View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by National Chung Hsing University Institutional Repository

Arch Virol (1997) 142: 2515–2520

Archives of Virology
© Springer-Verlag 1997
Printed in Austria

Cloning and nucleotide sequencing of the S4 genome segment of avian reovirus S1133*

Brief Report

C. J. Chiu and **L. H. Lee**

Department of Veterinary Medicine, National Chung Hsing University
Taichung, Taiwan

Accepted August 1, 1997

Summary. The sequence of RNA genome segment S4 of the avian reovirus (ARV) strain S1133 was determined. S4 RNA is 1185 base pairs long and contains one open reading frame encoding a protein of 367 amino acid residues (40.6 kDa), the similar size as the known S4 gene product (σ NS), with a net charge of -1 at neutral pH. The S4 RNA sequence possesses a pentanucleotide sequence UCAUC at the 3′-terminus of its plus strand like in ARV S1 and S3 segments and ten segments of mammalian reovirus (MRV). The predicted amino acid sequence comparison revealed that the homology is 44.02%, 45.71%, and 42.33% for ARV σ NS and three serotypes of MRV σ NS, respectively. The relatively high content of α-helix structure in the C-terminal portion of ARV σ NS suggests that this protein may functionally relate to MRV σ NS. Northern blot hybridization showed that a ^{32}P-labeled cDNA insert S4-49 from ARV S4 RNA cross-hybridized with the corresponding RNA segments of all seven strains of ARV tested.

*

Avian reovirus (ARV) is an important poultry pathogen involved in several diseases, among which viral arthritis and pale bird syndrome are the most important [2]. Genome and protein compositions of ARV are generally similar to those of mammalian reovirus (MRV), the prototype of the orthoreovirus genus [6, 17]. The genome of ARV consists of 10 segments of double-stranded RNA (dsRNA) that can be separated into three size classes: large (L), medium (M), and small (S), based on electrophoretic mobility [6, 19]. They code for 10

* The nucleotide sequence data reported in this paper have been submitted to the GenBank database and have been assigned the accession number U95952.

distinct virus-specific products, eight of which are viral structural proteins and two of which are nonstructural proteins [6, 16]. Furthermore, the gene protein coding assignment of strain S1133 have been made [16]. Protein σ NS is encoded by the S4 RNA and is one of two nonstructural proteins of ARV [16, 17]. The protein would be considered to be analogous to MRV σ NS which has been extensively studied and identified as the single-stranded RNA (ssRNA)-binding protein [7, 12].

To extend the available information on the S4 gene of ARV and its encoded protein σNS, the gene has been cloned and sequenced.

A vaccine strain of ARV S1133 (Vineland Laboratories) has been adapted for growth on chicken embryo fibroblast (CEF) and plaque purified twice [19]. Following propagation in CEF, virions were purified and viral dsRNA was extracted by the digestion with proteinase K and isolated as described previously [10]. Genomic dsRNA was separated on 6.5% polyacrylamide gel and the S4 RNA was eluted with an RNaid kit (Bio 101). S4 RNA was then polyadenylated at the 3′ end and the ds complementary DNA (dscDNA) synthesized using a commercial cDNA synthesis kit (BRL). The product was ligated into the *Pst* I site of pBR322 and subsequently used to transform *E. coli* JM103 [3]. Transformants that were resistant to tetracycline and sensitive to ampicillin were randomly selected. The cDNA insert of the largest clone (S4−49) was determined to be specific for S4 RNA by Northern blot hybridization. cDNA insert of clone S4−49 was subcloned into pUC18 and was sequenced by the methods of Sanger et al. [14], using DNA Sequenase (US Biochemicals). Clone S4−49 was a sequence of 692 base pairs (bp) long and appeared to be incomplete because its size is shorter than the expected size of S4 RNA [6, 19]. In order to complete the sequence and to determine the sequence of 3′ untranslational region of the gene, first strand synthesis of the cDNA was performed using an RNAtag primer, an adaptor primer having a poly(T) tail [CACTGGGATCCAGATG$(T)_{17}$] after polyadenylation of S4 RNA. An oligonucleotide primer (5′ TTGATGCTTCATTTGG 3′) was designed based on the sequence of the clone S4−49 and was used with a primer identical to RNAtag but without the poly(T) tail in rapid amplification of cDNA ends (RACE; [4]). This resulted in the amplification of a 542 bp fragment which was cloned in the pCR II vector and sequenced. The fragment showed 49 overlapping nucleotides with the S4−49 clone.

Based on the S4−49 and the sequence generated by means of 3′ RACE, the complete nucleotide sequence of the S4 RNA was determined to be 1185 bp long (Fig. 1). The distribution of the four bases was found to be relatively even: 21.86% A, 27.68%T, 25.82%C, and 24.64% G. The S4 RNA contains one long open reading frame (ORF) that starts with an ATG (at residues 17 to 19) and terminates at nucleotides 1118 to 1120 with a TGA codon. No additional ORF of significant length was detected in either the plus or minus strand RNA. The start codon at position 17 to 19 has a strong context for initiation, with a purine at position −3 and a guanine at position +4[8]. The ORF encodes a protein of 367 amino acid residues and a deduced molecular weight of approximately

```
   1  gagtccttgtgcagccatggacaacaccgtgcgtgttggagtttcccgcaacacatccgggccagctggtcagacactctttagaaacttctatttacta   100
                       M  D  N  T  V  R  V  G  V  S  R  N  T  S  G  P  A  G  Q  T  L  F  R  N  F  Y  L  L
                                                         ●  ●  ●  ●
 101  cgatgtaatatttcagctgatggccgtaatgcaacgaaggcggtacaatcccactttccattcctttcacgtgctgtgcgatgcctatcgcctcttgccg   200
      R  C  N  I  S  A  D  G  R  N  A  T  K  A  V  Q  S  H  F  P  F  L  S  R  A  V  R  C  L  S  P  L  A  A
            ●  ●  ●  ●           ●  ●  ●  ●
 201  ctcactgtgctgatagaacccttcgccgtgacaacgtgaaacagattcttactcgtgaactgccatttcctcggatctaatcaactacgcacaccatgt   300
        H  C  A  D  R  T  L  R  R  D  N  V  K  Q  I  L  T  R  E  L  P  F  S  S  D  L  I  N  Y  A  H  H  V
 301  caattcatcatcccttactacctctcaaggcgtcgaagcggctcgtttggtagctcaagtttatggggaacaagtaccgttcgatcacatttatcctact   400
       N  S  S  S  L  T  T  S  Q  G  V  E  A  A  R  L  V  A  Q  V  Y  G  E  Q  V  P  F  D  H  I  Y  P  T
       ●  ●  ●  ●
 401  ggttcagcgacatactgtcctggtgcaatcgcaaatgctatttctcgcattatggctggctttgtacctcgtgaaggtgatgactttgctccgagtggcc   500
      G  S  A  T  Y  C  P  G  A  I  A  N  A  I  S  R  I  M  A  G  F  V  P  R  E  G  D  D  F  A  P  S  G  P
 501  ctattgactacctcgctgctgacctgatcgcgtataagtttgtgctcccttacatgcttgacatggtagatggtcgtcctcagattgtcctgccgtctca   600
        I  D  Y  L  A  A  D  L  I  A  Y  K  F  V  L  P  Y  M  L  D  M  V  D  G  R  P  Q  I  V  L  P  S  H
 601  taccgtagaagaaatgttgaccaacaccagcttgctgaactcgattgatgcttcatttggtatcgaagcgcgcagtgatcaaaggatgactcgtgatgct   700
        T  V  E  E  M  L  T  N  T  S  L  L  N  S  I  D  A  S  F  G  I  E  A  R  S  D  Q  R  M  T  R  D  A
                             ●  ●  ●  ●
 701  gctgagatgagttctcgctccctcaatgaacttgaggatcatgatcagagaggtcgtatgccttggaagatcatgctagcgatgatggcggcccaattga   800
      A  E  M  S  S  R  S  L  N  E  L  E  D  H  D  Q  R  G  R  M  P  W  K  I  M  L  A  M  M  A  A  Q  L  K
 801  aggttgagttggacgcgctggcggacgagcgtacggagtcacaagctaatgctcacgttacatccttcggatcccgtttatttaatcagatgtcggcgtt   900
        V  E  L  D  A  L  A  D  E  R  T  E  S  Q  A  N  A  H  V  T  S  F  G  S  R  L  F  N  Q  M  S  A  F
 901  cgttactattgatcgtgaactgatggaactggcccttctcatcaaggaacagggcttcgccatgaatccgggtcagattgcatctaagtggtcgctgata  1000
        V  T  I  D  R  E  L  M  E  L  A  L  L  I  K  E  Q  G  F  A  M  N  P  G  Q  I  A  S  K  W  S  L  I
1001  cgtcgttccggtcctactcgtccactttcaggtgcccgtcttgaaatcaggaatggtaattggatgatccgtgagggtgaccaaacgctactgtctgtct  1100
      R  R  S  G  P  T  R  P  L  S  G  A  R  L  E  I  R  N  G  N  W  M  I  R  E  G  D  Q  T  L  L  S  V  S
1101  ctccagctaggatggcgtagacgggacccatggtgcgggtgaggggtcgccacaccctctgccgcgacttggactcttattcatc  1185
        P  A  R  M  A  *
```

Fig. 1. The S4 segment RNA of ARV S1133 is presented in the DNA form along with the encoded σ NS amino acid sequence. The open reading frame starts at nucleotide 17 and terminates at nucleotide 1 120. The 3′-terminal pentanucleotide common to ten segments of MRV, and S3 and S1 RNAs of ARV is underlined. Possible glycosylation sites (●) are indicated

40.6 kDa, the similar size of viral protein σ NS as determined previously [6, 17]. The 5′ and 3′ untranslated regions of the S4 genome segment were 16 and 65 nucleotides long, respectively. The translation of σ NS from the ORF corresponds to translation of 93.2% of the S4 RNA. This confirms previous observations [6, 16] that the σNS like other ARV proteins except the small size of σC does correlate with a full-length translation of the genomic segments. Its 3′-terminus possesses a nucleotide sequence UUAUUCAUC which is identical to that of the S3 RNA of the same ARV strain S1133 [20] and of the S1 RNA of a different strain of ARV [15] at the 3′-terminus of their plus strand. The 5′-terminal sequence, GAGUCCU was found to be the same as that of the S3 RNA of the same virus strain, whereas it differed from that of the S1 RNA of a different virus strain. Such conserved sequences at both the 5′ and 3′ termini were compared with those of MRV genome segments [1, 18]. The 3′-terminal pentanucleotide UCAUC is found to be the only sequence shared by the ARV and three serotypes of MRV. The 3′-terminal sequence of the plus strand is expected to play a significant role in the initiation of minus strand synthesis. Although the recent data come from only a limited numbers of genome segments of ARV, the presence of this identical conserved sequence at their 3′-terminus suggests that the viruses may share similar replication mechanisms. Also, there is no polyadenylation signal near the 3′ end of the S4 mRNA like in MRV.

The predicted amino acid sequence of σ NS showed that it has a net charge of −1 at neutral pH and there is a mainly hydrophilic region spanning residues 250–260 and no clear characteristics of hydrophobicity [9]. A search made against the GenBank and ARV σ NS amino acid sequences has revealed that the homology is higher between ARV σ NS and σ NS of three serotypes of MRV than any other sequences and is 44.02 %, 45.71%, and 42.33%, respectively. The possible secondary structure of σ NS has the following features: 33% of the residues are in the form of α-helices, 23.3% are in the form of β-sheets, 20.5% are within turns, and 22.2% are in random coils. The α helices are not uniformly distributed and about 66% of them are in the C-terminal one-third of the molecule [5] as previously reported for MRV σ NS of the three serotypes [18]. The C-terminal half of MRV σ NS that is rich in α-helix content has been specified for the responsibility for its affinity for the ssRNA and appears to function in MRV morphogenesis. Thus, such similar structure in C-terminal portion of ARV σ NS suggests that this protein may also functionally relate to MRV σ NS. The protein contains 5 potential glycosylation sites, but no potential phosphorylation site could be predicted.

Since the MRV S3 RNA, a counterpart of ARV S4 RNA, has been shown to be highly conserved among three serotypes [18], it was interest to determine whether the ARV S4 RNAs are also conserved. Here, we carried out Northern blot hybridization with the dsRNA of eight ARV isolates belonging to different serotypes [11] using the cDNA insert S4−49 as a probe. Under high stringency condition the cDNA hybridized to the corresponding RNA segments of all the

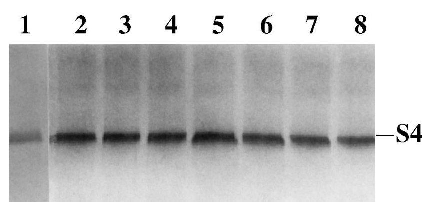

Fig. 2. Autoradiograms depicting hybridization of an ^{32}P-labeled cDNA insert S4-49 to Northern blots of dsRNA from the eight different strains of ARV. Plasmid DNA of clone S4-49 was isolated, and then ^{32}P-labeled by nick translation [13] and hybridized to ARV dsRNA which had been resolved in a 6.5% polyacrylamid gel at 70 volts for 36h and subsequently transferred to a Zeta probe blotting membrane. *1–7* indicated the samples of 601G, 750 505, R2, OS-161, 918, T6, and 1017-1 ARV local isolates [11], respectively. *8* representing the sample of S1133 is used as a positive control

strains (Fig. 2). The results suggests that the cDNA insert S4−49 may provide a suitable probe for the detection of ARV infection.

Acknowledgements

This work was equally supported by the National Science Council (NSC 85-2321-B005-050) and the council of Agriculture, Republic of China.

References

1. Antczak JB, Chmelo R, Pickup DJ, Joklik WK (1982) Sequences at both termini of the 10 genes of reovirus serotype 3 (strain Dearing). Virology 121: 307–319
2. Calnek BW, Barnes HJ, Beard CW, Reid WM, Yoder Jr HW (1991) Diseases of poultry, 9th ed. Iowa State University Press, Ames, Iowa
3. Cashdollar LW, Esparza J, Hudson GR, Chmelo R, Lee PWK, Joklik WK (1982) Cloning the double-stranded RNA genes of reovirus: sequence of the cloned S2 gene. Proc Natl Acad Sci USA 79: 7 644–7 648
4. Frohman MA, Dush MK, Martin GR (1988) Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. Proc Natl Acad Sci USA 85: 8 998–9 002
5. Garnier J, Osguthorpe DJ, Robson B (1978) Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J Mol Biol 120: 97–120
6. Gouvea VS, Schnitzer TJ (1982) Polymorphism of the genomic RNAs among the avian reovirus. J Gen Virol 61: 87–91
7. Huismans H, Joklik WK (1976) Reovirus-coded polypeptides in infected cells: isolation of two native monomeric polypeptides with affinity for single-straded and double-stranded RNA, respectively. Virology 70: 411–424
8. Kozak M (1981) Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. Nucleic Acids Res 9: 5 233–5 252
9. Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol 157: 105–132

10. Lee LH, Ting LJ, Shien JH, Shieh HK (1994) Single-tube, noninterupted reverse transcription-PCR for detection of infectious bursal disease virus. J Clin Microbiol 32: 1 268–1 272
11. Lee LH, Wang YH, Shien JH (1992) Serological characterization of reoviruses isolated from avian species in Taiwan. J Chin Soc Vet Sci 18: 67–77
12. Richardson MA, Furuichi Y (1983) Nucleotide sequence of reovirus genome segment S3, encoding non-structural protein sigma NS. Nucleic Acids Res 11: 6399–6408
13. Rigby PW, Dieckman JM, Rhodes C, Berg P (1977) Labelling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase. J Mol Biol 113: 237–251
14. Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74: 5 463–5 467
15. Shapouri MRS, Kane M, Letarte M, Bergeron J, Arella M, Silim A (1995) Cloning, sequencing and expression of the S1 gene of avian reovirus. J Gen Virol 76: 1515–1520
16. Varela R, Benavente J (1994) Protein coding assignment of the avian reovirus S1133. J Virol 68: 6 775–6 777
17. Wickramasinghe R, Meanger J, Enriquez CE, Wilcox GE (1993) Avian reovirus proteins associated with neutralization of virus infectivity. Virology 194: 688–696
18. Wiener JR, Joklik WK (1987) Comparison of the reovirus serotype 1, 2, and 3 S3 genome segments encoding the nonstructural protein σ NS. Virology 161: 332–339
19. Wu WY, Shien JH, Lee LH, Shieh HK (1994) Analysis of the double-stranded RNA genome segments avian reovirus field isolates. J Virol Methods 48: 119–122
20. Yin HS, Lee LH (1997) Characterization of the double-stranded RNA genome segment S3 of avian reovirus. J Virol Methods 67: 93–101

Authors' address: Dr. L. H. Lee, Department of Veterinary Medicine, National Chung Hsing University, Taichung, 403. Taiwan

Received June 6, 1997