JOURNAL OF VIROLOGY, Apr. 1990, p. 1808–1811
0022-538X/90/041808-04$02.00/0

Vol. 64, No. 4

Direct Sequencing of the HA Gene of Influenza (H3N2) Virus in Original Clinical Samples Reveals Sequence Identity with Mammalian Cell-Grown Virus

JACQUELINE M. KATZ,* MAOLIANG WANG,† AND ROBERT G. WEBSTER

Department of Virology and Molecular Biology, St. Jude Children's Research Hospital, 332 North Lauderdale, P.O. Box 318, Memphis, Tennessee 38101

Received 4 October 1989/Accepted 27 December 1989

When influenza (H3N2) viruses from infected individuals are grown in embryonated chicken eggs, viruses are isolated which differ antigenically and structurally from viruses grown in mammalian Madin-Darby canine kidney (MDCK) cell culture [G. C. Schild, J. S. Oxford, J. C. de Jong, and R. G. Webster, Nature (London) 303:706–709, 1983]. To determine which of these viruses is most representative of virus replicating in the infected individual, a region of the HA gene of virus present in original clinical samples was amplified by using the polymerase chain reaction and sequenced directly. Comparison of 170 amino acid residues of HA1 flanking and containing the receptor-binding site and antigenic sites indicated that over this region, the HA of virus replicating in the infected individual was identical to that of virus after growth in MDCK cells and was distinct from the HA of viruses grown in eggs. Therefore, cultivation of human influenza H3N2 virus in mammalian MDCK cells results in a virus similar to the predominant population of virus found in the infected individual.

Original clinical samples from patients infected with influenza viruses are routinely cultured in mammalian continuous cell lines or embryonated chicken eggs. It has been assumed that the antigenic and structural characteristics of viruses isolated in these host cells are identical to those of the virus which replicated and caused disease in the infected individual. However, in other viral systems some host cell types have been shown to impose selection on the replicating virus population, favoring a virus with amino acid mutations which provide a growth advantage for the virus in the particular host cell (2, 7).

Host cell variation of influenza viruses was first reported by Burnet and Bull (3) over 40 years ago when they observed that influenza viruses cultivated in the amniotic or allantoic cavities of eggs could exhibit altered receptor-binding properties relative to each other. In recent years, the molecular basis of this variation has been shown to reside largely in amino acid sequence changes in the region of the receptor-binding pocket of the HA molecule of influenza viruses grown in eggs (4, 10, 11).

When an original clinical sample is grown in eggs, multiple antigenically and structurally distinct variants can be isolated, whereas only a single population of virus is isolated from the same original sample grown in mammalian Madin-Darby canine kidney (MDCK) cell culture (5, 16). The predominant egg-grown virus is usually antigenically identical to MDCK cell-grown virus but, nevertheless, possesses a single amino acid difference compared to virus from the same source grown in mammalian cells. Typically, one or two other antigenically and structurally distinct variants are isolated in eggs. It is not known which of these viruses, if any, represent the virus replicating in the infected individual.

To address this question, we have sequenced directly a segment of the HA gene amplified by the polymerase chain reaction (PCR) (12, 13) technique from influenza (H3N2) virus present in an original throat wash sample. The region of the HA gene chosen for amplification encodes a 170-amino-acid stretch encompassing the receptor-binding site and antigenic regions around it. This region includes the majority of amino acid residues known to differ between egg-grown and MDCK cell-grown H3N2 viruses (4, 5, 16).

Throat washes (gargles) were obtained from two patients with clinical signs of influenza (virus strains A/Memphis/5/85 and A/Memphis/14/85). These samples were collected in phosphate-buffered saline to which 5% bovine serum albumin was added subsequently. Samples were aliquoted and stored at −70°C until use. Virus was concentrated from 2 to 4 ml of throat wash by centrifugation in a Beckman air-driven centrifuge (28 lb/in^2 gauge ≈ 90,000 rpm, 15 min at 20°C). Concentrated virus was resuspended in 0.25 ml of STE (0.05 M Tris, 0.1 M NaCl, 1.0 mM EDTA); the viral RNA was isolated by incubation (10 min at 20°C) of the virus with an equal volume of 2× proteinase K buffer (0.02 M Tris [pH 7.5], 0.01 M EDTA [pH 8.0], 1% sodium dodecyl sulfate) and 50 μl of proteinase K (2 mg/ml) followed by three extractions with phenol-chloroform and precipitation with ethanol (1). cDNA was synthesized in a volume of 10 μl from viral RNA by using 20 U of reverse transcriptase (Life Sciences), 1 μg of oligonucleotide primer identical to the 5′ end of the segment of the HA gene to be amplified (Fig. 1), 0.5 mM each 2′-deoxynucleotide 5′-triphosphate (dNTP), Schimke buffer (50 mM Tris, [pH 8.0], 10 mM $MgCl_2$, 10 mM dithiothreitol, 70 mM KCl) and 20 U of RNase inhibitor (Promega). Synthesis of cDNA was carried out at 42°C for 1 h, after which the samples were heated to 100°C for 5 min.

Each PCR tube contained 10 μl of cDNA template or mock template, 1 μM each of forward and reverse oligonucleotide primers (Fig. 1), 200 μM each dNTP, 50 mM KCl, 10 mM Tris (pH 8.3), 1.5 mM $MgCl_2$, 0.01% (wt/vol) gelatin, and 0.5 U of *Thermus aquaticus* (Taq) polymerase (Perkin Elmer Cetus). Reaction tubes were overlaid with 100 μl of mineral oil and were subjected to 35 cycles of amplification in a DNA thermal cycler (Perkin Elmer Cetus). The first cycle consisted of 1.5 min at 94°C, 2 min at 37°C, and 3 min

* Corresponding author.
† Present address: Virology Division, Fort Detrick-USAMRIID, Frederick, MD 21701.

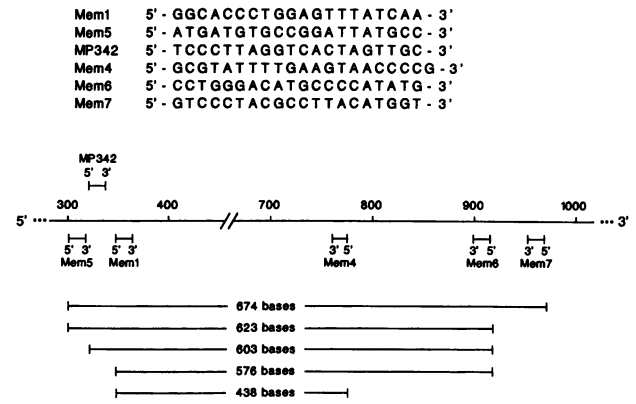

FIG. 1. Nucleotide sequences and position on the HA gene of oligonucleotide primers used for cDNA synthesis and for PCR amplification.

at 72°C. The next 34 cycles were similar except that denaturation at 94°C was carried out for 1 min. All experiments were performed by using equal numbers of potential positive samples and negative controls in which mock RNA extraction and cDNA synthesis had been carried out on samples of sterile water.

Figure 1 shows the sequences and position on the HA gene of the oligonucleotide primers used for cDNA synthesis and PCR amplification. Primer sequences were chosen on the basis of homologous nucleotide sequences obtained from viral RNA of MDCK cell- and egg-grown viruses. Three different positive-sense primers (Mem5, MP342, and Mem1) or negative-sense primers (Mem4, Mem6, and Mem7) were synthesized by using an Applied Biosystems 380A DNA synthesizer and phosphoramidite chemistry. The same positive-sense primer was used for both cDNA synthesis and PCR amplification. The primer combinations which optimized amplification differed between viruses, and therefore three to five independent amplifications were performed for each virus sample with different forward- and reverse-primer combinations. Depending on the primers used, amplification products ranged in size from 438 to 674 base pairs in length.

To detect PCR products, 20-μl samples were electrophoresed on 3% Nuseive, 1% Seakem agarose gels (FMC Bioproducts) containing 0.5 μg/μl ethidium bromide for 3.5 h at 60 V. The remainder of the PCR products (~80 μl) were subsequently run on 3% Nuseive, 1% Seakem agarose gels in TAE (40 mM Tris acetate, 1 mM EDTA [pH 6.2]), and the DNA bands were illuminated by UV light. Bands of the expected molecular sizes were excised from the gel, and their DNA was purified on a silica matrix (Geneclean; BIO 101 Inc.). DNA amplification products purified in this way were routinely used for two separate dideoxy-chain termination sequence determination reactions (14) using modified T7 DNA polymerase (United States Biochemical) and nested oligonucleotide primers which had been end labeled with [γ-^{32}P]ATP in the presence of T4 polynucleotide kinase (Bethesda Research Laboratories). A total of six different primers, five priming for positive-sense DNA and one priming for negative-sense DNA, were used to determine overlapping sequences of DNA products obtained from multiple independent amplification reactions. Therefore, any given sequence is the result of at least two independent determinations.

A 600-base-pair region of the HA gene was amplified from virus present in throat wash samples obtained from two individuals (patients 5 and 14) infected with type A (H3N2) virus in the 1985 influenza virus outbreak in Memphis, Tenn. These samples contained $10^{3.5}$ to $10^{4.5}$ 50% tissue culture infectious doses of virus on the basis of titration of the original samples on MDCK cells. For comparison, cDNA prepared from MDCK cell-grown virus or the predominant egg-grown variant diluted to a titer of virus similar to that in the original throat wash samples was also amplified. Care was taken not to work with MDCK cell- or egg-grown virus and the corresponding throat wash sample at the same time. As a positive control, the cultured virus isolated from patient 5 was amplified at the same time as the original gargle from patient 14 and vice versa. Since these viruses possessed a patient-specific amino acid difference at 173 (patient 14, Lys; patient 5, Arg), sequence analysis of the PCR products ensured that amplified DNA obtained from original throat wash samples was not merely a result of contamination from other reaction tubes.

TABLE 1. Amino acid sequence comparison of critical residues in HA1 obtained by direct sequencing of PCR amplification products of H3N2 influenza viruses from two patients

Patient no. and source of virus	Amino acid at HA1 residue no.						
	145	156	173	186	199	246	248
5							
Original throat wash	Asn	Glu	Arg	Ser	Ala	Asn	Thr
MDCK cell grown (5-1)	Asn	Glu	Arg	Ser	Ala	Asn	Thr
Egg grown (5-15)	Asn	Glu	Arg	Ser	Ala	Ser[a]	Thr
14							
Original throat wash	Asn	Glu	Lys	Ser	Ser	Asn	Thr
MDCK cell grown (14-1)	Asn	Glu	Lys	Ser	Ser	Asn	Thr
Egg grown (14-1-1)	Asn	Glu	Lys	Ser	Ser	Asn	Ile

[a] Underlined type indicates an amino acid sequence change.

By using nested primers for dideoxy-chain termination sequencing, the nucleotide sequence encoding a 170-amino-acid region of HA1 (from residues 130 to 300) was obtained from the purified products of amplification. Table 1 gives a comparison of sequences deduced from PCR amplification products at amino acid residues implicated in the adaptation of these viruses to growth in in vitro host cell systems. For comparison, Table 2 gives the range of antigenic variants obtained by in vitro cultivation of the original clinical samples from patients 5 and 14 and their relevant amino acid sequences deduced by sequencing from viral RNA directly. In both cases, antigenic group I viruses (5-15 and 14-1-1) were the predominant virus type isolated in eggs.

No silent base changes or base mutations due to Taq polymerase incorporation errors were detected. For patient 5, the nucleotide sequence obtained from virus in the original throat wash was identical over the entire region to the sequence obtained from the amplification of cDNA prepared from comparable amounts of MDCK cell-grown virus. The mutation present in HA1 of the most frequently isolated egg-grown variant 5-15 at residue 246 (Asn→Ser) was not detected in virus from the original sample, nor were three other mutations at residues 156, 186, and 199 identified by RNA sequencing of other minor egg-grown variants isolated from patient 5 (Table 2). Similar analysis of the viruses from patient 14 (Table 1) shows that the amino acid change at residue 248 (Thr→Ile) found in the predominant egg-grown variant 14-1-1 did not occur in the DNA amplified from virus present in throat washings from this patient. The sequence obtained from this original sample was again identical to the

TABLE 2. Summary of antigenic and amino acid sequence differences between MDCK cell- and egg-grown viruses isolated from two patients[a]

Patient no.	Host cell isolated in:	Virus clone no.	Antigenic group	Amino acid at HA1 residue no.						
				145	156	173	186	199	246	248
5	MDCK	5-1	I	Asn	Glu	Arg	Ser	Ala	Asn	Thr
	Egg	5-15	I	—[b]	—	—	—	—	Ser	—
	Egg	5-4	II	—	Lys	—	—	—	—	—
	Egg	5-6	III	—	—	—	Ile	Ser	—	—
14	MDCK	14-1	I	Asn	Glu	Lys	Ser	Ser	Asn	Thr
	Egg	14-1-1	I	—	—	—	—	—	—	Ile
	Egg	14-2-4	III	Lys	—	—	—	—	—	—
	Egg	14-2-6	IV	—	—	—	—	—	Ser	—

[a] These patients are described in reference 16. Corrections have been made to typographical errors made in original data. (i) Sequence changes were observed at residue 186, not 185, in HA1. (ii) Patient 5 viruses have Ala at 199 except egg-grown variant 5-6, which has Ser. (iii) Patient 14 viruses have Asn at 246 except egg-grown variant 14-2-6, which has Ser.

[b] —, No change compared to the amino acid sequence of MDCK cell-grown virus.

sequence obtained from amplification of the MDCK cell-grown clone isolated from patient 14. These results indicate that over the region studied, the HA of virus replicating in the infected individual is identical to that of the virus after growth in mammalian MDCK cells and is distinct from the HA of virus grown in eggs.

Amplification by using PCR of the genome of very low amounts ($10^{3.5}$ to $10^{4.5}$ 50% tissue culture infectious doses) of influenza virus present in original throat wash material from infected individuals has enabled the direct sequence analysis of the viruses replicating in humans. This amount of virus is at least 10- to 100-fold less than is readily detectable by hemagglutination assay. For two separate cases, we have shown that the predominant population of influenza virus present in the infected individual bears an HA gene identical to that of virus isolated in vitro in mammalian cell culture. These results supply direct evidence that mammalian (MDCK) cell culture provides a host system for the isolation of human influenza viruses which, unlike the embryonated egg, does not select for variants with structural alterations in the receptor-binding region of the HA gene which are not the predominant virus species present in the infected individual. Support for these findings comes from studies by Robertson et al. (9), who have reported similar sequence identity between the HA of a type B virus amplified from an original sample and the HA of the virus after growth in MDCK cells.

The sequence identity between the HA of MDCK cell-grown virus and virus in the original sample demonstrated here is consistent with the body of circumstantial evidence which has accumulated to suggest that mammalian cell-grown virus is more similar to the virus replicating in the infected individual than is egg-grown virus. We have previously shown that original throat wash samples contain 10- to 100-fold-greater titers of virus which grows in MDCK cells than virus which can grow in the amniotic cavities of eggs. Furthermore, multiple antigenic and structural variants were only obtained by growth of virus in eggs. Mammalian cell-grown viruses have also been shown to be antigenically more similar to the virus replicating in humans than their egg-grown counterpart on the basis of their greater ability to detect neutralizing and hemagglutination-inhibition antibody in postinfection human sera (8, 15).

We have shown that cultivation of human influenza H3N2 viruses in mammalian cell culture results in a virus similar to the predominant population of virus that is replicating in the infected individual. Previous studies have shown that MDCK cell-grown viruses are more broadly cross-reactive with related viruses isolated in previous years and can provide superior protection against infection in animal models compared to variant egg-grown viruses from the same source (6, 17). However, MDCK cells, while suitable for virus isolation and characterization, cannot be used for vaccine strain isolation or production of the vaccine itself. Since egg-grown influenza virus vaccines are likely to continue, it is obviously important to select the virus which is antigenically similar to the virus grown in MDCK cells and which has minimal differences in biological properties. Care must be taken to obtain this virus. Although it may be readily isolated at limit dilution, early passaging of uncloned populations of egg-grown virus may result in the overgrowth of this predominant population by a faster-growing minor variant. This can result in selection of a minor variant which is more likely to be antigenically distinct and possess altered receptor-binding properties compared to the predominant virus present in the original throat wash sample. To detect such minor subpopulations of variant viruses directly, it would be necessary to molecularly clone and sequence the products of DNA amplification from original throat wash samples.

The amplification and direct sequence analysis of the influenza virus genome isolated directly from humans has proven to be a powerful research tool with which the sequence of critical amino acid residues known to affect the antigenicity and biological properties of a virus can be determined in the absence of pressures imposed by an in vitro host cell system.

This work was supported by Public Health Service grant AI-20591 from the National Institute of Allergy and Infectious Diseases, Cancer Center Support (CORE) grant CA-21765-10, and American Lebanese Syrian Associated Charities.

We thank Tommy Mills and Sarah Young for technical assistance and Dayna Anderson for secretarial help.

LITERATURE CITED

1. **Bean, W. J., G. Sriram, and R. G. Webster.** 1983. Electrophoretic analysis of iodine-labeled influenza virus RNA segments. Anal. Biochem. **102**:228–232.
2. **Bolwell, C., A. L. Brown, P. V. Barnett, R. O. Campbell, B. E. Clarke, N. R. Parry, E. J. Auldridge, F. Brown, and D. J. Rowlands.** 1989. Host cell selection of antigenic variants of foot-and-mouth disease virus. J. Gen. Virol. **70**:45–57.
3. **Burnet, F. M., and D. R. Bull.** 1943. Changes in influenza virus associated with adaptation to passage in chick embryos. Aust. J. Exp. Biol. Med. Sci. **21**:55–69.
4. **Katz, J. M., C. W. Naeve, and R. G. Webster.** 1987. Host

cell-mediated variation in H3N2 influenza viruses. Virology **156**:386–395.
5. **Katz, J. M., and R. G. Webster.** 1988. Antigenic and structural characterization of multiple subpopulations of H3N2 influenza virus from an individual. Virology **165**:446–456.
6. **Katz, J. M., and R. G. Webster.** 1989. Efficacy of inactivated influenza A virus (H3N2) vaccines grown in mammalian cells or embryonated eggs. J. Infect. Dis. **160**:191–198.
7. **Mettenleiter, T. C., B. Lemniozi, N. Sugg, C. Schreurs, and T. Ben-Porat.** 1988. Host cell-specific growth advantage of pseudorabies virus with a deletion in the genome sequences encoding a structural glycoprotein. J. Virol. **62**:12–19.
8. **Oxford, J. S., T. Corcoran, R. Knott, J. Bates, O. Bartolomei, D. Major, R. W. Newman, P. Yates, J. Robertson, R. G. Webster, and G. C. Schild.** 1987. Serological studies with influenza A (H1N1) viruses cultivated in eggs or in a canine kidney cell line (MDCK). Bull. W.H.O. **65**:181–187.
9. **Robertson, J. S., J. S. Bootman, and D. Major.** 1989. The nature of influenza viruses in clinical specimens, p. 274–278. *In* D. Kolakofsky and B. Mahy (ed.), Genetics and pathogenicity of negative strand viruses. Elsevier Biomedical Press, New York.
10. **Robertson, J. S., J. S. Bootman, R. Newman, J. S. Oxford, R. S. Daniels, R. G. Webster, and G. C. Schild.** 1987. Structural changes in the haemagglutinin which accompany egg adaptation of an influenza A (H1N1) virus. Virology **160**:31–37.
11. **Robertson, J. S., C. W. Naeve, R. G. Webster, J. S. Bootman, R. Newman, and G. C. Schild.** 1985. Alterations in the hemagglutinin associated with adaptation of influenza B virus to growth in eggs. Virology **143**:166–174.
12. **Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich.** 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239**:487–491.
13. **Saiki, R. K., S. Scharf, F. Faloona, K. B. Mullis, G. T. Horn, H. A. Erlich, and N. Arnheim.** 1985. Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science **230**:1350–1354.
14. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
15. **Schild, G. C., J. S. Oxford, J. C. de Jong, and R. G. Webster.** 1983. Evidence for host-cell selection of influenza virus antigenic variants. Nature (London) **303**:706–709.
16. **Wang, M., J. M. Katz, and R. G. Webster.** 1989. Extensive heterogeneity in the hemagglutinin of egg-grown influenza viruses from different patients. Virology **171**:275–279.
17. **Wood, J. M., J. S. Oxford, U. Dunleavy, R. W. Newman, D. Major, and J. S. Robertson.** 1989. Influenza A (H1N1) vaccine efficacy in animal models is influenced by two amino acid substitutions in the haemagglutinin molecule. Virology **171**: 214–221.