the backbone of recombinant viruses generated for this study. Twelve plasmids were constructed, four plasmids for producing the polymerase protein, and seven plasmids each containing the cDNA of one of the eight gene segments of the wild-type virus. A plasmid that contained the cDNA of the NA gene of the systematically mutated were also constructed as previously described (18).

Constructed plasmids were sequenced to ensure that they were identical to the field strain. Recombinant viruses that were identical, with the exception of the serial introduced NA mutations, were generated by transfecting 293T cells with the wild-type NA plasmid (WT-NA) or the NA mutant plasmids that contained the serial mutations (Figure 1). Stock viruses were prepared by three passages of the viruses rescued from 293T in MDCK cells, in Dulbecco's minimal essential medium (DMEM) supplemented with 0% fetal bovine serum (FBS) and 1 μg/mL trypsin.
2.3 Transfection of 293T cells

Transient calcium phosphate-mediated transfections of 293T cells were performed essentially as described (20). Cells were plated the day before transfection in 6 wells culture plates to obtain 50% confluent monolayers. After overnight transfection with 12 μg/mL plasmid DNA, the transfection medium was replaced with fresh medium supplemented with 0.5% FBS for virus production or 10% FBS for all other transfections. Cells were incubated for 96 h, after which supernatants were harvested. Virus-containing supernatants were cleared by centrifugation for 10 min at 300 x g. Virus titers in the supernatant were determined either directly or upon storage at 4 °C for less than 1 week, or at -80 °C for longer than 1 week.

2.4 Plaque assay in MDCK cells

Confluent MDCK cells were incubated for 1 h at 37°C with 10-fold serial dilutions of virus in 0.5 ml infection medium. The cells were then washed and overlaid with freshly prepared DMEM containing 0.2% FBS, 0.3% agar, and 1 μg/mL trypsin. The plaques were visualized after incubation at 37°C for 3 days by staining with 0.1% crystal violet solution containing 10% formaldehyde.

2.5 Sequence analysis

The RT-PCR product was purified using the QIAquick Gel Extraction Kit (Qiagen, Valencia, California). The nucleotide sequence was determined using an automated DNA sequencer. Sequence editing and processing was performed with Lasergene, version 3.18 (DNASTAR, Madison, Wisconsin). Multiple sequence alignment was performed with Clustal W version 1.83. Amino acid sequences were translated from coding sequences and aligned by BioEdit version 7.0 (25). Figure presentation were shown by GeneDoc version 2.6.002 (26). All amino acid numberings are based on A/WSN/33.
2.6 *In silico* based NA structure modification analysis

The mutant-NAs protein structure were analysis and modification by Swiss-Model program (27). The references structure was 2hty from the RCSB Protein Data Bank (28). Figure presentation were showed by PyMOL v 0.99.
CHAPTER 3  Result

3.1 *In silico* based surveillance assay

A total of 3,034 influenza A virus NA sequences were downloaded from GenBank up to date 28th February 2006. After assembling using SeqMan (Lesergene) software, only 2,827 non-redundant NA sequences sample were available for consensus analysis. There are nine influenza A virus NA subtypes (namely, N1 to N9), found from those 2,827 sequences, each of them contains 714, 1,720, 96, 16, 27, 107, 38, 76, and 33 sequences.

The reverse genetic system we used in mutagenesis analysis is of H1N1 subtype. We further grouped those 714 NA sequences of N1 subtype with their HA subtype. There are 184 H1N1 and 428 H5N1 sequences, which consist of 85.4% among those 714 NA sequences with N1 subtype, and 21.6% among all 2,827 NA sequences for all nine NA subtype.

Due to the significant genetic diversity among various NA subtype, the attempted analysis of the multiple aligned 714 NA sequences from Clustal W failed to produce a satisfactory result in GeneDoc, with only 52% identify. When lowered the threshold frequency from 80% to 25%~60%, a consensus of each subtype was successfully generated and the predicted consensus region are highlighted and displayed in figure 1. There are 95 residues found highly conserved among all nine NA subtypes. In figure 1 we also included A/WSN/33 NA sequence and shows the diagram of those 95 found residues on N1 consensus versus A/WSN/33 NA sequence.

3.2 Generation of recombinant influenza A virus with mutation on NA

Based on our *in silico* analysis, recombinant viruses with alanine substitution at highly conserved region have been generated.
The viability of the NA-mutant viruses has been examined. They were 7 of 28 viruses successfully rescued (Table 1). In the first passage, only 4 recombinant viruses have presented the cytopathic effect (CPE) (T132A, T210A, D228A, and P286A), the G332A and G388A recombinant viruses have no CPE but present in quantitative plaque assay. The E356A recombinant virus has been observed in passage 2. For the other unobservable plaque forming recombinant viruses would transflect and passages with *C. perfringens* NA to improve its viability. Compare with the control group, without adding *C. perfringens* NA, they are 6 NA-mutant viruses be recombined in additional condition.

Followed the table 1 result, in quantitative plaque assay without additional NA in overlay agar, differently sized plaques are observed in rescued viruses (figure 2A). With the help of additional NA, had 12 out of 28 mutants successfully rescued, but only 7 mutants were rescued without the additional NA. The weak mutant viruses which were rescued by additional NA produced a very small plaque in these assay (figure 2B). These observations will provide an evident to prove that the mutant of NA conserve region will critically influence its replication or viability. This observation suggesting that the NA plays an important function in the replication.

### 3.3 Computational prediction for NA-mutants Structure

The three-dimensional structure of A/WSN/33 NA protein structure was predicted using the Swiss-Model tool. Those 28 conserved residues were marked and highlighted (Figure 3). We found that the marked positions are mostly at the active site.

In figure 4, A/WSN/33 NA structure was predicted by the SWISS Model server as an example used to show these 3 categorical phenomena. In the first category of serial mutants, the reverse genetic approach was used in successful rescues of 7/28 viruses (T132A, T210A, D228A, P286A, G332A, E356A, and G388A). Most of the NA mutation points are located in the inner protein; only T132A is located on the top of the surface (figure 4A). Figure 4B showed NA mutant virus which were rescued with additional *C. perfringens* NA in
transfection and replication. Five out of 21 mutant viruses are rescued in this category (P104A, T227A, E261A, N279A, and R352A); most of them appear to affect the active side of NA. Only P104A was not located on the active side, this protein was more compact than the others. In the category of non-rescued NA mutant viruses (figure 4C), 53.3% exhibits modified NA side-surface structures. 20% increases their internal space compared with wild type and 13.3% loses disulfide bond. Exchanging residues 354G and 413G to A increases the effect of the hydrogen bond.

In the other hand, we found that there were 9/16 non-rescued NA mutants localized at the side-surface of monomer may affect NA oligomerization (figure 5), including 4/16 mutants which were near the active site. Under the transparent view, we found that other identical residues play a role to interact with another monomer.

**CHAPTER 4  Discussion**

*In silico* analysis was used to verify that 75% of mutant viruses are not recombinant or a functional virus. This finding indicates that the highly conserved regions of NA - especially the residues that surround the side surfaces - are very important to viral survival.

Exchanging the NA origin residue to Ala may cause loss of protein function, possibly an interaction with other viral protein, cleavage of the sialid acid or attraction to other cytoplasm protein.

The residues in the surface protein may constitute a good target of anti-viral drug design, as they can effectively inhibit viral growth or combination.
Reference


Table 1: Viability of rescued NA mutant viruses by reverse genetics system.

<table>
<thead>
<tr>
<th>Viruses</th>
<th>Mock</th>
<th>Medium with C. perfringens NA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P1°</td>
<td>P3°</td>
</tr>
<tr>
<td>Wild type</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>K86A</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>D87A</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>P104A</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>C108A</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>C113A</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>T132A</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>D135A</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>R136A</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>R140A</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>L142A</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>G181A</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Y192A</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>T210A</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>R209A</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>T227A</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>D228A</td>
<td>±</td>
<td>+</td>
</tr>
<tr>
<td>E261A</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Y266A</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>N279A</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>P286A</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>G332A</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>R352A</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>G354A</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>E356A</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>G388A</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>G413A</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>S425A</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>D443A</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>H259A</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

±, a little cells death;  
+, cytopathic effect (CPE) and plaque forming presented.  
-, cytopathic effect (CPE) and plaque form absent.
Figure 1: Alignment of NA from consensus sequences with various subtypes.

A total of 2,827 influenza A virus NA sequences were collected (as of 28th February 2006) from the NCBI database, N1 to N9 were associated with 714, 1,720, 96, 16, 27, 107, 38, 76 and 33 samples, respectively. Based on HA, the 714 N1 sequences comprised with 11 subtypes, more were H1N1 (25.8%) and H5N1 (59.7%). Nine NA consensus sequences were aligned with A/WSN/33 NA. The highlighted black amino acid residues are highly conserved regions of all NA. Twenty-eight residues with corresponding conditions were labeled below the sequence.
Figure 2: Plaque morphology of successfully rescued viruses.

All the NA mutant viruses were rescued by additional *C. perfringens* NA. In quantitative plaque assay without additional NA in overlay agar, differently sized plaques are observed in rescued viruses (A). With the help of additional NA, had 12 out of 28 mutants successfully rescued, but only 7 mutants were rescued without the additional NA. The weak mutant viruses which were rescued by additional NA produced a very small plaque in these assay, suggesting that the NA plays an important function in the replication (B).
Figure 3: NA structure labeled with 28 amino acid residues. NA monomer is formed by 6 beta-sheet domains. Our analyzed results show that conserved amino residues were distributed in each domain. (A) Top view; (B) side view. (Reference: 2HTY, N1 Neuraminidase, 86% identity)
Figure 4: Three viability phenomena of NA mutant virus found.

A/WSN/33 NA structure was predicted by the SWISS Model server as an example used to show these 3 categorical phenomena. (A) In the first category of serial mutants, the reverse genetic approach was used in successful rescues of 7/28 viruses (T132A, T210A, D228A, P286A, G332A, E356A, and G388A). Most of the NA mutation points are located in the inner protein; only T132A is located on the top of the surface. (B) NA mutant virus which were rescued with additional *C. perfringens* NA in transfection and replication. Five out of 21 mutant viruses are rescued in this category (P104A, T227A, E261A, N279A, and R352A); most of them appear to affect the active side of NA. Only P104A was not located on the active side, this protein was more compact than the others. (C) In the category of non-rescued NA mutant viruses, 53.3% exhibits modified NA side-surface structures. 20% increases their internal space compared with wild type and 13.3% loses disulfide bond. Exchanging residues 354G and 413G to A increases the effect of the hydrogen bond.
Figure 5: The interrupted mutagenesis may affect NA oligomerization.

There were 9/16 non-rescued NA mutants localized at the side-surface of monomer, including 4/16 mutants which were near the active site (A). Under the transparent view, we found that other identical residues play a role to interact with another monomer (B and C), such as the yellow monomer interacts with the blue monomer.
Vita

Chee-Keng MOK

1980 Born on Feb 1, in Seremban, state of Negeri Sembilan Darul Khusus, Malaysia.

1999-2003 B.S., Division of Medical Technology, School of Medical Technology, College of Medicine, Chang Gung University, Tao-Yuan, R.O.C.

Publication


