In silico mutation analysis of non-structural protein-5 (NS5) dengue virus

R D Puspitasari and U S F Tambunan
Bioinformatics Research Group, Department of Chemistry, Faculty of Mathematics and Natural Sciences Universitas Indonesia, Kampus UI Depok, Depok 16424, Indonesia

Corresponding author's e-mail: usman@ui.ac.id

Abstract. Dengue fever is a world disease. It is endemic in more than 100 countries. Information about the effect of mutations in the virus is important in drug design and development. In this research, we studied the effect of mutation on NS5 dengue virus. NS5 is the large protein containing 67% amino acid similarity in DENV 1-4 and has multifunctional enzymatic activities. Dengue virus is an RNA virus that has very high mutation frequency with an average of 100 times higher than DNA mutations, and the accumulation of mutations will be possible to generate the new serotype. In this study, we report that mutation occurs in NS5 of DENV serotype 3, glutamine mutates into methionine at position 10 and threonine mutates into isoleucine at position 55. These residues are part of the domain named S-Adenosyl-L-Methionine-Dependent Methyltransferase (IPR029063).

Keywords: NS5, DENV3, mutation, glutamine, threonine

1. Introduction

Until now, dengue fever remains a global health problem that endemic in more than 100 countries [1-2]. In 2015, an increasing number of cases were reported in Brazil and neighboring countries. An estimated 500,000 people have severe dengue fever, most of them were children, and around 25,000 were dead [3-4]. Dengue virus is a flaviridae virus transmitted by the Aedes aegypti and Ae. albopictus mosquito. Dengue virus has five serotypes (DENV1-5) [5-7]. DENV-1 and DENV-4 are often associated with primary infection, while DENV-2 and DENV-3 are twice more likely to infect and causing dengue hemorrhagic fever (DHF) to human as the secondary infection [8-9]. DENV-5 has been found in Sarawak Malaysia in 2007. The sample was taken from a farmer and has been officially announced in 2013 [6]. The discovery of the new serotypes possibly delays the development of vaccines and antiviral therapy, which have been done before [7].

Dengue virus is a ssRNA positive-strand, packed in nucleocapsid and enveloped in a lipid bilayer-glycoprotein. A genomic RNA consists of a single open reading frame (ORF) which encoded into 10 proteins. The three structural proteins (Capsid, pre-Membrane and Envelope) and seven non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5) [10]. Dengue virus is an RNA virus that has very high mutation frequency with an average of 100 times higher than the genomic DNA mutations. The accumulation of mutations is a continuous process, together with the possibility of intramolecular recombination for their simultaneous infection with a different serotype that could cause the appearance of a new serotype of dengue virus [11].
Mutation can affect the protein structure and function. It is located in the protein-coding region of the genome. The information about structural and functional effects of protein can contribute to further drug design research [12]. In this study, we used non-structural protein 5 (NS5) dengue virus as a protein target. The non-structural protein 5 (NS5) has been reported as the best characterized DENV non-structural protein. NS5 is a large protein contains 67% amino acid similarity with DENV 1-4 and has multifunctional protein with several enzymatic activities [13]. The purpose of this study was to evaluate the effects of mutation on NS5 Dengue virus.

2. Methods
The NS5 amino acid sequence of dengue virus was downloaded in GenBank at NCBI (http://www.ncbi.nlm.nih.gov/). The 3D structure was downloaded from RCSB PDB (http://www.rcsb.org), and we used GIRAF [14] server to get the best protein model.

Amino acid sequence of dengue serotype was aligned using online Basic Local Alignment Search Tool (BLAST) for proteins at NCBI (http://blast.ncbi.nlm.nih.gov/), Multiple Sequences Alignment were conducted by using Clustal Omega at EMBL-EBI (http://www.ebi.ac.uk/Tools/msa/clustalo/) to search the mutated region. In order to find the conserved region, we used Consurf [15] server (http://consurf.tau.ac.il/). We used HOPE server [12] (www.cmbi.ru.nl/hope/) to predict the effects of the mutation on NS5 DENV. MOE 2014.09 was used to predict the binding site position.

3. Results and discussion
NS5 is the protein that plays an important role in the dengue virus replication [13]. Information about the effects of mutations on the NS5 is important in drug design and development. Comparison and classification of protein structure are essential to understand protein function. Due to the problem in computational capabilities and the increasing amount of data structures, it is not possible to compare all the existing structures using standard methods. In this research, we used GIRAF [14] server to search the best protein model using a similarity ligand binding site. From this server, we obtained the highest similarity score of ligand binding sites, a protein with PDBID 3P8Z chain A. This protein is a model of NS5 methyltransferase dengue serotype 3. Additionally, through this method, we obtained the protein function prediction. The function prediction of NS5 methyltransferase is shown in table 1. Researchers have reported that NS5 methyltransferase is one of the most important parts of dengue virus replication [13]. Based on the result of function prediction, NS5 proved to be one of the most important parts of dengue virus replication. Some researchers have used NS5 methyltransferase to design dengue drugs [16-17].

The sequence of NS5 methyltransferase dengue virus serotype 3 was aligned using BLAST, and then an alignment of multiple sequences was conducted by Clustal Omega at EMBL-EBI. Based on the results of multiple sequence alignment (MSA) in figure 1, amino acid glutamine (Q) at position 10 was mutated into methionine (M) and threonine (T) at position 55 was mutated to isoleucine (I). Amino acids have specific characteristics which can play a role in protein folding, stability, complex interactions of proteins and protein function [18]. Mutations can cause changes in the secondary and tertiary structure.

![Multiple sequence alignment of NS5 methyltransferase DENV3](http://www.ebi.ac.uk/Tools/msa/clustalo/)
of proteins. Changes in the secondary and tertiary structures affect the protein function and folding. The change of secondary and tertiary structure can affect the molecular docking simulation during drug design process.

The sequence was analyzed by Consurf server. From this server we obtained the variable region as well as the conserved region, which were marked using different colors on each amino acid residue. The amino acids glutamine and threonine are located in the variable regions (figure 2). At this position, the amino acids can be mutated.

Mutation analysis was done using HOPE server [12]. The structure of glutamine and methionine have different sizes (figure 3). The methionine residue has a smaller size than the glutamine residue. These residues are part of the domain named S-Adenosyl-L-Methionine-Dependent Methyltransferase (IPR029063). Although a structure was found, the protein might be still disordered in its native state under certain conditions. The protein is predicted to be partially disordered (61%) and the residue of interest is located in a disordered region. This residue is part of a domain named S-Adenosyl-L-Methionine-Dependent Methyltransferase (IPR029063). The other residue is part of a domain named mRNA Cap 0/1 Methyltransferase (IPR026490). This domain is annotated with the following Gene-Ontology (GO) terms to indicate its function as mRNA (nucleoside-2-O-)-Methyltransferase Activity (GO:0004483) and mRNA (guanine-N7-)-Methyltransferase Activity (GO:0004482).

The mutated residue is located in a domain that important for the protein activity. It is possible that this interaction is important for the correct function of the protein. The mutation can affect this interaction and protein function.

Table 1. Molecular and biological function of non-structural protein 5 (NS5)

| Chain | GOId     | Contents                                                                 |
|-------|----------|--------------------------------------------------------------------------|
| A     | 0005524  | ATP binding                                                              |
| A     | 0008026  | ATP-dependent helicase                                                   |
| A     | 0003725  | Double-strand RNA binding                                                |
| A     | 0046872  | Metal ion binding                                                       |
| A     | 0004482  | mRNA (guanine-N7-) Methyltransferase activity                            |
| A     | 0004483  | mRNA (nucleoside-2’-O-) Methyltransferase activity                      |
| A     | 0003724  | RNA Helicase activity                                                    |
| A     | 0003968  | RNA-directed RNA polymerase activity                                     |
| A     | 0004252  | Serine-type endopeptidase activity                                       |
| A     | 0070008  | Serine-type exopeptidase activity                                       |
| A     | 0005198  | Structural molecule activity                                             |
| A     | 0075512  | Calthrin-mediated endocytosis of virus by host                           |
| A     | 0039654  | Fusion of virus membrane with host endosome                               |
| A     | 0039520  | Induction by virus of host autophagy                                     |
| A     | 0006355  | Regulation of transcription, DNA-templated                              |
| A     | 0039564  | Supression by virus of host STAT2 activity                               |
| A     | 0039574  | Supression by virus of host TYK2 activity                                |
| A     | 0039502  | Supression by virus of host type 1 interferon-mediated signaling pathway |
| A     | 0006351  | Transcription, DNA-templated                                             |
| A     | 0039694  | Viral genome replication                                                 |
| A     | 0019062  | Virion attachment to host cell                                           |
Figure 2. The conserved region of NS5 methyltransferase DENV3 (http://consurf.tau.ac.il/).

Figure 3. The structure of glutamin (green) mutates into methionine (red).

The structure of threonine and isoleucine have different sizes (figure 4). Isoleucine residue has a larger size and more hydrophobic than threonine. The mutation is located within a domain, annotated in
UniProt as mRNA cap 0-1 NS5-type MT. This residue is part of a domain named *S-Adenosyl-L-Methionine-Dependent Methyltransferase* (IPR029063). It is located on the surface of a domain with an unknown function. However, contact with other domain can affect protein function.

The residue is predicted (by KMAD) to be a phosphorylation site. Mutation into isoleucine disturbs the phosphorylation modification, because only serine, threonine and tyrosine can be phosphorylated.

In the process of in silico drug design and discovery, it is important to know the position of the ligand binding site. Mutations that occur in the ligand binding site would affect the interaction of the ligand to the receptor. Based on the result of the ligand binding site visualization, mutation does not occur in the ligand binding site, probably because it does not disturb the interaction of ligand to the receptor during drug design process (table 2).

### 4. Conclusions

The results of mutation analysis using HOPE server provide information about the possibility that the mutation would affect the structure and function of proteins. The mutation occurs in NS5 of DENV serotype 3, glutamine mutates into methionine at position 10 and threonine mutates into isoleucine at position 55. These residues are part of the domain named *S-Adenosyl-L-Methionine-Dependent Methyltransferase* (IPR029063). The amino acid residues Gln10 and Thr55 are not available in a binding site.

#### Table 2. Amino acid residue ligand binding side

| PDB ID 3P8Z | PDB ID 4V0Q | PDB ID 5CCV |
|-------------|-------------|-------------|
| Val 132     | Val 132     | Val 132     |
| His 110     | His 110     | His 110     |
| Asp 146     | Asp 131     | Asp 131     |
| Gly 58      | Gly 58      | Gly 58      |
| Ser 56      | Ser 56      | Ser 56      |
| Gly 85      | Gly 86      | Cys 82      |
| Gly 86      | Cys 82      | Asp 146     |
| Cys 82      | Asp 146     | Glu 111     |
|             | Lys 105     |             |

![Figure 4](image-url)
site pocket of proteins, so it might not affect the interaction of the ligand to the receptor in drug design. The result of this mutation will be used for further research design in dengue fever drug discovery.

Acknowledgements
The authors would like to thank Erwin Prasetya Toepak, who has proofread this manuscript. This research was funded by PITTA Grant Universitas Indonesia with contract No. 2049/UN2.R12/HKP.05.00/2016.

References
[1] Idrees S and Ashfaq U A 2014 Asian Pac. J. Trop. Biomed. 7 513-6
[2] Mahmood S and Ashfaq U A 2015 Sciforum Electronic Conference Series 1 b032
[3] Mustafa M S, Bansal A S and Rastogi V 2011 Med. J. Armed Forces India 67 192-3
[4] World Health Organization. Fact Sheet 117: Dengue and severe dengue [Internet]. Fact Sheet. World Health Organization; updated July 2016. p. 1-4. Available from: http://www.who.int/mediacentre/factsheets/fs117/en/
[5] Bartenschlager R and Miller S 2008 Future Microbiol. 3 155-65
[6] Normile D 2013 Science 342 415
[7] Mustafa M S, Rasotgi V, Jain S and Gupta V 2015 Med. J. Armed Forces India 71 67-70
[8] Fried J R, Gibbons R V, Kalayanarooj S, Thomas S J, Srikiatkhachorn A, Yoon I K, Jarman R G, Green S, Rothman A L and Cummings D A 2010 PLoS Negl. Trop. Dis. 4 1-6
[9] Garcia R A, Chismark E and Eggert J 2015 J. Nurse Pract. 11 34-40
[10] Herrero L J et al. 2013 Pharmacol. Ther. 137 266-82
[11] Monath T P 1994 Proc. Natl. Acad. Sci. U S A 91 2395-400
[12] Venselaar H, Te Beek T A, Kuipers R K, Hekkelman M L and Vriend G 2010 BMC Bioinformatics 11 548
[13] Noble C G and Shi P Y 2012 Antiviral Res. 96 115-26
[14] Kinjo A R and Nakamura H 2012 Biophysics (Nagoya-shi) 18 79-94
[15] Ashkenazy H, Erez E, Martz E, Pupko T and Ben-Tal N 2010 Nucleic Acids Res. 38 (Web Server Issue) W529-33
[16] Tambunan U S, Zahroh H, Utomo B B and Parikesit A A 2014 Bioinformation 10 23-7
[17] Podvimec M, Lim S P, Schmidt T, Scarsi M, Wen D, Sonntag L S, Sanschagrin P, Shenkin P S, and Schwede T 2010 J. Med. Chem. 53 1483-95
[18] Teng S, Madej T, Panchenko A and Alexov E 2009 Biophys. J. 96 2178-88