sample buffer, were then applied to an SDS–polyacrylamide gel (10% to 15% polyacrylamide gradient) and subjected to electrophoresis as described (*1*). Electroblot transfer of the separated proteins onto Schleicher & Schull nitrocellulose paper was carried out with the use of a Bio-Rad apparatus. Blots were immersed in Denhardt's solution [0.3% Ficoll 400, 0.3% radioimmunoassay (RIA)-grade bovine serum albumin (Sigma), and 0.3% polyvinylpyrollidone] for 2 hours and in 1% RIA-grade bovine serum albumin (BSA) (dissolved in Dulbecco's phosphate-buffered saline) for 1 hour. Rabbit antiserum to mouse cachectin was diluted 1:200 in a solution containing 0.85% NaCl, 0.05% Tween 20, 1% RIA-grade BSA, and 0.01*M* tris-Cl buffer, *p*H 8.0 (TTB). The blot was incubated with the antiserum for 3 hours, and then rinsed three times with a solution of 0.05% Triton X-100 and 2% SDS in water. Affinity-purified goat antiserum to rabbit immunoglobulin (Miles-Yeda, Elkhart, IN) was then diluted 1:1000 in TTB, and applied to the blot for 3 hours. The blots were rinsed again with Triton/SDS solution and exposed to ^{125}I-labeled protein A (New England Nuclear) diluted to a concentration of 0.15 μCi/ml in TTB. After 3 hours the blots were rinsed with Triton/SDS solution, then with distilled water, and were dried and used to produce autoradiograms.

24. BALB/c mice were injected intraperitoneally with 3 ml of sterile Brewer's thioglycollate broth (Difco). After 5 days, macrophages were harvested by peritoneal lavage with sterile Hanks balanced salt solution (HBSS), washed once in the same, and plated at a confluent density in 3-cm tissue culture dishes (Becton-Dickinson) with RPMI 1640 medium supplemented with 5% fetal bovine serum (Gibco). After 1 hour, the adherent cells were washed twice with serum-free RPMI 1640 and stimulated by addition of *Escherichia coli* strain 0127:B8 lipopolysaccharide (LPS; Difco) (final concentration 1 μg/ml). Control monolayers did not receive LPS. After 16 hours, the medium was aspirated for use in immunoblotting (Fig. 1B) and the cells were lysed with 5 ml of a solution containing 4*M* guanidinium isothiocyanate, 5 m*M* sodium citrate, *p*H 7.0, 0.1*M* β-mercaptoethanol, and 0.5% Sarcosyl (Pharmacia). CsCl (2 g) was added to each sample, and total cellular RNA was isolated by centrifugation over a 5.7*M* CsCl cushion. RNA (approximately 2 μg per lane) was subjected to electrophoresis in a 1.2% agarose gel containing 2.2*M* formaldehyde as a denaturant. Nitrocellulose blots were allowed to hybridize with a nick-translated pUC-9 plasmid containing a mouse cachectin complementary DNA insert (specific activity = 10^8 dpm/μg). Hybridization was carried out for 12 hours at 43°C in the presence of 50% formamide and 10% dextran sulfate. Blots were then washed with two changes of 2× standard saline citrate (SSC) containing 0.1% SDS, and with two changes of 0.1× SSC containing 0.1% SDS at 60°C, and used in autoradiography.

25. J. Weber, W. Jelinek, J. E. Darnell, Jr., *Cell* **10**, 611 (1977).

26. We thank M. Salditt-Georgieff for help in performing the nuclear transcription assays. Supported by NIH grants AM01314 and AI21359 and Rockefeller Foundation grant RF 84077.

24 October 1985; accepted 14 March 1986

Evolution of Human Influenza A Viruses over 50 Years: Rapid, Uniform Rate of Change in NS Gene

DEBORAH A. BUONAGURIO, SUSUMU NAKADA, JEFFREY D. PARVIN, MARK KRYSTAL, PETER PALESE, WALTER M. FITCH

Variation in influenza A viruses was examined by comparison of nucleotide sequences of the NS gene (890 bases) of 15 human viruses isolated over 53 years (1933 to 1985). Changes in the genes accumulate with time, and an evolutionary tree based on the maximum parsimony method can be constructed. The evolutionary rate is approximately 2×10^{-3} substitution per site per year in the NS genes, which is about 10^6 times the evolutionary rate of germline genes in mammals. This uniform and rapid rate of evolution in the NS gene is a good molecular clock and is compatible with the hypothesis that positive selection is operating on the hemagglutinin (or perhaps some other viral genes) to preserve random mutations in the NS gene.

INFLUENZA A VIRUSES HAVE A SINGLE-stranded RNA genome of eight segments of negative polarity, with the shortest segment coding for the nonstructural proteins (NS1 and NS2) (*1*). Figure 1 shows the nucleotide sequences of the NS genes of 15 human influenza A virus strains. The viruses were isolated over a 53-year period and represent all three human hemagglutinin serotypes (H1, H2, and H3). Except for the three Houston isolates, the strains were obtained from diverse geographical locations. The 15 sequences are easily aligned for analysis because of the size conservation of the NS gene segment of 890 bases. Nucleotide substitutions occur at 149 positions scattered throughout the gene and usually, once a base change is observed in a virus isolate, it is found in subsequent strains.

The sequence information as presented in Fig. 1 was analyzed by maximum parsimony (*2*) to determine the phylogenetic tree of minimum length. The best tree found contains a total of 186 substitutions and is illustrated in Fig. 2. The parsimony method also yielded four alternative trees containing 187 substitutions. These alternative trees contain only minor branch perturbations of the best tree.

Figure 3 shows the number of nucleotide substitutions between the origin of the best tree and the tip of each branch (Fig. 2) plotted against the date of isolation of the viruses whose NS gene is represented by that tip. The major line, derived by linear regression analysis, shows that these sequences are evolving at the steady rate of 1.73 ± 0.08 nucleotide substitutions per year, or $1.94 \pm 0.09 \times 10^{-3}$ substitution per nucleotide site per year. The WSN/33 and PR/34 strains appear to have more substitutions per year than expected and therefore were excluded from the evolutionary rate calculation. Since these strains were isolated before refrigeration became available in the laboratory, we believe that continuous passaging in animal hosts and in embryonated eggs (particularly in the first 10 to 15 years after isolation of the strains) may have introduced additional mutations not present in the original isolates. Figure 3 also shows that the group of H1N1 subtype strains, which reemerged in the human population in 1977 and after a 27-year absence (*3*), is evolving at the same rate. These "new" H1N1 viruses have been cocirculating with the H3N2 viruses since 1977 and form a separate evolutionary branch (Fig. 2). In reality, the H1N1 branch should be directly connected to the FW/50 branch of the main tree, since there are only five nucleotide differences between the FW/50 and USSR/77 virus NS genes. However, the viruses were isolated 27 years apart and, on the basis of the calculated evolutionary rate of 1.73 substitutions per year, we would predict approximately 46 additional substitutions in the NS gene of USSR/77 (represented by the broken line in Fig. 2). The observed data thus suggest a unique epidemiology of the new H1N1 isolates.

Several points can be made from the analysis of the data. First, calibration of the molecular clock is not affected by inaccurate paleontological dates, since the time of fossilization (isolation) of these strains is recorded. This may partly explain why the NS gene of influenza A viruses behaves as an accurate molecular clock (*4*). Thus, given only the NS gene sequence of a main line isolate, one can closely estimate the year of its isolation (Fig. 3). Although fewer points are available for measuring the rate in the new H1N1 strains (1977 to 1985), the data (filled squares in Fig. 3) are compatible with a molecular clock ticking at the same evolutionary rate for these NS genes. The mutations seen in the NS genes of the new H1N1 strains (1977 to 1985) are different from those seen in the 1950–1957 H1N1 strains.

The second point that can be made is that

D. A. Buonagurio, S. Nakada, J. D. Parvin, M. Krystal, P. Palese, Department of Microbiology, Mount Sinai School of Medicine, City University of New York, New York, NY 10029.
W. M. Fitch, Department of Physiological Chemistry, University of Wisconsin, Center for Health Sciences, Madison, WI 53706.

the evolutionary rate of $1.94 \pm 0.09 \times 10^{-3}$ substitution per site per year appears to be approximately 10^6 times the rate for mammalian germline genes (*5*). (The difference calculated on a per generation basis would be much less.)

Third, we have concentrated on an analysis of the NS genes because there was evidence that the NS gene was not exchanged (*6*) during the reassortment of influenza viruses leading to new A virus subtypes (*7*). Furthermore, preliminary data from several different laboratories had suggested that substitutions in influenza virus NS genes (as well as other genes) are generally retained when strains obtained at later times are compared with earlier isolates (*8*). We have compared NS gene sequences of nine new strains with six previously published sequences (*8*). Thus we can rely on a much larger data base and on isolates obtained over a longer period of time than was previously possible.

Yet another point concerns the long, slender nature of the evolutionary tree of the NS genes. This appears to be a consequence of the short life-span of any lineage other than the one that gives rise to the future generations (the average age of the side branches is only 3 years). This is reminiscent of periodic selection in chemostats (*9*), where the number of accumulated mutants in *Escherichia coli* suddenly drops to zero when a new favorable mutant replaces the older strains, so that at any point in time all coexisting

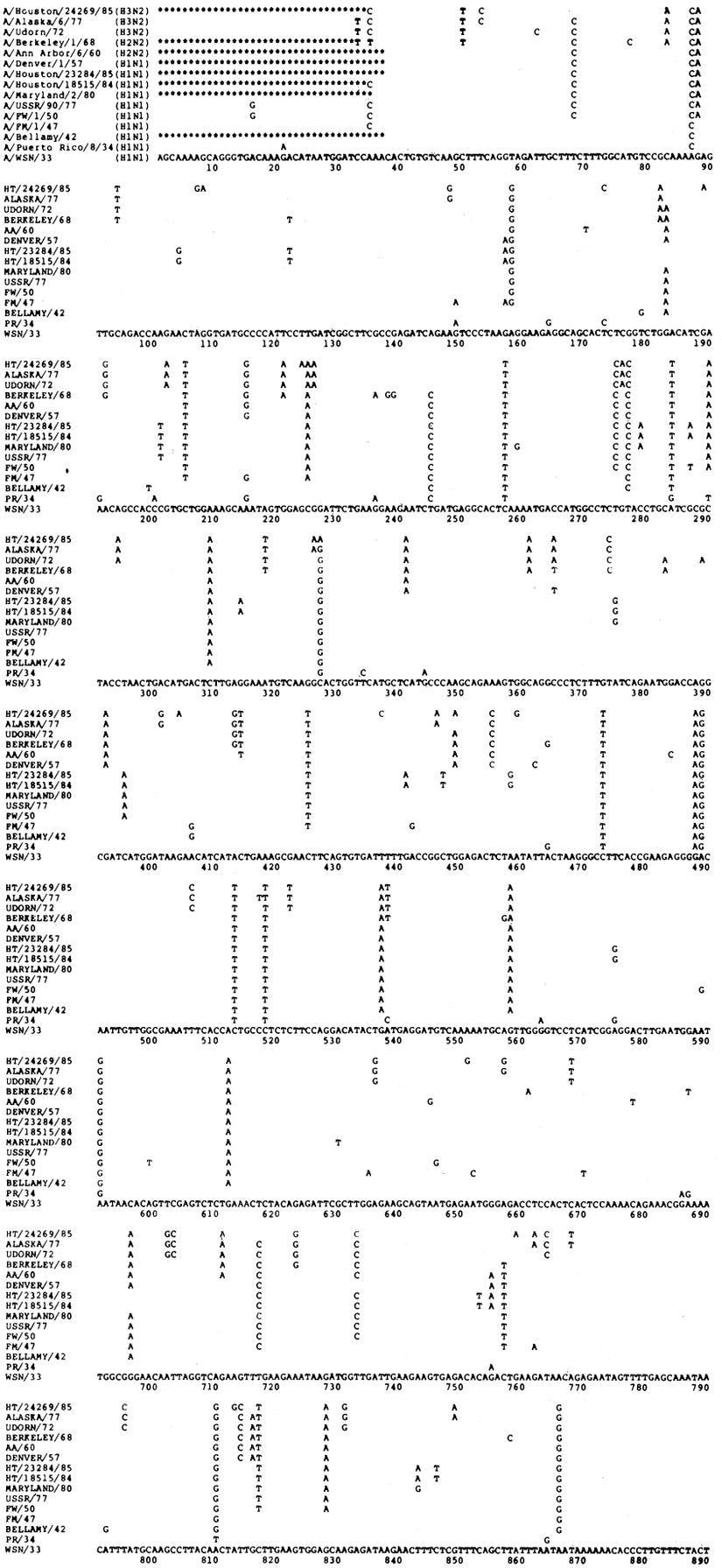

Fig. 1. Nucleotide sequences of 15 NS genes of human influenza A viruses. The NS sequences of the A/WSN/33, A/Bellamy/42, A/Maryland/2/80, A/Houston/18515/84, A/Houston/23284/85, A/Denver/1/57, A/Ann Arbor/6/60, A/Berkeley/1/68, and A/Houston/24269/85 viruses were determined by direct RNA sequencing of purified viral RNA by the dideoxy chain terminator method (*19*). Sequencing primers were oligonucleotides 15 to 20 bases in length synthesized with the phosphoramidite or phosphotriester chemistry on the Biosearch SAM One automated DNA synthesizer (*11*). The primers correspond to the following positions of the NS gene: 10–29, 133–147, 240–255, 293–312, 344–359, 518–536, 593–612, 616–632, and 739–755. Nucleotide sequences were stored, edited, and compared in an IBM 370 computer at the City University of New York by using published programs (*20*). The sequences of the remaining NS genes have been published (*8, 21*). The NS sequence of the oldest isolate, A/WSN/33, is shown in its entirety and serves as the reference sequence. Only nucleotide positions of the other strains that differ from the reference sequence are listed. The NS sequences of the A/Houston/18515/84 and A/Houston/23284/85 viruses are identical. Asterisks represent nucleotide positions that could not be determined because the first primer used for sequencing corresponds to positions 10 to 29 of the NS gene. The 5′ terminal nucleotides of the A/WSN/33 NS gene were determined by dideoxy sequencing with an M13 complementary DNA clone and the universal primer.

strains are recently diverged from the main lineage. Animal genes and genes from viruses, except possibly those of enterovirus 70 (*10*), do not follow this pattern. Rather, they appear to have multiple surviving lineages undergoing slower change, for example, the influenza C viruses in man (*11*). In addition, the length of time that a vaccine is effective against a viral pathogen may be correlated with the evolutionary rate of the virus. Vaccines for all three poliovirus types are made with isolates obtained approximately five decades ago. Similarly, the yellow fever virus vaccine was developed more than 50 years ago and current isolates have not sufficiently changed to warrant a new vaccine formula. Also, the currently used rabies vaccine strain goes back to Pasteur's time, whereas the influenza A viruses used for vaccine manufacture are changed every 2 to 3 years (even in the absence of a subtype shift) to accommodate changes in the prevalent strains.

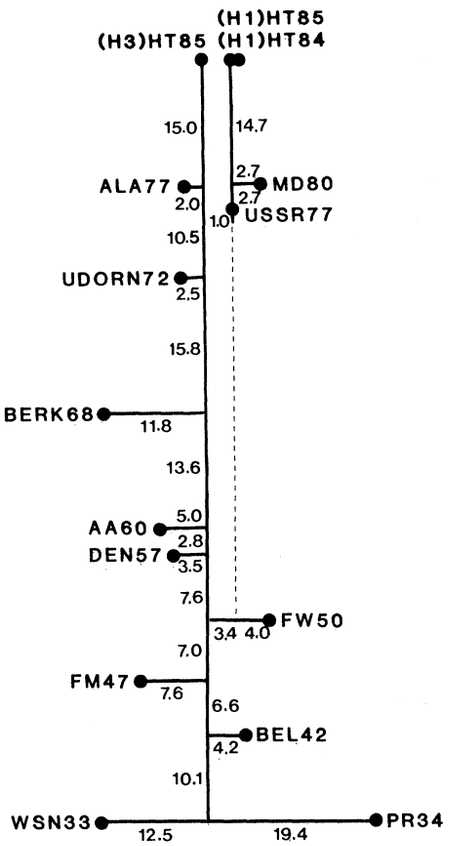

Fig. 2. Most parsimonious evolutionary tree for 15 influenza A virus NS genes. The nucleotide sequences shown in Fig. 1 were analyzed by the method of Fitch (*2*). The length of the trunk and the side branches of the evolutionary tree are proportional to the minimal number of substitutions required to account for the differences in sequence. Nonintegral numbers arise from averaging over all possible minimal solutions. The broken line represents the predicted number of additional substitutions between the NS genes of FW/50 and USSR/77 based on the calculated evolutionary rate. The abbreviations PR34, BEL42, FM47, FW50, DEN57, AA60, BERK68, ALA77, (H3)HT85, MD80, USSR77, (H1)HT84, and (H1)HT85 are used for influenza viruses A/Puerto Rico/8/34, A/Bellamy/42, A/Fort Monmouth/1/47, A/Fort Warren/1/50, A/Denver/1/57, A/Ann Arbor/6/60, A/Berkeley/1/68, A/Alaska/6/77, A/Houston/24269/85, A/Maryland/2/80, A/USSR/90/77, A/Houston/18515/84, and A/Houston/23284/85, respectively.

What, then, could explain the rapid evolutionary change in influenza virus genes? One possibility is that the unusually high variation in influenza A viruses in nature is the result of a high mutation rate (*12*). However, many other viruses, including vesicular stomatitis virus (*13*), parainfluenza virus (*14*), foot-and-mouth disease virus (*15*), RNA tumor viruses (*16*), and herpesviruses (*17*) also have the potential to undergo rapid genetic change. Moreover, a high rate of change may not explain the "slender" nature of the genealogy.

An explanation for the unusual pattern of NS gene evolution may be found in positive selection of influenza A virus variants. There is no evidence for immune surveillance of the NS gene products, nor have other selectional forces on the NS gene been identified. It may be that only one influenza virus gene (most likely the hemagglutinin) needs to be subject to selection. In the brief time before immunity develops to a new (antigenic) variant, that strain may sweep through the population, carrying with it whatever variant of the NS gene happens to be present. In this way, the NS gene's phylogeny may be linked to that of another gene undergoing extensive positive selection. The fixation of substitutions in the NS genes is not, therefore, simply the result of random genetic drift, and one probably cannot calculate a "neutral mutation rate" (*18*) for the A virus NS gene from these data. We suggest that positive selection represents a significant factor in the unusual pattern and high rate of influenza A virus evolution in humans.

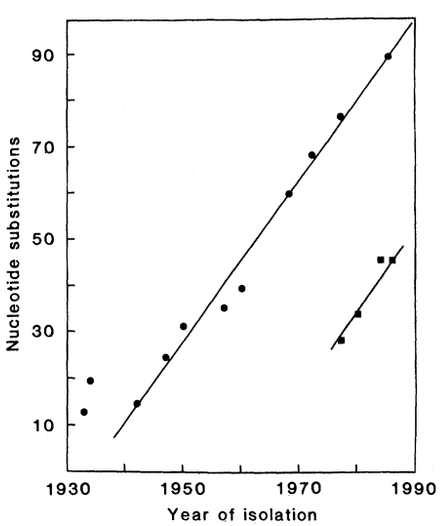

Fig. 3. Linearity with time of number of substitutions in the NS genes of influenza A viruses. The abscissa represents the year of isolation of the influenza A viruses used in the analysis. The ordinate indicates the number of substitutions observed in their NS genes between the first branching point formed by the WSN/33 and PR/34 sequences in Fig. 2 and the tips of all branches of the evolutionary tree. A line, generated by linear regression analysis, is drawn through the points. The slope of the line is 1.73 ± 0.08 substitutions per year. In addition to the sequences found on the trunk of the evolutionary tree (filled circles), the NS genes of the four new H1N1 viruses are also represented in this graph (filled squares). The line through the squares is arbitrarily drawn parallel to the other line.

REFERENCES AND NOTES

1. R. A. Lamb, in *Genetics of Influenza Viruses*, P. Palese and D. W. Kingsbury, Eds. (Springer-Verlag, Vienna, 1983), pp. 21–69.
2. W. M. Fitch, *Syst. Zool.* **20**, 406 (1971).
3. A. P. Kendal *et al.*, *Virology* **89**, 632 (1978); K. Nakajima, U. Desselberger, P. Palese, *Nature (London)* **274**, 334 (1978).
4. A. C. Wilson, S. S. Carlson, T. J. White, *Annu. Rev. Biochem.* **46**, 573 (1977).
5. W.-H. Li and T. Gojobori, *Mol. Biol. Evol.* **1**, 94 (1983).
6. C. Scholtissek *et al.*, *Virology* **87**, 13 (1978).
7. R. G. Webster, W. G. Laver, G. M. Air, G. C. Schild, *Nature (London)* **296**, 115 (1982).
8. G. M. Air, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 7639 (1981); G. W. Both, M. J. Sleigh, N. J. Cox, A. P. Kendal, *J. Virol.* **48**, 52 (1983); H. Hayashida, H. Toh, R. Kikuno, T. Miyata, *Mol. Biol. Evol.* **2**, 289 (1985); M. Nei, in *Evolution of Genes and Proteins*, M. Nei and R. K. Koehn, Eds. (Sinauer, Sunderland, MA, 1983), pp. 165–190; F. L. Raymond, A. J. Caton, N. J. Cox, A. P. Kendal, G. G. Brownlee, *Nucleic Acids Res.* **11**, 7191 (1983); J. J. Skehel, R. S. Daniels, A. R. Douglas, D. C. Wiley, *Bull. WHO* **61**, 671 (1983); C. Martinez, L. Del Rio, A. Portela, E. Domingo, J. Ortin, *Virology* **130**, 539 (1983); J. Ortin *et al.*, *Gene* **23**, 233 (1983); M. Krystal, D. A. Buonagurio, J. F. Young, P. Palese, *J. Virol.* **45**, 547 (1983); D. A. Buonagurio, M. Krystal, P. Palese, D. C. DeBorde, H. F. Maassab, *ibid.* **49**, 418 (1984).
9. K. C. Atwood, L. K. Schneider, F. J. Ryan, *Proc. Natl. Acad. Sci. U.S.A.* **37**, 146 (1951).
10. M. Tanimura, K. Miyamura, N. Takeda, *Jpn. J. Genet.* **60**, 137 (1985).
11. D. A. Buonagurio, S. Nakada, U. Desselberger, M. Krystal, P. Palese, *Virology* **146**, 221 (1985).
12. Influenza Workshop, *J. Infect. Dis.* **128**, 361 (1973).
13. J. Holland *et al.*, *Science* **215**, 1577 (1982); K. R. Spindler, F. M. Horodyski, J. J. Holland, *Virology* **119**, 96 (1982); P. J. O'Hara, S. T. Nichol, F. M. Horodyski, J. J. Holland, *Cell* **36**, 915 (1984).
14. A. Portner, R. G. Webster, W. J. Bean, *Virology* **104**, 235 (1980).
15. F. Sobrino, M. Davila, J. Ortin, E. Domingo, *ibid.* **128**, 310 (1983).
16. D. A. Zarling and H. M. Temin, *J. Virol.* **17**, 74 (1976); T. Gojobori and S. Yokoyama, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 4198 (1985).
17. T. C. Holland, S. D. Marlin, M. Levine, J. Glorioso, *J. Virol.* **45**, 672 (1983).
18. J. F. Crow and M. Kimura, *An Introduction to Population Genetics Theory* (Harper & Row, New York, 1970), pp. 368–369.
19. F. Sanger, S. Nicklen, A. R. Coulson, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463 (1977).
20. R. Staden, *Nucleic Acids Res.* **6**, 2601 (1979).
21. M. Baez *et al.*, *ibid.* **8**, 5845 (1980); R. A. Lamb and C.-J. Lai, *Cell* **21**, 475 (1980).

22. We thank J. L. Schulman for helpful discussions. Most of the influenza virus isolates were provided by A. P. Kendal and H. Six. This work was supported in part by PHS grants AI-11823 (P.P.) and AI-18998 (P.P.) and NSF grant BSR-8400682 (W.M.F.). M.K. is the recipient of a Sinsheimer Scholar award.

1 November 1985; accepted 21 February 1986