JOURNAL OF VIROLOGY, Oct. 1976, p. 307-313

Vol. 20, No. 1
Printed in U.S.A.

Mapping of the Influenza Virus Genome

III. Identification of Genes Coding for Nucleoprotein, Membrane Protein, and Nonstructural Protein

MARY B. RITCHEY, PETER PALESE,* AND JEROME L. SCHULMAN

Department of Microbiology, Mount Sinai School of Medicine of the City University of New York, New York, New York 10029

Received for publication 17 May 1976

In previous communications we reported that the eight RNA segments of influenza A/PR/8/34 (HON1) virus could be distinguished from corresponding segments of influenza A/Hong Kong/8/68 (H3N2) virus by migration on polyacrylamide-urea gels. Examination of the RNA patterns of the two parent viruses and recombinants derived from them in concert with serological identification of surface proteins and analysis of the other proteins on sodium dodecyl sulfate gradient gels permitted the identification of the genes coding for hemagglutinin, neuraminidase, and the P1, P2, and P3 proteins (Palese and Schulman, 1976; P. Palese et al., Virology, in press). In the present report we have extended these observations using similar techniques to examine other recombinants and have identified the genes coding for the remaining virus-specific proteins. We demonstrate that segment 5 of PR8 virus (counting the slowest moving RNA segment as 1) and segment 6 of Hong Kong virus code for the respective nucleoproteins, and that segment 7 of both viruses codes for the membrane protein and RNA segment 8 codes for the nonstructural protein. This completes the mapping of the influenza A virus genome.

In previous communications we reported that the eight RNA segments of influenza A/PR/8/34 (HON1) virus (PR8 virus) had migration rates on polyacrylamide-urea gels that could be distinguished from those of corresponding RNA segments of influenza A/Hong Kong/8/68 (H3N2) virus (HK virus) (P. Palese, M. B. Ritchey, and J. L. Schulman, Virology, in press; 5, 6). By examination of the RNA patterns of antigenically characterized recombinant viruses it was possible to determine from which parent each gene was derived and to identify the 4th-slowest-moving segment of both viruses as the gene for hemagglutinin and segment 5 of HK virus and segment 6 of PR8 virus as the genes for the respective neuraminidases (6).

Next, we observed that the remaining (nonsurface) proteins of PR8 virus could be distinguished from corresponding proteins of HK virus by analysis on sodium dodecyl sulfate gradient-polyacrylamide gels (M. B. Ritchey, P. Palese, and J. L. Schulman, Virology, in press). From analysis of the parent viruses and appropriate recombinants with respect to both RNA and protein patterns we found that the slowest-moving RNA (RNA 1) codes for the P3 protein, RNA 2 codes for P1, and RNA 3 codes for the second largest P protein, P2 (Palese et al., Virology, in press).

In the present communication we describe how similar analysis of other recombinants has made it possible to identify the RNA segments coding for nucleoprotein, membrane protein, and nonstructural protein (NS) and thus complete the mapping of the influenza A virus genome.

MATERIALS AND METHODS

Viruses. The two parent viruses, influenza A/PR/8/34 (HON1) and influenza A/Hong Kong/8/68 (H3N2) (PR8 and HK virus) have been described previously (5, 6, 8, 9). The three recombinants used in the present experiments were derived as follows. Recombinant 1, influenza A/Hong Kong/8/68 (H3)-A/Hong Kong/8/68 (N2) (HK-HK virus), was obtained from a mixed infection of MDCK (canine kidney) cells with HK and PR8 viruses following procedures described previously (Palese et al., Virology, in press; 5, 6, 9). Plaques were picked from MDCK cells infected with the recombinational mixture in the presence of antibody to PR8 hemagglutinin (HO) and HK neuraminidase (N2) (anti-HON2) incorporated in the agar overlay. Recombinant 1 was isolated after a second plaque to plaque passage in MDCK cells, this time without antiserum in the agar overlay. This procedure was expected to yield an H3N1 recombinant, but recombinant 1 was found

to contain both the hemagglutinin and neuraminidase of the HK parent virus (H3N2). Nevertheless, it was included in the present studies because its replication in embryonated eggs was sufficiently different from that of the parent HK virus to suggest that it was a recombinant.

Recombinant 2, which contains PR8 (HO) hemagglutinin and HK (N2) neuraminidase, PR8-HK (HON2) virus, was obtained after mixed infection of eggs with a mixture of HK virus and UV-irradiated PR8 virus. Recombinants selected after mixed infection in which one parent is subjected to UV irradiation have been shown to derive most of their genes from the nonirradiated parent virus (9). Thus recombinant 2 derives most of its genes from the HK virus parent.

Recombinant 3, also a PR8-HK (HON2) virus, was obtained from the same recombinational mixture as recombinant 1 after selection of plaques in the presence of antiserum to H3N1 in the agar overlay.

All the viruses employed were purified by plaque to plaque passage in MDCK cells and then injected into 11-day-old embryonated eggs to prepare seed virus. Antigenic analysis of all viruses was done according to published procedures (1, 4).

RNA and proteins. Labeling of virus-specific RNA with [^{32}P]phosphate (as orthophosphoric acid in sterile water, New England Nuclear) was done in MDCK cells by published procedures (6). Purification of the virus and isolation of the RNA followed standard techniques (5, 8, 9). Separation of RNAs on 2.6% urea-polyacrylamide gels was performed by methods described by Floyd et al. (3) and as modified by us (5, 6, 8, 9). In contrast to conditions used previously, electrophoresis at elevated temperatures (25°C) permitted the separation of the 7th RNA segments of PR8 and HK viruses (Palese et al., Virology, in press).

Virus-specific proteins were labeled with [^{35}S]methionine for 15 min at 6 h postinfection in MDCK cells (Ritchey et al., Virology, in press). Infected cell extracts were analyzed on 5 to 13% sodium dodecyl sulfate gradient-polyacrylamide gels as described (Ritchey et al., Virology, in press; 7). These methods separate and distinguish all proteins of these two viruses except for neuraminidase, which can be distinguished by serological analysis (1).

RESULTS

Figure 1 demonstrates the RNA patterns of PR8 virus (lanes 1 and 6), HK virus (lane 2), and three recombinant viruses (lanes 3 to 5). Comparison of the RNA patterns of the parent viruses (lanes 1 and 2) confirms previous observations that all eight RNA segments of PR8 virus migrate differently from comparable segments of HK virus. Analysis of the RNA pattern of recombinant 1 (H3N2) (lane 3) reveals that it derives all of its RNA segments from HK virus (lane 2), except for the fastest-moving segment (RNA 8). Examination of the protein pattern of recombinant 1 (Fig. 2, lane 1) demonstrates that all proteins co-migrate with equivalent proteins of HK virus (see arrows in lane 1), except for the NS protein which co-migrates with the NS protein of PR8 virus (lane 2). RNA and protein analysis of recombinant 1 thus reveals that RNA segment 8 codes for the NS protein.

Although the conditions under which this gel was run do not separate HK and PR8 virus hemagglutinins, recombinant 2 was shown serologically to derive its hemagglutinin from PR8 virus, which is consistent with the observation that its RNA segment 4 is derived from PR8 virus (Fig. 1, lane 4) (6). Because recombinant 2 derives only one other RNA segment (RNA 7) (Fig. 1, lane 4) from PR8 virus, analysis of the protein pattern in Fig. 2 (lane 4) establishes that RNA 7 codes for the membrane protein.

Comparison of the RNA pattern of recombinant 3, another PR8-HK (HON2) virus, with that of recombinant 2 (Fig. 1, lane 4) demonstrates that recombinant 3 (Fig. 1, lane 5) also derives RNA segments 4 and 7 from PR8 virus. In addition, recombinant 3 derives segment 6 (which corresponds to segment 5 of PR8 virus in lanes 1 and 6) and segment 8 from PR8 virus. Analysis of the proteins of recombinant 3 reveals that it derives its nucleoprotein, membrane protein, and nonstructural protein from PR8 virus (Fig. 2, lanes 3 and 5). All other proteins are derived from the HK virus parent (except for hemagglutinin, coded for by RNA segment 4). As discussed above, analysis of the RNAs and proteins of recombinant 1 demonstrated that the NS protein is coded for by RNA 8, and similar analysis of recombinant 2 revealed that the membrane protein is coded for by RNA 7. Hence, by a process of elimination we conclude that RNA segment 6 of HK virus segment 5 of PR8 virus code for nucleoprotein.

All of the recombinant viruses employed in the present experiments were shown by serological analysis to contain HK virus neuraminidase. Analysis of their RNAs confirms that all contain a segment that migrates in the position of band 5 of HK virus, and none have RNA bands that comigrate with RNA 6 of PR8 virus. This confirms our previous observation that neuraminidase is coded for by RNA 5 of HK virus and RNA 6 of PR8 virus (6). Likewise, serological identification and the migration of the 4th segments of all three recombinants confirms our earlier identification of the gene coding for hemagglutinin (6).

In addition, all three recombinants derive their three slowest-moving RNA segments from HK virus and, although the conditions

under which proteins were analyzed in the present experiments were not optimal for separation of P3 proteins, the results shown in Fig. 2 are consistent with our previous identification of the genes coding for the P proteins (Palese et al., Virology, in press).

DISCUSSION

Traditionally, genetic maps have been constructed by analyzing multifactor crosses of two organisms, which permits an assessment of relative distances, and hence the position, of markers. The ability to construct such maps depended upon finding a sufficient number of distinct markers to show linkage across the entire, linear genome. The segmented nature of the influenza virus genome does not lend itself readily to this type of analysis. Recently, however, we have developed polyacrylamide gel systems that allow us to distinguish each and every RNA segment of PR8 virus from equivalent segments of HK virus, with which PR8 virus readily recombines (Palese et al., Virology, in press; 6, 9). Similarly, systems were developed to separate and distinguish all proteins of these two viruses except for the neuraminidases, which can readily be distinguished serologically. Analysis of recombinants containing one or few genes from one parent and all other genes from the other parent allowed us to identify the gene products of specific RNA segments. Thus by analyzing a recombinant that contains only one RNA segment and one protein from PR8 virus, deriving the rest from HK virus, we have shown that RNA segment 8 codes for the nonstructural protein (Fig. 3). Analysis of a second recombinant possessing two proteins and two RNA segments from PR8 virus confirmed the previous result that the 4th largest RNA segment codes for hemagglutinin (6) and showed that segment 7 codes for the membrane protein. Finally, analysis of a third recombinant, which derived four RNA segments and four proteins from PR8 virus, established that band 5 of PR8 virus and band 6 of HK virus code for nucleoprotein (Fig. 3). This completes the mapping of the influenza virus genome by a positive identification of the products of all eight genes.

It should be noted that the gene products of each of the RNA segments have molecular weights that are in reasonable agreement with the coding capacity of the corresponding RNA segment as determined previously (8). Possible exceptions to this are the P proteins, one of which, the smallest (P3 protein), is coded for by the slowest-moving RNA segment and the neuraminidases, which are coded for by either RNA segment 5 or 6. However, the latter is a glycoprotein, which makes molecular weight determinations less reliable.

In addition to identifying the gene products of specific RNA segments, the results described here and in the two previous communications (Palese et al., Virology, in press; 6) have other implications. First of all, the assumption that each RNA segment codes for a virus-specific protein and that all eight segments are translated is confirmed.

A second conclusion from our results is that influenza virus neuraminidases are coded by only one RNA segment and that consequently the enzyme consists of identical subunits and not of two different subunits as has been reported (2).

In addition, the availability of recombinant viruses that can be defined with respect to the derivation of each gene and each gene product makes it possible to examine the relationship of differences in specific proteins to virus-related differences in biological activity. In this connection we suggest that the current system of nomenclature that defines influenza virus recombinants only with respect to hemagglutinin and

FIG. 1. *Analysis of [^{32}P]RNAs of influenza PR8 and HK viruses and three recombinant viruses derived from them. Lanes 1 and 6, PR8 virus; lane 2, HK virus; lane 3, recombinant 1; lane 4, recombinant 2; lane 5, recombinant 3. Migration is from top to bottom on a 2.6% urea-polyacrylamide gel (23 cm long) containing 0.15% N,N'-methylenebisacrylamide. The RNA segments of PR8 and HK virus are numbered 1 through 8. The letters next to the RNA segments of the recombinant viruses indicate their derivation from PR8 virus (P) or HK virus (H).*

FIG. 2. *Protein analysis of PR8 and HK viruses and three recombinants derived from them. Lane 1, MDCK cells infected with recombinant 1; lane 2, MDCK cells infected with PR8 virus; lanes 3 and 5, MDCK cells infected with recombinant 3; lane 4, MDCK cells infected with recombinant 2; lane 6, MDCK cells infected with HK virus. Extracts of [^{35}S]methionine-labeled cells were analyzed on 5 to 13% sodium dodecyl sulfate gradient-polyacrylamide gels (23 cm long) (see Materials and Methods). Electrophoresis was for 19 h at 100 V. All proteins of the recombinant viruses that were derived from HK virus are identified by arrows. Proteins of recombinants derived from PR8 virus are unmarked.*

FIG. 3. *Complete map of the influenza virus genome. The RNAs of PR8 and HK virus were separated on a polyacrylamide gel as shown in Fig. 1. Based on this and two previous communications (Palese et al., Virology, in press; 6) it is now possible to assign a gene product to each of the eight influenza virus RNA segments.*

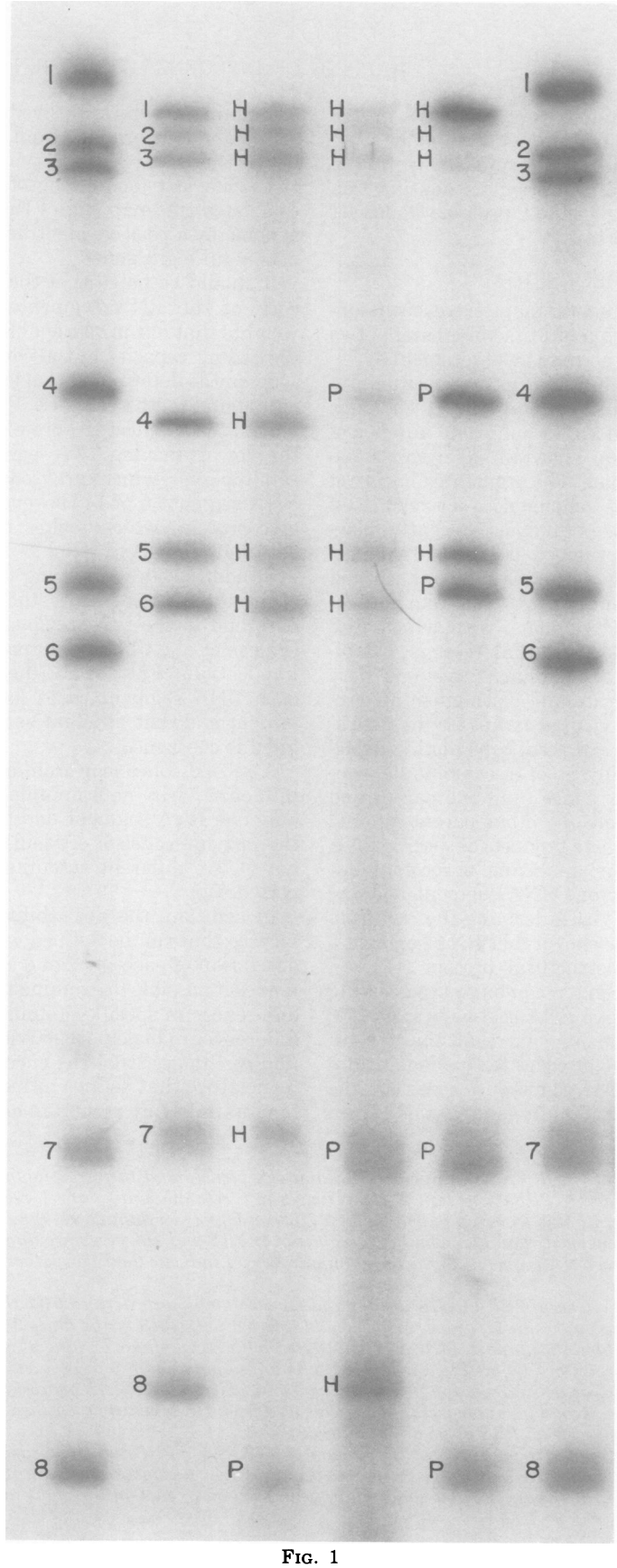
1
2
3
4
5
6
7
8
1
2
3
4
5
6
7
8
H
H
H
H
H
H
H
H
H
H
H
H
H
H
H
H
H
H
H
H
P
P
P
P
P
P
1
2
3
4
5
6
7
8

FIG. 1

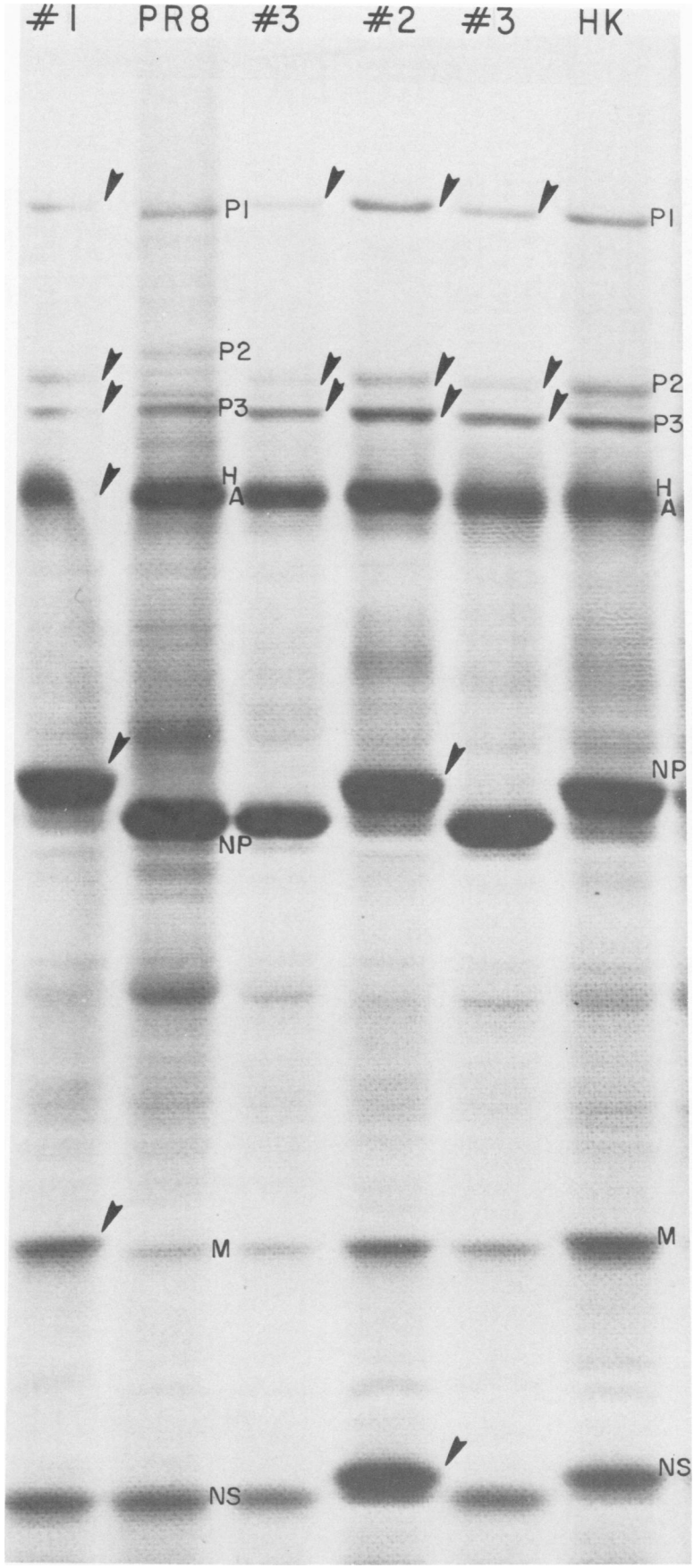

FIG. 2

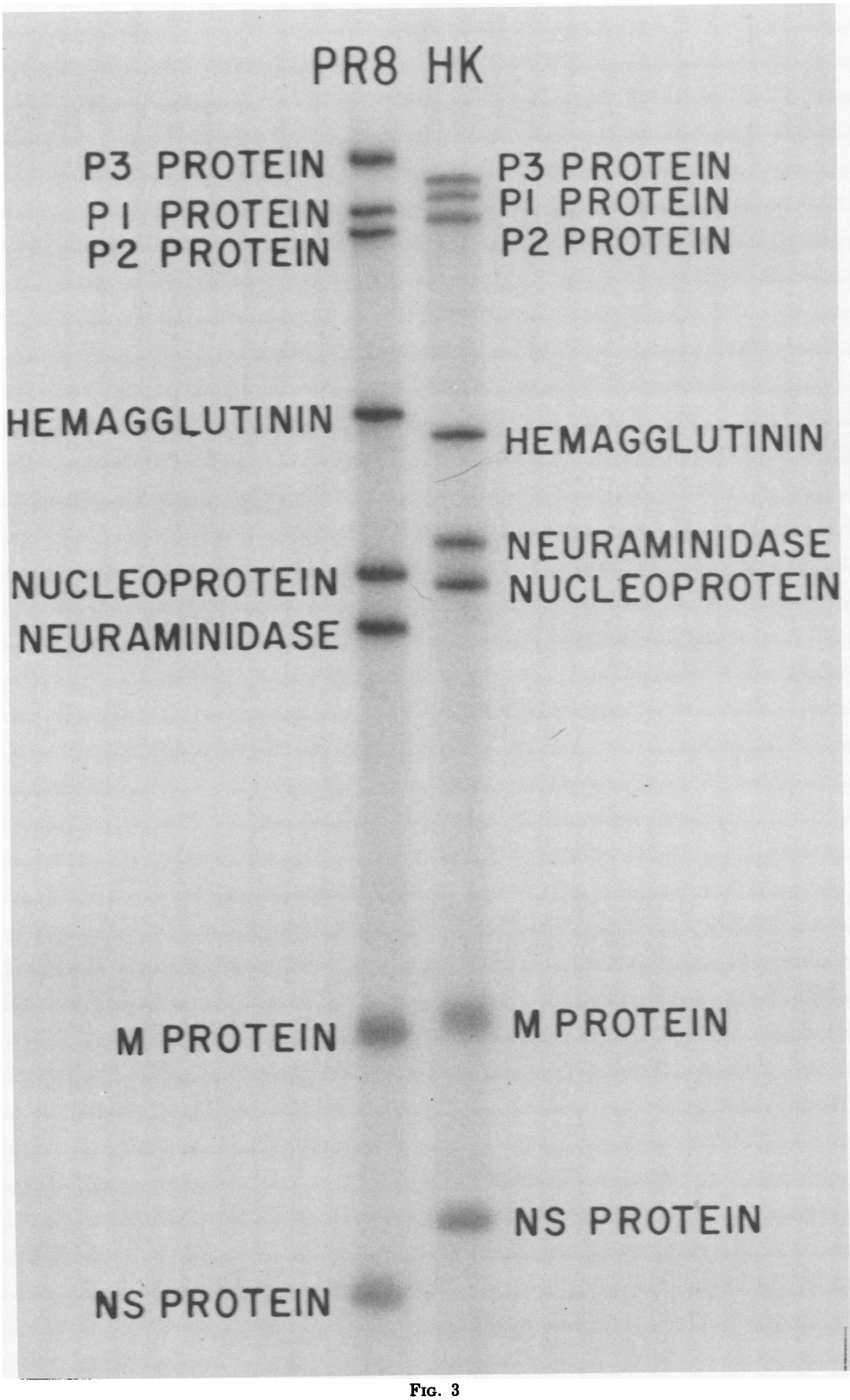
PR8 HK
P3 PROTEIN
P1 PROTEIN
P2 PROTEIN
HEMAGGLUTININ
NUCLEOPROTEIN
NEURAMINIDASE
M PROTEIN
NS PROTEIN
P3 PROTEIN
P1 PROTEIN
P2 PROTEIN
HEMAGGLUTININ
NEURAMINIDASE
NUCLEOPROTEIN
M PROTEIN
NS PROTEIN

FIG. 3

neuraminidase is not adequate to meet advancing knowledge of influenza virus genetics and propose that in the future recombinant viruses should be defined with respect to the derivation of all of their genes.

ACKNOWLEDGMENTS

We wish to express our appreciation for the excellent technical assistance of Marlene Lin and Kaye Leitzinger.

This work was supported by Public Health Service grants AI-11823 and AI-09394 from the National Institute of Allergy and Infectious Diseases and by National Science Foundation grant PCM-76-11066. M. B. R. was supported by Public Health Service training grant I-TOI-AI-00439 from the National Institute of Allergy and Infectious Diseases.

LITERATURE CITED

1. **Aminoff, D.** 1961. Methods for the quantitative estimation of N-acetylneuraminic acid and their application to hydrolysates of sialomucoids. Biochem. J. 81:384–392.
2. **Bucher, D. J., and E. D. Kilbourne.** 1972. A2 (N2) neuraminidase of the X-7 influenza virus recombinant: determination of molecular size and subunit composition of the active units. J. Virol. 10:60–66.
3. **Floyd, R. W., M. P. Stone, and W. K. Joklik.** 1974. Separation of single stranded ribonucleic acids by acrylamide-agarose-urea gel electrophoresis. Anal. Biochem. 59:599–609.
4. **Palese, P., and J. L. Schulman.** 1974. Isolation and characterization of influenza virus recombinants with high and low neuraminidase activity. Virology 57:227–237.
5. **Palese, P., and J. L. Schulman.** 1976. Differences in RNA patterns of influenza A viruses. J. Virol. 17:876–884.
6. **Palese, P., and J. L. Schulman.** 1976. Mapping of the influenza virus genome: identification of the hemagglutinin and the neuraminidase genes. Proc. Natl. Acad. Sci. U.S.A. 73:2142–2146.
7. **Palese, P., K. Tobita, M. Ueda, and R. W. Compans.** 1974. Characterization of temperature sensitive influenza virus mutants defective in neuraminidase. Virology 61:397–410.
8. **Ritchey, M., P. Palese, and E. D. Kilbourne.** 1976. RNAs of influenza A, B, and C viruses. J. Virol. 18:738–744.
9. **Schulman, J. L., and P. Palese.** 1976. Selection and identification of influenza virus recombinants of defined genetic composition. J. Virol. 20:248–254.