PENERBITAN PEGAWAI

Existence of two forms of L protein of Newcastle disease virus isolates due to a compensatory mutation in Domain V

E. Kusumaningtyas, W. S. Tan, Z. Zamrod, M. Eshaghi and K. Yusoff

Arch Virol (2004) 149: 1859–1865
DOI 10.1007/s00705-004-0336-5

Archives of Virology
Printed in Austria

Existence of two forms of L protein of *Newcastle disease virus* isolates due to a compensatory mutation in Domain V

Brief Report

E. Kusumaningtyas[1], W. S. Tan[1,2], Z. Zamrod[3], M. Eshaghi[1], and K. Yusoff[1,2]

[1]Department of Biochemistry and Microbiology, Faculty of Science and Environmental Studies, Serdang, Malaysia
[2]Institute of Bioscience, Universiti Putra Malaysia, Serdang, Malaysia
[3]Centre for Gene Analysis and Technology, School of Bioscience and Biotechnology, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia

Received September 30, 2003; accepted March 3, 2004
Published online April 27, 2004

Summary. Nucleotide sequence comparison of the L gene of the Malaysian neurotropic-viscerotropic velogenic NDV strain AF2240 with other NDV strains revealed a single nucleotide insertion at position 3870. This mutation is compensated by a nucleotide deletion downstream at position 3958 which results in two forms of the L proteins containing a 30-amino acid substitution in Domain V. This compensatory mutation does not correlate with the pathogenicity of the viral strains but it may affect the viral replication as Domain V is believed to play an important role in the replication of paramyxoviruses.

*

Newcastle disease virus (NDV) is an economically important virus that causes a fatal disease in chicken with a devastating effect on commercial poultry production worldwide [14]. The virus is grouped in the genus *Avulavirus* and family *Paramyxoviridae* [7]. It has a single stranded negative-sense RNA genome which contains six major genes encoding six structural proteins; nucleocapsid (NP), phospho (P), matrix (M), fusion (F), haemagglutinin-neuraminidase (HN) and large (L) proteins [17].

The viral genome is approximately 15 kb of which 8.5 kb comprises the first five genes and the remaining 6.5 kb contains the L gene [16] that encodes the L

protein with a molecular weight of approximately 250 kDa [6]. The L protein is believed to be responsible for all of the catalytic activities of the viral polymerase associated with transcription, including initiation and elongation of transcripts, as well as cotranscriptional modification of RNAs such as capping, methylation, and polyadenylation. In addition, the function of the RNA polymerase is dictated by the start, stop and restart template's *cis* signals at the borders of each transcription unit [1, 13].

This paper describes the determination and analysis of the L gene sequence of NDV strain AF2240, a Malaysian neurotropic-viscerotropic velogenic strain which is often used as the challenge virus in vaccine trials. The amino acid sequences of the L protein of the strain AF2240 was compared with the published sequences of other strains including Beaudette C, LaSota, B1, B1 Takaaki, clone 30, F48E9 and ZJ1.

NDV strain AF2240 was propagated and harvested as described [15]. Viral RNA was extracted with the Trizol LS reagent (Gibco BRL, U.S.A.). The L gene was divided into 10 overlapping fragments which were amplified by the Access RT-PCR kit (Promega, U.S.A.). The primers used were designed based on the homology between NDV strains LaSota, B1 and Beaudette C as described in Kusumaningtyas [5]. The PCR products were then cloned into pGEM-T Easy vector system I (Promega, U.S.A.) and introduced into competent *E. coli* TOP 10 (Invitrogen, U.S.A.). The transformants were screened on Luria Bertani agar supplemented with ampicillin (100 μg/ml), IPTG (0.5 mM) and X-Gal (80 μg/ml). The recombinant plasmids were screened by PCR and verified by restriction enzyme digestion.

Plasmid extraction was carried out using the Qiagen Plasmid Purification Kit (Qiagen, Germany). Sequencing was performed using an automated sequencer (ABI PRISM.3100 Genetic Analyzer). Sequencing data from all of the DNA fragments were joined and compared with other strains using the Clustalw Biology workbench (http://workbench.sdsc.edu). The predicted protein sequence of the L gene was generated using the ExPASy tool analysis (http://expasy.hcuge.ch/www/dna.html). The ORF finder was used to search for an open reading frame (ORF) of the L protein from NDV AF2240 (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). Hydropathy profile was generated and displayed graphically using the Kyte-Doolittle Hydropathy Profile (http://expasy.hcuge.ch/www/dna.html). Phylogenetic analysis was done using the Biology Workbench Alignment Tool Analysis (http://workbench.sdsc.edu).

A comparison of the L gene sequence of NDV strain AF2240 (GenBank accession no AY262106) with those of strains Beaudette C (X05399), LaSota (AF077761), B1 (AF309418), B1 Takaaki (AF375823), clone 30 (Y18898), F48E9 (AY049766) and ZJ1 (AF431744) revealed that they all had the same number of nucleotides (6704). Amino acid sequence analysis showed that they share more than 90% homology: Beaudette C (93%); LaSota (92%); B1 (92%); B1 Takaaki (92%); Clone 30 (93%); F48E9 (94%); and ZJ1 (93%).

Nucleotide sequence comparison of NDV strain AF2240 and the above strains also revealed that they can be divided into 2 groups based on the presence or

A

```
            3846                                                    3913
              :                                                       :
B1            TTCACCCCTGCATCTCTCTACAGG-TGTCACCTTACATTCACATATCCAATGATTCTCAAAGGCTGTT
LaSota        TTCACCCCTGCATCTCTCTACAGG-TGTCACCTTACATTCACATATCCAATGATTCTCAAAGGCTGTT
Beaudette C   TTCACCCCTGCATCTCTCTACAGG-TGTCACCTTACATTCACATATCCAATGATTCTCAAAGGCTATT
ZJ1           TTTACCCCTGCATCTCTCTACAGG-TGTCGCCTTACGTTCACATATCCAATGATTCTCAAAGGCTATT
B1 Takaaki    TTCACCCCTGCATCTCTCTACAGGGTGTCACCTTACATTCACATATCCAATGATTCTCAAAGGCTGTT
Clone 30      TTCACCCCTGCATCTCTCTACAGGGTGTCACCTTACATTCACATATCCAATGATTCTCAAAGGCTGTT
F48E9         TTCACCCCTGCATCTCTATACAGGGTGTCACCTTACATTCACATATCTAATGATTCTCAAAGGCTATT
AF2240        TTTACCCCTGCATCTCTTTACAGGGTATCACCTTACATTCACATATCTAATGATTCTCAAAGGCTATT

            3914                                                    3981
              :                                                       :
B1            CACTGAAGAAGGAGTCAAAGAGGGGAATGTGGTTTACCAACAGAGTCATGCTCTTGGGTTTATCTCTA
LaSota        CACTGAAGAAGGAGTCAAAGAGGGGAATGTGGTTTACCAACAGAGTCATGCTCTTGGGTTTATCTCTA
Beaudette C   CACTGAAGAAGGAGTCAAAGAGGGGAATGTGGTTTATCAACAGAGTCATGCTCTTGGGTTTATCTCTA
ZJ1           CACCGAAGAAGGGGTCAAAGAGGGAAACGTGGTTTACCAACAAATTCATGCTCTTGGGTTTATCCCTA
B1Takaaki     CACTGAAGAAGGAGTCAAAGAGGGGAATGTGGTTTACCAACAGA-TCATGCTCTTGGGTTTATCTCTA
Clone 30      CACTGAAGAAGGAGTCAAAGAGGGGAATGTGGTTTACCAACAGA-TCATGCTCTTGGGTTTATCTCTA
F48E9         CACTGAAGAAGGAGTCAAAGAAGGGAATGTGATTTATCAGCAAA-TCATGCTCTTGGGTTTATCTCTT
AF2240        TACCGAAGAAGGAATCAAAGAGGGGAATGTAGTTTATCAACAAA-TTATGCTCTTGGGCTTATCACTA
```

B

```
                        1279                                    1323
                          :                                       :
          B1              FTPASLYRCHLTFTYPMILKGCSLKKESKRGMWFTNRVMLLGLSL
Group     LaSota          FTPASLYRCHLTFTYPMILKGCSLKKESKRGMWFTNRVMLLGLSL
  A       Beaudette C     FTPASLYRCHLTFTYPMILKGYSLKKESKRGMWFINRVMLLGLSL
          ZJ1             FTPASLYRCRLTFTYPMILKGYSPKKGSKRETWFTNKFMLLGLSL

Group     B1 Takaaki      FTPASLYRVSPYIHISNDSQRLFTEEGVKEGNVVYQQIMLLGLSL
  B       Clone 30        FTPASLYRVSPYIHISNDSQRLFTEEGVKEGNVVYQQIMLLGLSL
          F48E9           FTPASLYRVSPYIHISNDSQRLFTEEGVKEGNVIYQQIMLLGLSL
          AF2240          FTPASLYRVSPYIHISNDSQRLFTEEGIKEGNVVYQQIMLLGLSL
```

Fig. 1. Compensatory mutation in the L gene sequence (**A**) and its subsequent amino acid substitutions the L protein (**B**). Strains B1 Takaaki, Clone 30, F48E9 and AF2240 contain a G at position 3870 which is absent in strains ZJ1, La Sota, Beaudette C and B1. On the other hand, at position 3958, strains B1, La Sota, Beaudette C and ZJ1 contain a nucleotide (G/T) while strains F48E9, B1 Takaaki, Clone 30 and AF 2240 do not. This mutation results in a 30 amino acid substitution (shaded) within the Domain V of the L protein, dividing the NDV strains into Groups A (Strains B1, La Sota, Beaudette C and ZJ1) and B (AF2240, B1 Takaaki, Clone 30, and F48E9)

absence of an insertion at nucleotide position with respect to the L gene sequence of Beaudette C [16]. Figure 1A shows a purine (G) nucleotide insertion at position 3870 of strains AF2240, B1 Takaaki, Clone 30 and F48E9. A similar observation was made by Romer-Oberdorfer et al. [11] on Clone 30, LaSota and Komarov. However, this frameshift mutation is compensated by a nucleotide deletion (G or T) downstream at position 3958 in all of these strains. As shown in Fig. 1B, NDV strain AF2240, which has similar amino acid sequence with B1 Takaaki, Clone 30 and F48E9, can be placed in one group while B1, La Sota, Beaudette C and ZJ1 are placed in other group.

This difference takes place in Domain V which contains the conserved motif MWFINRV (from positions 1310 to 1316) common to member of *Mononegavirales* [10]. Domain V has been proposed to play an important role in transcription and thermosensitivity of isolates of Vesicular stomatitis virus and the replication of Sendai virus [2, 3]. Site directed mutagenesis across Domain V (amino acids 1129–1378) of the *Sendai virus* isolates showed that majority of the mutants could synthesize mRNA, but not the genomic RNA *in vitro*. Thus, this difference in the amino acid sequence within Domain V may also affect the replication process of NDV. In addition, Park [9] suggested that the conserved region in Domain V is located between two independently folding halves of the L protein. Therefore, the difference between these sequences may correlate with the L protein folding which will subsequently influence the protein conformation. The change in structure may be related with the function of the L protein in the viral replication. It is possible that this compensatory frameshift mutation is able to restore the fitness of the L protein during an adaptation process in the environment.

This compensatory mutation in NDV AF2240 has changed the hydrophobicity and charge of the L protein. The polarity of the amino acid residues in this region, between positions 1287 to 1316, may affect the stability of the protein and consequently affects the function of the L protein. Figure 2 shows the hydropathy profile of the two different regions. The graphs indicate that the region from amino acids at positions 1292 to 1312 in Beaudette C contains hydrophobic amino acids (Fig. 2A) while AF2240 contains hydrophilic residues (Fig. 2B). The hydrophobic residues tend to form cluster and are normally buried in the internal part of the folded protein. Therefore, it is believed that these amino acids of group A may be buried inside whilst those of group B are exposed on the surface of the protein.

The estimated pI of the 30-amino acid substituted region for Beaudette C is 10.02 while AF2240 is 4.83. Despite the large difference in the pI of this region, it does not significantly affect the pI value of the whole protein of these two strains: Beaudette C, pI = 7.0; AF2240, pI = 6.85. This suggests that the two forms of L protein may possess the same net charge at a particular pH. In addition, secondary structural analysis of the L protein sequences revealed that some differences of the amino acid residues took place in the coil and helix regions (positions 1287 to 1316). As these coil regions are responsible in protein folding, it is likely that the L proteins in groups A and B may have folded differently. A consequence of this may be that the contact sites with other interacting molecules [8] involved in transcription may be affected. However, this has yet to be proven by using specific monoclonal antibodies against this region to detect any significant structural changes.

Interestingly, the conserved motif (at positions 1310–1316) in strain AF2240 (NVVYQQI) is more similar with that of Sendai virus (NLVYQQI; at positions 1341–1347) than Beaudette C (MWFINRV; at positions 1310–1316). If this conserved motif plays a specific function, the process related with this function in NDV strain AF2240 may be more similar with Sendai virus than with Beaudette C.

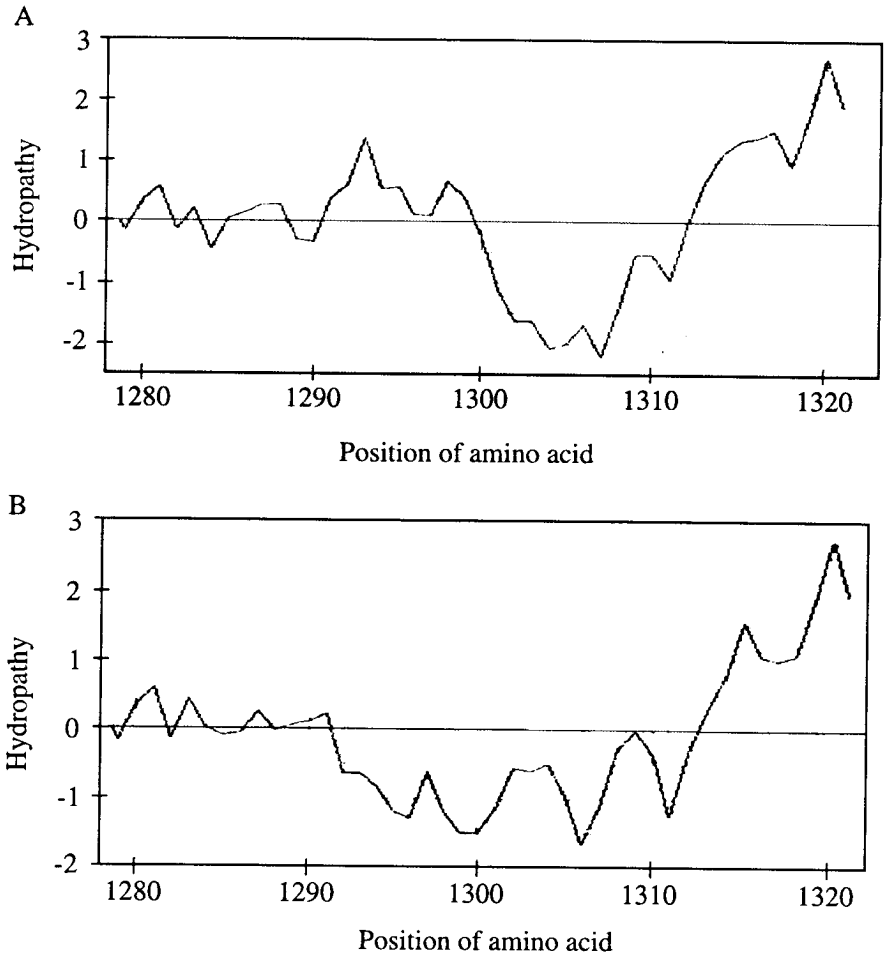

Fig. 2. Hydropathy profile of the frameshift region of NDV strain Beaudette C (**A**) and AF2240 (**B**). A window of 9 amino acids was used to calculate local hydrophobicity of each position using the procedure of Kyte and Doolittle [4]. Hydrophobic regions are shown above and hydrophilic regions are shown below the horizontal line

Phylogenetic analysis of the L proteins showed that AF2240 is closest to ZJ1 [5] although they belong to different groups based on the frameshift mutation grouping (Fig. 1). This frameshift does not significantly affect the phylogeny because the residues involved only constitute 1.36% of the whole L protein. Thus it is more practical to look at the functions and activities of the two forms of L protein because the changes took place in a highly conserved domain V.

In conclusion, we have identified a compensatory frameshift mutation in the L proteins of NDV that changes not only the hydrophobicity but also the charge of amino acids. The NDV strains can be divided into two groups based on the different forms of their L protein. This difference is not correlated to pathogenicity because each group contains virulent and avirulent strains. Further research on this region

will be useful to gain better understanding about the structure and function of the L protein of NDV.

Acknowledgements

This study was supported by IRPA Grant No 01-02-04-0107 and 01-02-04-003 BTK/ER/006 of the Ministry of Science, Technology and the Environment of Malaysia. E. Kusumaningtyas was supported by the PAATP ADB Loan No 1526-INO of the Agency for Agricultural Research and Development (AARD), Government of Indonesia.

References

1. Banerjee AK (1987) Transcription and replication of rhabdoviruses. Microbiol Rev 51: 66–87
2. Cortese CK, Feller JA, Moyer SA (2000) Mutations in domain V of the Sendai virus L polymerase protein uncouple transcription and replication and differentially affect replication in vitro and in vivo. Virology 277: 387–396
3. Hunt DM, Hutchinson KL (1993) Amino acids changes in the L polymerase protein of vesicular stomatitis virus which confer aberrant polyadenylation and temperature-sensitive phenotypes. Virology 193: 786–793
4. Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol 157: 105–132
5. Kusumangingtyas E (2003) Sequence analysis of the L gene of Newcastle disease virus strain AF2240. MSc thesis, Universiti Putra Malaysia
6. Lamb RA, Mahy BWJ, Chopin PW (1976) The synthesis of Sendai virus polypeptides in infected cells. Virology 69: 116–131
7. Mayo MA (2002) A summary of taxonomic changes recently approved by ICTV. Arch Virol 147: 1655–1656
8. Ortiz AR, Skolnick J (2000) Sequence evolution and the mechanism of protein folding. Biophys J 79: 1787–1799
9. Park GD (1994) Mapping of a region of the paramyxovirus L protein required for the formation of a stable complex with the viral phosphoprotein. J Virol 68: 4862–4872
10. Poch O, Blumberg BM, Bouguelered L, Tordo N (1990) Sequence comparison of five polymerase (L proteins) of unsegmented negative-strand RNA viruses: theoretical assignment of functional domains. J Gen Virol 71: 1153–1162
11. Romer-Oberdorfer A, Mundt E, Mebatsion T, Buchholz UJ, Mettenleiter TC (1999) Generation of recombinant lentogenic Newcastle disease from cDNA. J Gen Virol 80: 2987–2995
12. Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4: 406–425
13. Smallwood S, Easson CD, Feller JA, Horikami SM, Moyer SA (1999) Mutation in conserved domain II of the large (L) subunit of the Sendai virus RNA polymerase abolish RNA synthesis. Virology 30: 375–385
14. Spradbrow PB (1999) Newcastle Disease – an overview. In: Copland JW (ed), Newcastle disease in Poultry. Canberra, Australia
15. Tan WS, Lau CH, Ng BK, Ibrahim AL, Yusoff K (1995) Nucleotide sequence of the haemagglutinin-neuraminidase (HN) gene of Malaysian heat resistant viscerotropic velogenic Newcastle disease (NDV) strain AF2240. DNA Seq 6: 47–50

16. Yusoff K, Millar NS, Chambers P, Emmerson PT (1987) Nucleotide sequence analysis of the L gene of Newcastle disease virus: homologies with Sendai and vesicular stomatitis viruses. Nucleic Acids Res 15: 3961–3976
17. Yusoff K, Tan WS (2001) Newcastle disease virus: macromolecules and opportunities. Avian Pathol 30: 439–455

Author's address: Prof. Khatijah Yusoff, Department of Biochemistry and Microbiology, Faculty of Science and Environmental Studies, Universiti Putra Malaysia, 43400, Serdang, Selangor, Malaysia; e-mail: khatijah@fsas.upm.edu.my