Phylogenetic analysis of influenza C virus nonstructural (NS) protein genes and identification of the NS2 protein

A. S. M. Alamgir, Yoko Matsuzaki, Seiji Hongo, Emi Tsuchiya, Kanetsu Sugawara, Yasushi Muraki and Kiyoto Nakamura

Department of Bacteriology, Yamagata University School of Medicine, Iida-Nishi, Yamagata 990-9585, Japan

The nucleotide sequences of RNA segment 7 (nonstructural protein gene; NS) were compared among 34 influenza C virus strains isolated between 1947 and 1992. The results showed that all the NS genes analysed had the potential to encode NS1 and NS2 proteins of 246 and 182 amino acids, respectively. The deduced amino acid sequence of the previously unidentified NS2 was fairly well conserved, although it was more divergent than the NS1 protein sequence. Moreover, immunoprecipitation experiments with rabbit immune serum against a glutathione S-transferase fusion protein containing the C-terminal region of the 182 amino acid NS2 protein revealed synthesis of a protein with an apparent molecular mass of ~ 22 kDa in infected cells. A phylogenetic analysis showed that the 34 NS genes were split into two distinct groups, A and B. Comparison of the phylogenetic positions of the individual isolates in the NS gene tree with those in the haemagglutinin–esterase (HE) gene tree suggested that most of the influenza C viruses currently circulating in Japan, irrespective of their HE gene lineage, had acquired group B NS genes through reassortment events that presumably occurred either in the 1970s or in the early 1980s.

Introduction

The influenza C virus genome consists of seven single-stranded RNA segments. Nakada et al. (1985) determined the nucleotide sequence of the shortest RNA segment (non-structural protein gene; NS) of C/California/78 (CAL78) and showed that the gene contained 934 nucleotides with an open reading frame (ORF) capable of encoding a 286 amino acid NS1 protein. The same research group also detected a second mRNA (a spliced mRNA derived from NS1 mRNA) in infected cells that can encode a 121 amino acid NS2 protein (Nakada et al., 1986). Furthermore, Buonagurio et al. (1986) analysed the NS gene sequences of seven virus strains and found that all of the sequences were the same length as that of CAL78. On the other hand, Hongo et al. (1992) reported that, compared with the CAL78 virus NS gene, an insertion of one extra G residue occurs in the NS gene of C/Yamagata/1/88, which results in frame-shifts within both the NS1 and NS2 coding regions, directing synthesis of NS1 and NS2 proteins of 246 and 182 amino acids, respectively, and these authors claimed that the NS gene sequences of other influenza C virus strains, including that of CAL78, must be reinvestigated. Recently, Marschall et al. (1999) identified an NS1 ORF of 246 amino acids in length in the NS genes of four virus strains. One of the aims of the present study was to establish that the influenza C virus NS gene invariably contains NS1 and NS2 ORFs identical to those proposed by Hongo et al. (1992).

Recently, Marschall et al. (1999) detected a 27 kDa polypeptide in influenza C virus-infected cells that was reactive with antiserum against a glutathione S-transferase (GST) fusion protein constructed to contain NS1. However, influenza C virus NS2 has not yet been identified unequivocally in infected cells, although there are two earlier papers that reported synthesis of a protein of about 15 kDa that might be a counterpart of the NS2s of influenza A and B viruses (Petri et al., 1980; Nakada et al., 1986). The second aim of this study was to confirm that the NS2 amino acid sequence is conserved among different influenza C virus strains.

By comparing the haemagglutinin–esterase (HE) and NS gene sequences among various strains in different parts of the world over a long period of time, Buonagurio et al. (1985) suggested that influenza C virus epidemiology may be characterized by the presence of many co-circulating variants belonging to different lineages. We also compared previously the HE gene sequence among 25 isolates obtained during

Author for correspondence: Kiyoto Nakamura.
Fax +81 23 628 5250.
The DDBJ/EMBL/GenBank accession numbers of the nucleotide sequence data reported in this paper are AB034155–AB034181.
1964–1988 and showed the existence of four discrete lineages, represented by C/Yamagata/26/81, C/Aichi/1/81, C/Aomori/74 and C/Mississippi/80, three of which (C/Yamagata/26/81, C/Aichi/1/81- and C/Mississippi/80-related lineages) co-circulated in the 1980s in Japan (Muraki et al., 1996). Therefore, mixed infection with influenza C viruses belonging to different lineages is likely to occur in nature, resulting in the emergence of reassortment viruses. In fact, evidence was obtained that several virus strains (C/Yamagata/64, C/Kanagawa/1/76, C/Miyagi/77, C/England/83, C/Nara/1/85, C/Yamagata/9/88 and C/Yamagata/5/92) are naturally occurring reassortants (Peng et al., 1994, 1996; Tada et al., 1997). However, the significance of genetic reassortment in influenza C virus epidemiology is totally unknown. The third aim of this report was to identify additional reassortants by comparing the phylogenetic positions of the individual isolates between the evolutionary trees for the HE and NS genes and to obtain information about the epidemiological significance of reassortment of influenza C viruses. For these purposes, we compared the nucleotide sequences of the NS gene among 34 influenza C virus isolates obtained between 1947 and 1992. We also present evidence for the synthesis of the NS2 protein with a molecular mass of 22 kDa in influenza C virus-infected cells.

Methods

■ Viruses, RNA extraction and cells. A total of 34 strains of influenza C virus isolated in Japan, China, the UK, the Republic of South Africa and the USA during the period 1947–1992 were used (Table 1). They were each cloned twice by means of the limiting dilution method (Clontech) and the following primers: 5′ dAAAATGTCCGACAAAACTAG (positions 25–44), 5′ dCTCCTCTTTTGACCTTAGAAC (172–191), 5′ dGATGGTGACCTTCTATTCGACG (211–232), 5′ dGACCAATTGGCCAAAATCCC (328–347), 5′ dCCATGATTATCTACTTGGTG (361–383), 5′ dCCGGATTTGATTTAGG (617–631) and 5′ dGCTTATAATGCTCCGGTTTAG (733–752) for minus-strand sequencing and 5′ dTGCACTTGACGTGTC (125–126), 5′ dCGCTTGAAATGCTCACC (233–234), 5′ dCTAAGGGAAGCATAAAGG (752–733) and 5′ dAAGGGGATTTTTAACCTTGG (927–908) for plus-strand sequencing. Sequence data were analysed with the PHYLIP program version 3.54c (Felsenstein, 1989) and phylogenetic trees were constructed by the maximum-parsimony method (Fitch, 1971) using the same software. The probabilities of the internal branches were determined by bootstrap analysis (n = 100).

■ Production of antisera against the C-terminal regions of the NS1 and NS2 proteins. To create a GST fusion protein containing the C-terminal region (residues 225–246) of NS1 (GST/NS1-C), a 69 bp DNA fragment corresponding to positions 700–768 of the NS gene was prepared by PCR by using plasmid pCN5-8 (which contains nucleotides 5–923 of the YA188 virus NS gene; Hongo et al., 1992) as a template and two primers: a plus-sense primer (5′ dCTGGGATCCGGGAAATGAACACCGATATTGCACAGAGAC) containing a Bam HI site (underlined) followed by sequence corresponding to positions 700–726 and a minus-sense primer (5′ dCTCGAATTCTTATGCGAGTCTTTCAACGGCAG-AG) containing an EcoRI site (underlined) followed by sequence corresponding to positions 768–744. This DNA fragment was digested with Bam HI and EcoRI and then cloned into the Bam HI and EcoRI sites of pGEX-2T (Pharmacia) to give pGEX/NS1-C. A GST fusion protein containing the C-terminal region (residues 121–182) of NS2 (GST/NS2-C) was generated as follows: a 189 bp DNA fragment corresponding to positions 701–889 of the NS gene was amplified by PCR with a plus-sense primer composed of a Bam HI site followed by sequence corresponding to positions 701–728 (5′ dCTGGGATCCGGAGAAATGAAACACCGATATTGCACAGAGAC; Bam HI site underlined) and a minus-sense primer composed of a Smal site following by sequence corresponding to positions 889–859 (5′ dCTCCCGGTTATATAAGTTGGAATTACACAAAGATTTC; Smal site is underlined). The PCR product, after digestion with Bam HI and Smal, was cloned into the Bam HI and Smal sites of pGEX-2T to give pGEX/NS2-C. pGEX/NS1-C and pGEX/NS2-C were each transformed into E. coli strain DH5α. Cultures of the bacteria were grown to mid-exponential phase and treated with 0.1 M IPTG for 4 h. The cells were then collected by centrifugation, resuspended in Bug Buster reagent (Novagen) and incubated on a rotating mixer for 10 min at room temperature. After centrifugation at 16 000 g for 20 min, the GST fusion protein was purified from the supernatant by glutathione–Sepharose 4B (Pharmacia) affinity chromatography. Antiseria against the GST/NS1-C and GST/NS2-C proteins were raised in New Zealand white rabbits according to procedures described previously (Hongo et al., 1994).

■ Radioimmunoprecipitation. HMV-II cells infected with YA188 virus at a m.o.i. of 10 p.f.u. per cell were labelled for 1 h at 12 h post-infection (p.i.) with 50 μCi/ml [35]S)methionine (Amersham) in methionine-free RPMI 1640 medium. Cells were then disrupted in 0.01 M Tris–HCl, pH 7.4, containing 1% Triton X-100, 1% sodium deoxycholate, 0.1% SDS, 0.15 M NaCl and a cocktail of protease

1934

Downloaded from www.microbiologyresearch.org by
IP: 54.70.40.11
On: Sat, 10 Mar 2018 18:25:45
Table 1. Influenza C virus strains used in this study

<table>
<thead>
<tr>
<th>Virus strain</th>
<th>Abbreviation</th>
<th>HE gene*</th>
<th>NS gene†</th>
<th>Accession no.</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>C/Taylor/47</td>
<td>TAY47</td>
<td>ND</td>
<td>A</td>
<td>D00028</td>
<td>Buonagurio et al. (1986)</td>
</tr>
<tr>
<td>C/Ann Arbor/50</td>
<td>AA50</td>
<td>ND</td>
<td>A</td>
<td>D00029</td>
<td>Buonagurio et al. (1986)</td>
</tr>
<tr>
<td>C/Great Lakes/54</td>
<td>GL54</td>
<td>ND</td>
<td>A</td>
<td>D00030</td>
<td>Buonagurio et al. (1986)</td>
</tr>
<tr>
<td>C/Yamagata/64</td>
<td>YA64</td>
<td>AO</td>
<td>A</td>
<td>AB034155</td>
<td>This report</td>
</tr>
<tr>
<td>C/Johannesburg/4/67</td>
<td>JHG67</td>
<td>AI</td>
<td>A</td>
<td>AB034156</td>
<td>This report</td>
</tr>
<tr>
<td>C/Sapporo/71</td>
<td>SA71</td>
<td>YA</td>
<td>A</td>
<td>AB034157</td>
<td>This report</td>
</tr>
<tr>
<td>C/Aomori/74</td>
<td>AO74</td>
<td>AO</td>
<td>A</td>
<td>AB034158</td>
<td>This report</td>
</tr>
<tr>
<td>C/Kanagawa/1/76</td>
<td>KA176</td>
<td>AO</td>
<td>A</td>
<td>AB034159</td>
<td>This report</td>
</tr>
<tr>
<td>C/Miyagi/77</td>
<td>MI77</td>
<td>AO</td>
<td>A</td>
<td>AB034160</td>
<td>This report</td>
</tr>
<tr>
<td>C/California/78</td>
<td>CAL78</td>
<td>YA</td>
<td>A</td>
<td>M10087</td>
<td>Nakada et al. (1985)</td>
</tr>
<tr>
<td>C/Kyoto/1/79</td>
<td>KY179</td>
<td>YA</td>
<td>A</td>
<td>AB034161</td>
<td>This report</td>
</tr>
<tr>
<td>C/Mississippi/80</td>
<td>MS80</td>
<td>MS</td>
<td>A</td>
<td>D00032</td>
<td>Buonagurio et al. (1986)</td>
</tr>
<tr>
<td>C/Aichi/1/81</td>
<td>AI181</td>
<td>AI</td>
<td>A</td>
<td>AB034162</td>
<td>This report</td>
</tr>
<tr>
<td>C/Yamagata/26/81</td>
<td>YA2681</td>
<td>YA</td>
<td>B</td>
<td>AB034163</td>
<td>This report</td>
</tr>
<tr>
<td>C/Pig/Beijing/115/81</td>
<td>PB11581</td>
<td>YA</td>
<td>B</td>
<td>AB034164</td>
<td>This report</td>
</tr>
<tr>
<td>C/Nara/82</td>
<td>NA82</td>
<td>MS</td>
<td>A</td>
<td>AB034165</td>
<td>This report</td>
</tr>
<tr>
<td>C/Kyoto/41/82</td>
<td>KY14182</td>
<td>MS</td>
<td>A</td>
<td>AB034166</td>
<td>This report</td>
</tr>
<tr>
<td>C/Hyogo/1/83</td>
<td>HY183</td>
<td>MS</td>
<td>A</td>
<td>AB034167</td>
<td>This report</td>
</tr>
<tr>
<td>C/England/83</td>
<td>ENG83</td>
<td>YA</td>
<td>B</td>
<td>D00034</td>
<td>Buonagurio et al. (1986)</td>
</tr>
<tr>
<td>C/Nara/1/85</td>
<td>NA185</td>
<td>MS</td>
<td>A</td>
<td>AB034168</td>
<td>This report</td>
</tr>
<tr>
<td>C/Nara/2/85</td>
<td>NA285</td>
<td>YA</td>
<td>B</td>
<td>AB034169</td>
<td>This report</td>
</tr>
<tr>
<td>C/Nara/1/86</td>
<td>NA186</td>
<td>AI#</td>
<td>B</td>
<td>AB034171</td>
<td>This report</td>
</tr>
<tr>
<td>C/Yamagata/1/86</td>
<td>YA186</td>
<td>AI#</td>
<td>B</td>
<td>AB034170</td>
<td>This report</td>
</tr>
<tr>
<td>C/Yamagata/1/88</td>
<td>YA188</td>
<td>YA#</td>
<td>B</td>
<td>AB002469</td>
<td>Hongo et al. (1992)</td>
</tr>
<tr>
<td>C/Yamagata/7/88</td>
<td>YA788</td>
<td>YA</td>
<td>B</td>
<td>AB034172</td>
<td>This report</td>
</tr>
<tr>
<td>C/Yamagata/9/88</td>
<td>YA988</td>
<td>AI</td>
<td>B</td>
<td>AB034173</td>
<td>This report</td>
</tr>
<tr>
<td>C/Yamagata/5/89</td>
<td>YA589</td>
<td>AI#</td>
<td>B</td>
<td>AB034174</td>
<td>This report</td>
</tr>
<tr>
<td>C/Yamagata/10/89</td>
<td>YA1089</td>
<td>YA</td>
<td>B</td>
<td>AB034175</td>
<td>This report</td>
</tr>
<tr>
<td>C/Yamagata/1/90</td>
<td>YA190</td>
<td>YA#</td>
<td>B</td>
<td>AB034176</td>
<td>This report</td>
</tr>
<tr>
<td>C/Miyagi/1/90</td>
<td>MI190</td>
<td>AI#</td>
<td>B</td>
<td>AB034177</td>
<td>This report</td>
</tr>
<tr>
<td>C/Miyagi/5/91</td>
<td>MI591</td>
<td>AI#</td>
<td>B</td>
<td>AB034178</td>
<td>This report</td>
</tr>
<tr>
<td>C/Miyagi/7/91</td>
<td>MI791</td>
<td>AI#</td>
<td>B</td>
<td>AB034179</td>
<td>This report</td>
</tr>
<tr>
<td>C/Miyagi/9/91</td>
<td>MI991</td>
<td>YA</td>
<td>B</td>
<td>AB034180</td>
<td>This report</td>
</tr>
<tr>
<td>C/Yamagata/5/92</td>
<td>YA592</td>
<td>MS</td>
<td>B</td>
<td>AB034181</td>
<td>This report</td>
</tr>
</tbody>
</table>

* The HE gene lineages to which the individual isolates belong were determined previously (Peng et al., 1996; Muraki et al., 1996; Kimura et al., 1997). AO, AO74-related lineage; AI, AI181-related lineage; YA, YA2681-related lineage; MS, MS80-related lineage.
† Determined in this study.
‡ The HE gene sequences of these virus strains have not yet been determined, but their HE antigenicity was indistinguishable from that of AI181 or YA2681.
ND, Not determined.

Inhibitors (Hongo et al., 1997) and immunoprecipitated as described previously (Sugawara et al., 1986), utilizing rabbit antisera against GST/NS1-C or GST/NS2-C. The immunoprecipitates obtained were analysed by SDS–PAGE on 17.5% gels containing 4 M urea under reducing conditions and processed for analysis by fluorography (Yokota et al., 1983).

Results

Nucleotide sequence of the NS gene

The NS gene sequences (nucleotides 24–905) were determined here for 27 influenza C virus strains isolated between 1964 and 1992 (see Table 1). Compared with the published sequence of the YA188 virus NS gene (Hongo et al., 1992), no deletions or insertions were detected in the coding region of any of the virus strains analysed. However, an insertion of one extra A residue was observed in the 5’ noncoding region (position 24) of the NS gene of strain MI190 and a deletion of two residues (positions 892 and 893) in the 3’ noncoding region of the NS genes of strains YA1089 and MI991. More importantly, the 27 NS genes sequenced were all found to possess four G residues at positions 698–701, as has been reported for the NS gene of YA188 (Hongo et al., 1992) as well...
as for those of TAY47, AA50, Johannesburg/1/66 and England/84 (Marschall et al., 1999). Therefore, we reexamined the sequences around nucleotide position 700 of the NS genes of six strains (TAY47, AA50, GL54, CAL78, MS80 and ENG83) that have been reported by Nakada et al. (1986) and Buonagurio et al. (1986) to contain only three G residues in this region. The representative sequencing patterns obtained are shown in Fig. 1(a). Clearly, all six NS genes investigated had four G residues around the area in question. Thus, it appears that the NS genes of most of the influenza C viruses listed in Table 1 are 935 nucleotides in length, with the exception of three genes composed of 936 (MI190) or 933 nucleotides (YA1089 and MI991).

Comparison of the NS gene sequences among the 34 isolates showed that nucleotide substitutions occurred at 86 positions (9.8%), including an insertion at position 24 and deletions at positions 892 and 893. However, the initiation codon (positions 28–30) and the termination codons for synthesis of NS1 (positions 766–768) and NS2 (positions 887–889), as well as the sequences around the 5’ and 3’ end junctions (positions 214 and 528) of the NS2 mRNA, were all strictly conserved among the 34 NS gene sequences, which suggests strongly that all influenza C virus strains analysed can synthesize NS1 and NS2 proteins of 246 (predicted molecular mass 28 kDa) and 182 amino acids (predicted molecular mass 21 kDa), respectively, according to the coding strategy illustrated in Fig. 1(b).

**Phylogenetic tree for NS genes**

A phylogenetic tree for the NS genes, constructed by using the nucleotide sequences of 27 virus strains determined here as well as the previously determined sequences of seven strains (Nakada et al., 1985; Buonagurio et al., 1986; Hongo et al., 1992) (nucleotide positions 24–41 of these strains, except for CAL78 and YA188, were sequenced in this study and were found to possess the same sequences as those of the other 29 strains, including CAL78 and YA188) is shown in Fig. 2. The 34 NS genes were split into two distinct groups, designated A and B. Group A contains six strains isolated between 1947 and 1980 in the USA (TAY47, AA50, GL54, CAL78 and MS80) and the Republic of South Africa (JHG67) in addition to 10 strains isolated in Japan between 1964 and 1983. It should be noted that, among the 34 strains analysed, the 12 isolates obtained before 1980 were all within this group. Group B contains 16 virus strains isolated in Japan during 1981–1992 as well as two foreign isolates from China (PB11581) and the UK (ENG83). Interestingly, the 15 isolates obtained after 1985 were all found to belong to this group. Pairwise comparison of the NS gene sequences among the 34 isolates showed that nucleotide sequence differences between the isolates of the same group were 0–1–2.5% (group A) and 0–1–8% (group B), while differences between the isolates belonging to different groups ranged from 1.5 to 4.0%.

There were several noticeable differences between the phylogenetic trees of the NS and HE genes. The HE gene of YA988, together with those of AI181 and Johannesburg/1/66, formed an AI181 virus lineage (Muraki et al., 1996). In the NS gene tree, however, not only YA988 but six strains isolated during 1986–1991 in Japan (YA186, NA186, YA589, MI190, MI591 and MI791), all of which had HE antigenicity indistinguishable from that of AI181 (Ohyama et al., 1992; Matsuoka et al., 1994; Kimura et al., 1997), were located within the branch cluster of group B, separate from AI181, which fell within group A. It was also impressive that the NS gene sequences of these seven 1986–1991 isolates (including YA988) were identical or closely similar to that of YA2681. These observations suggest strongly that the recent Japanese isolates that possess HE genes on the AI181 virus lineage have emerged by reassortment from two viruses closely related to AI181 and YA2681, having obtained a group B NS gene from the latter parent. We reported previously that strains YA2681, NA285 and YA788 possess HE genes highly homologous to those of strains SA71, CAL78 and KY179 (Muraki et al., 1996). It became evident here, however, that the former three isolates had group B NS genes, whereas the latter three had group A NS genes, which raised the possibility that the recent Japanese
isolates having HE genes of the YA2681 virus lineage are reassortants that acquired their HE and NS (group B) genes from an SA71-like parent and another parent as yet not identified. The HE genes of strains NA185 and YA592 (having group B NS genes), as well as those of strains NA82, KY4182 and HY183 (having group A NS genes), belong to the MS80 virus lineage (Muraki et al., 1996; Peng et al., 1996). By comparing partial sequences of the individual RNA segments among a limited number of virus strains, we provided evidence that suggests that MS80-like viruses reassorted separately with YA2681-like and PB11581-like viruses to generate NA185 and YA592, respectively (Peng et al., 1994, 1996). Here, it became clear that NA185 and YA592 had acquired group B NS genes by these reassortments from YA2681-like and PB11581-like parents, respectively.

**Deduced amino acid sequences of NS1 and NS2**

Deduced amino acid sequences of the NS1 and NS2 proteins were compared among the 34 isolates and the results are summarized in Fig. 3. Amino acid substitutions were observed at 24 positions (9.8%) among the 246 amino acid residues of NS1. The degrees of NS1 protein sequence identity between isolates belonging to the same group were 97–2–100% (group A) or 98–4–100% (group B), while those between isolates belonging to different groups ranged from 96–3 to 99–6%. It was impressive that the NS1 protein of the prototype strain TAY47 (group A) differed by only one or two amino acids from those of several group A strains (MS80, AI181, NA82, KY4182 and HY183) isolated at least 33 years later and that two strains belonging to different groups (HY183 of group A and MI991 of group B) that were isolated 8 years apart had NS1 proteins that differed by only one amino acid residue, indicating that the amino acid sequence of NS1 is highly conserved.

Among the 182 amino acid residues of NS2, substitutions occurred at 21 positions (11.3%). This protein exhibited an amino acid sequence similarity of 96–1–100% within each of groups A and B and similarity between the two groups ranged from 94–0 to 98–4%. Furthermore, it was found that the NS2 proteins of strains AA50 and MS80, isolated 30 years apart, had amino acid sequences that differed by only one residue and that the sequences of two strains (SA71 and CAL78) isolated 7 years apart in different countries (Japan and the USA) were completely identical to each other. All of these observations suggest that the NS2 amino acid sequence is conserved fairly well, although it is slightly more divergent than the sequence of NS1. Visual inspection of the NS2 sequence showed that it contains two heptad repeat motifs, at positions 84–105 (HR1) and 130–158 (HR2), the former being a leucine zipper consisting of three heptad repeats. HR1 was found to be conserved almost completely, except that KY4182 had a Lys to Ile change at position 92. In HR2, amino acid changes were detected at three positions, 142, 144 and 147. However, the
Fig. 3. Comparison of deduced amino acid sequences of the NS1 and NS2 proteins among 34 influenza C virus strains. Only amino acids that were different from those of TAY47 are shown.

Fig. 4. Synthesis of the NS1 and NS2 proteins in influenza C virus-infected cells. Mock-infected (lanes 2 and 5) and YA188 virus-infected (lanes 1, 3, 4 and 6) HMV-II cells were labelled with $^{35}$S methionine for 1 h at 12 h p.i. and then immunoprecipitated with pre-immune rabbit serum (lane 1), anti-GST/NS1-C serum (lanes 2 and 3), rabbit antiserum against AA50 virions (lane 4) or anti-GST/NS2-C serum (lanes 5 and 6) and the resulting precipitates were analysed by SDS-PAGE.

Detection of NS2 protein in influenza C virus-infected cells

In order to identify the NS2 protein unequivocally in influenza C virus-infected cells as well as to obtain direct evidence that NS1 and NS2 are synthesized according to the coding strategy shown in Fig. 1(b), we produced antisera against the GST fusion proteins GST/NS1-C and GST/NS2-C, which contain residues 225–246 of the 246 amino acid NS1 protein and residues 121–182 of the 182 amino acid NS2 protein, respectively. It should be stressed that the sequence of residues 225–246 of the 246 amino acid NS1 is completely different from that of the putative 286 amino acid NS1 and that residues 121–182 of the 182 amino acid NS2 protein, respectively. It should be stressed that the sequence of residues 225–246 of the 246 amino acid NS1 is completely different from that of the putative 286 amino acid NS1 and that residues 121–182 of the 182 amino acid NS2 are not contained in the putative 121 amino acid NS2 protein. YA188 virus-infected HMV-II cells were labelled with $^{35}$S methionine for 1 h at 12 h p.i. and then subjected to immunoprecipitation with each of these two antisera. As seen clearly in Fig. 4, immune
sera against GST/NS1-C and GST/NS2-C immunoprecipitated polypeptides with apparent molecular masses of 31 and 22 kDa, respectively.

Discussion

Four G residues were found invariably at nucleotide positions 698–701 of the NS genes of 34 influenza C virus strains analysed here, which included those of six strains (TAY47, AA50, GL54, CAL78, MS80 and ENG83) reported previously to contain only three G residues in this region (Nakada et al., 1985; Buonagurio et al., 1986), which indicates that RNA segment 7 of this virus has the potential to encode NS1 and NS2 proteins composed of 246 and 182 amino acids, respectively, as was first proposed by Hongo et al. (1996). This notion was substantiated by the finding that rabbit antisera against GST fusion proteins containing the C-terminal sequences of the 246 amino acid NS1 and 182 amino acid NS2 proteins (which are not present in the putative 286 amino acid NS1 and 121 amino acid NS2 proteins) immunoprecipitated the NS1 (molecular mass 31 kDa) and NS2 (molecular mass 22 kDa) proteins from lysates of YA188 virus-infected HVM-II cells. The same results were obtained with lysates of HVM-II cells infected with any of several other strains including CAL78 (data not shown). Previously, Nakada et al. (1986) reported synthesis of a protein of 15 kDa in AA50 virus-infected MDCK cells that was immunoprecipitated with antiserum to a synthetic pentadecapeptide corresponding to the C-terminal sequence of the putative 121 amino acid NS2 of CAL78 (residues 107–120 of a 182 amino acid NS2 followed by the C-terminal Lys). It seems unlikely, however, that this polypeptide represents NS2, since its molecular mass (15 kDa) was much smaller than that predicted for a 182 amino acid NS2 (21 kDa).

Previously, we provided evidence that suggested that reassortment of the genome between different influenza C virus strains occurs frequently in nature (Peng et al., 1994, 1996; Kimura et al., 1997; Tada et al., 1997). Here, we have obtained data that show that seven 1986–1991 isolates with HE genes on the AI181 virus lineage (YA186, NA186, YA988, YA589, MI190, MI591 and MI791) are all reassortants that inherited their HE and NS genes from AI181-like and YA2681-like viruses, respectively, and that all five 1981–1990 strains (YA2681, NA285, YA188, YA788 and YA190) with HE genes on the YA2681 virus lineage are also reassortants, which acquired HE genes from an SA71-like virus and NS genes from an unknown parent. These observations, together with those reported previously (Peng et al., 1996; Kimura et al., 1997), suggest that most if not all influenza C viruses currently circulating in Japan arose by reassortment, which presumably occurred either in the 1970s or in the early 1980s. A phylogenetic tree of the NS genes was unique in that 34 influenza C virus strains isolated during 1947–1992 were split into two different groups (A and B) and that the recent isolates, irrespective of their HE gene lineage, had group B NS genes, whereas the older ones had group A NS genes. In any of the trees for the other six genes, constructed on the basis of total (HE and M genes) or partial nucleotide sequences (PB2, PB1, P3 and NP), three or four distinct lineages were identified, each of which contained older isolates as well as recent ones (Peng et al., 1996; Muraki et al., 1996; Kimura et al., 1997; Tada et al., 1997). These observations lead us to postulate that influenza C viruses that acquired group B NS genes through the reassortment events described above dominantly replaced the parental viruses with group A genes, forming stable viral lineages.

The roles of NS1 and NS2 in influenza C virus replication are not known. The C-terminal half of influenza A virus NS1 is highly variable (Nakajima et al., 1990; Ludwig et al., 1991; Kawaoka et al., 1998; Suarez & Perdue, 1998) and may be dispensable, since its deletion does not lead to loss of virus infectivity (Norton et al., 1987). In contrast, the C-terminal half of influenza C virus NS1 is highly conserved, suggesting that the NS1 protein of this virus may be structurally and functionally different from that of influenza A virus. Indeed, it has been shown that, while the steady-state levels of M gene-derived spliced mRNAs are only 5–10% of that of the unspliced mRNA in influenza A virus-infected cells (Lamb et al., 1981), the predominant M gene-derived mRNA synthesized in influenza C virus-infected cells is a spliced one (Yamashita et al., 1988; Hongo et al., 1994), raising the possibility that influenza C virus NS1 may lack the ability to inhibit splicing of pre-mRNAs, one of the important functions identified for influenza A virus NS1 (Forbes et al., 1994; Lu et al., 1994). The functions of the influenza C virus NS gene products must be investigated extensively in future studies.

Interestingly, one and three amino acid positions were found in the NS1 and NS2 proteins, respectively, that differentiate between groups A and B. Group A NS1 proteins have Gln at position 212 in common, whereas all group B NS1 proteins have Arg at this position. This was the only amino acid difference detected between the NS1 proteins of strains NA82 (group A) and NA185 (group B) as well as between those of strains HY183 (group A) and M991 (group B). In the NS2 protein, group-specific amino acid differences were found at positions 76 (Lys in group A; Arg in group B), 108 (Lys or Arg in group A; Glu in group B) and 142 (Leu in group A; His in group B), the former two positions flanking the first heptad repeat motif HR1 (residues 84–105) and the last being located within the second motif HR2 (residues 130–158). It may be interesting to investigate the possibility that one or more of these amino acid changes might have caused the apparent increase in the epidemiological potential of influenza C viruses with group B NS genes compared with that of viruses with group A NS genes.

We thank Drs S. Nakada (Yamanouchi Pharmaceutical Co.) and N. J. Cox (CDC) for supplying several influenza C virus strains and T. Itoh
(Yamagata University Laboratory Animal Center) for excellent technical assistance. This work was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture, Japan.

References


Received 8 February 2000; Accepted 27 March 2000