JOURNAL OF VIROLOGY, Feb. 1986, p. 697–700
0022-538X/86/020697-04$02.00/0

Vol. 57, No. 2

# Characterization of a Gene Coding for M Proteins Which Is Involved in Host Range Restriction of an Avian Influenza A Virus in Monkeys

ALICIA J. BUCKLER-WHITE,[1]* CLAYTON W. NAEVE,[2] AND BRIAN R. MURPHY[1]

*Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, Bethesda, Maryland 20892,[1] and Division of Virology and Molecular Biology, St. Jude Children's Research Hospital, Memphis, Tennessee 38101[2]*

Received 17 April 1985/Accepted 24 October 1985

**The nucleotide sequence of the region of RNA segment 7 coding for the M1 and M2 proteins of avian influenza A/Mallard/New York/6750/78 was determined, and the deduced amino acid sequences were compared to other avian and human M protein sequences. The M2 proteins of the avian and human viruses have diverged much more than the M1 proteins, although amino acids specific for avian and human viruses were found in both M1 and M2 proteins.**

An avian influenza virus, A/Mallard/New York (NY)/6750/78, that was markedly restricted in replication in the trachea of squirrel monkeys, was evaluated as a donor of its non-surface protein genes for attenuation of virulent human influenza A viruses. A "six-gene" avian-human influenza A reassortant virus, which contained the hemagglutinin and neuraminidase surface antigen genes from the human A/Udorn/307/72 strain and the six non-surface antigen genes (internal genes) from the avian A/Mallard/NY/78 strain, was as restricted in replication in the respiratory tract of squirrel monkeys as its avian influenza A virus parent, indicating that restriction of replication of the avian influenza virus is a function of one or more of its internal genes (7). To investigate which avian influenza gene or combination of genes was responsible for restricted replication in primates, reassortant viruses were produced that contained human influenza A/Udorn/72 virus surface antigen genes and one or more internal genes derived from the avian influenza virus donor. Reassortants that contained only a nucleoprotein (segment 5) or matrix protein (segment 7) RNA segment of avian influenza origin were markedly restricted in replication, as were reassortants containing either of these segments in combination with one or more other avian influenza genes (13). The avian nucleoprotein and matrix protein genes must, therefore, independently play a major role in the restriction of replication exhibited by the A/Mallard/78 virus and reassortants containing its internal genes.

RNA segment 7 is 1,027 nucleotides long and codes for two proteins, M1 and M2, and possibly an additional small peptide (1, 3, 14). The M1 protein is a structural protein that contains 252 amino acids. The M2 protein shares the first 9 amino acids of the M1 protein, while the remainder of its 97 amino acids are translated in the +1 reading frame from a splice acceptor site. The extent of nucleotide divergence between the matrix protein gene of the A/Mallard/78 virus and the corresponding gene from human influenza viruses was determined to initiate studies aimed at understanding the molecular basis for this host range restriction.

The nucleotide sequence of the coding region of the A/Mallard/78 M protein RNA segment was determined by extending a series of synthetic oligonucleotide primers by reverse transcription with dideoxynucleotides using the influenza virus genome RNA as a template (11). Virus was purified and virion RNA extracted as described previously (4). Virion RNA was fractionated by electrophoresis on a 2.6% polyacrylamide gel containing 4.5 M urea (6), and segment 7 was eluted from the gel in 0.5 M NaCl–0.1 M Tris hydrochloride (pH 8.0)–5 mM EDTA. The oligonucleotide primers, 1 AGCAAAAGCAGG, 220 GCTCACCGTGCCC, 488 GCTGACTCCCAG, and 736 GCAGGCCTATCA, were selected within regions of homology between the A/Udorn/72 (3) and A/PR/8/34 (1, 14) matrix protein genes and synthesized on an Applied Biosystems DNA synthesizer. Dideoxynucleotide chain termination sequencing reactions and gels were as described by Naeve et al. (8).

The sequence of nucleotides 15 to 1024 of the positive strand of segment 7 of A/Mallard/78 is shown in Fig. 1. Nucleotides 26 to 781, which constitute the single large open reading frame coding for the M1 protein, are indicated by a line above the sequence, while nucleotides 26 to 51 and 740 to 1004 in the spliced reading frame for the M2 protein are underlined. The sequence of nucleotides 1 to 14 and 1025 to 1027 was not determined. However, nucleotides 1 to 12 and 1025 to 1027 are not within the coding region of the RNA segment and fall within regions of base homology in all segments of influenza A viral RNA (10, 12).

The predicted amino acid sequence of the A/Mallard/78 M1 and M2 proteins was compared to the corresponding sequences of the A/Udorn/307/72 human influenza virus (3). Amino acid sequence divergence was much greater in the M2 protein, with 14 differences out of the 97 amino acids (Table 1). These differences were clustered within three regions of the M2 protein. In contrast, there were only seven amino acid differences in the M1 protein, although it is more than 2.5 times larger than the M2 protein, and the differences were spread more uniformly throughout the protein. A comparison of the specific amino acids at each of the sites at which a difference was documented indicated that none of these differences within the M1 protein involved a change of charge (Table 1), while a charged amino acid was substituted for a neutral one at six sites within the M2 protein.

Comparison of M protein sequences was extended to include the other available human, A/PR/8/34 (1, 14) and A/Bangkok/1/79 (9), and avian, A/fowl plague virus (FPV)/Rostock/34 (5), M gene sequences (Table 2). Although there is more sequence divergence between the A/Mallard/NY/78

* Corresponding author.

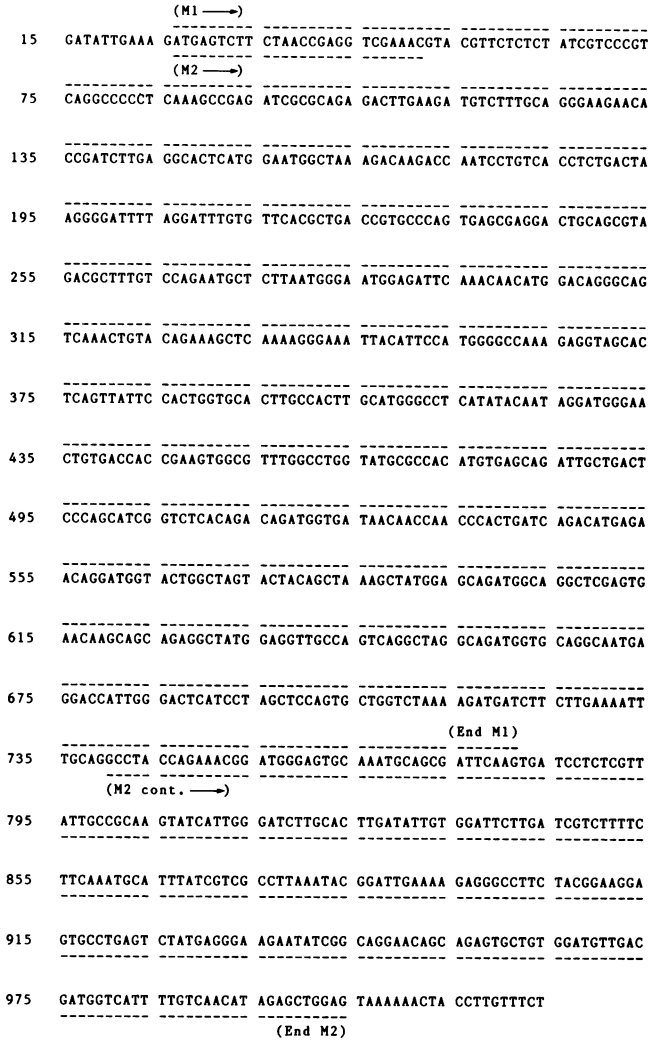


FIG. 1. Sequence of nucleotides 15 to 1024 in the mRNA sense of RNA segment 7 of the influenza A/Mallard/NY/6750/78 strain. The probable coding sequence for the M1 protein is indicated by a line above the nucleotide sequence; the probable spliced sequence encoding M2 is indicated by a line under the nucleotide sequence.

and A/Udorn/72 strains in the M2 protein than in the M1 protein, this pattern is reversed when the mallard strain is compared to another avian influenza A strain, A/FPV/Rostock/34. The M2 proteins differ in only 4 amino acids, while there are 10 amino acid differences in the M1 proteins. However, since the M1 protein is more than twice as large as the M2 protein, the percent divergence is approximately the same for each protein. The relatively high degree of divergence of the avian and human M2 proteins compared with the M1 proteins suggests independent evolution of the M1 and M2 proteins.

The NP genes of influenza A viruses can be divided into five classes, including one human and two avian classes, on the basis of hybridization analysis (2). The similarity in degree of divergence between the mallard strain and each of the human strains suggested the possibility of separate classes of M proteins in influenza A viruses of human and avian origin. A comparison of amino acid differences between the known human and avian influenza M proteins supports this interpretation. Of the 14 amino acid differences

TABLE 1. Amino acid sequence divergence between the M1 and M2 proteins of influenza A/Udorn/307/72 and A/Mallard/NY/6750/78 viruses

| Protein | Amino acid sequence divergence | | |
|---|---|---|---|
| | Position of amino acid difference | Udorn | Mallard |
| M1 | 90 | Pro | Ser |
| | 115 | Ile | Val |
| | 121[a] | Ala | Thr |
| | 126 | Ser | Thr |
| | 137[a] | Ala | Thr |
| | 167 | Ala | Ile |
| | 218[a] | Ala | Thr |
| M2 | 11[a] | Ile | Thr |
| | 14[b] | Glu | Gly |
| | 16[b] | Gly | Glu |
| | 18 | Arg | Lys |
| | 20 | Asn | Ser |
| | 28 | Val | Ile |
| | 54[b] | Phe | Arg |
| | 55 | Phe | Leu |
| | 56[b] | Glu | Lys |
| | 57[b] | His | Tyr |
| | 78[b] | Lys | Gln |
| | 86 | Ala | Val |
| | 89 | Ser | Gly |
| | 93 | Ser | Asn |

[a] Positions with amino acids differing in hydrophobicity.
[b] Positions with amino acids differing in charge.

between the A/Udorn/72 and A/Mallard/78 M2 proteins, 12 are also present in a comparison of the A/Udorn/72 and A/FPV/34 virus strains.

Comparison of all of the known avian and human influenza A M gene sequences, including sequences from strains isolated more than 40 years apart, indicates that there are 10 sites with amino acids specific for either avian or human influenza strains, i.e., a specific amino acid is conserved in the three human influenza strains while a different amino acid is present in both avian strains at that site (Fig. 2). Three of these sites fall within the M1 protein, while seven are located in the M2 protein. The three sites that possibly define an avian or human class of M1 protein are clustered within the center of the molecule at residues 115, 121, and 137. Three of the M2 amino acids specific to avian or human strains are located within the seven amino acids immediately downstream from the splice site (residues 11, 14, and 16), one is located in the center of the molecule at residue 55, and the remaining three are clustered near the carboxy terminus of the protein (residues 78, 86, and 93). Preliminary sequence data from the M protein RNA segment of two other avian

TABLE 2. Sequence divergence of the M proteins of avian and human influenza A viruses

| Protein[a] | No. (%) of differences in the amino acid sequence of the A/Mallard/NY/6750/78 M Proteins vs. the indicated avian or human influenza virus M proteins | | | |
|---|---|---|---|---|
| | A/FPV/Rostock/34 | A/Udorn/307/72 | A/PR/8/34 | A/Bangkok/1/79 |
| M1 | 10 (4.0) | 7 (2.8) | 10 (4.0) | 9 (3.6) |
| M2 | 4 (4.1) | 14 (14.4) | 14 (14.4) | 15 (15.5) |

[a] The M1 protein contains 252 amino acids; the M2 protein contains 97 amino acids.

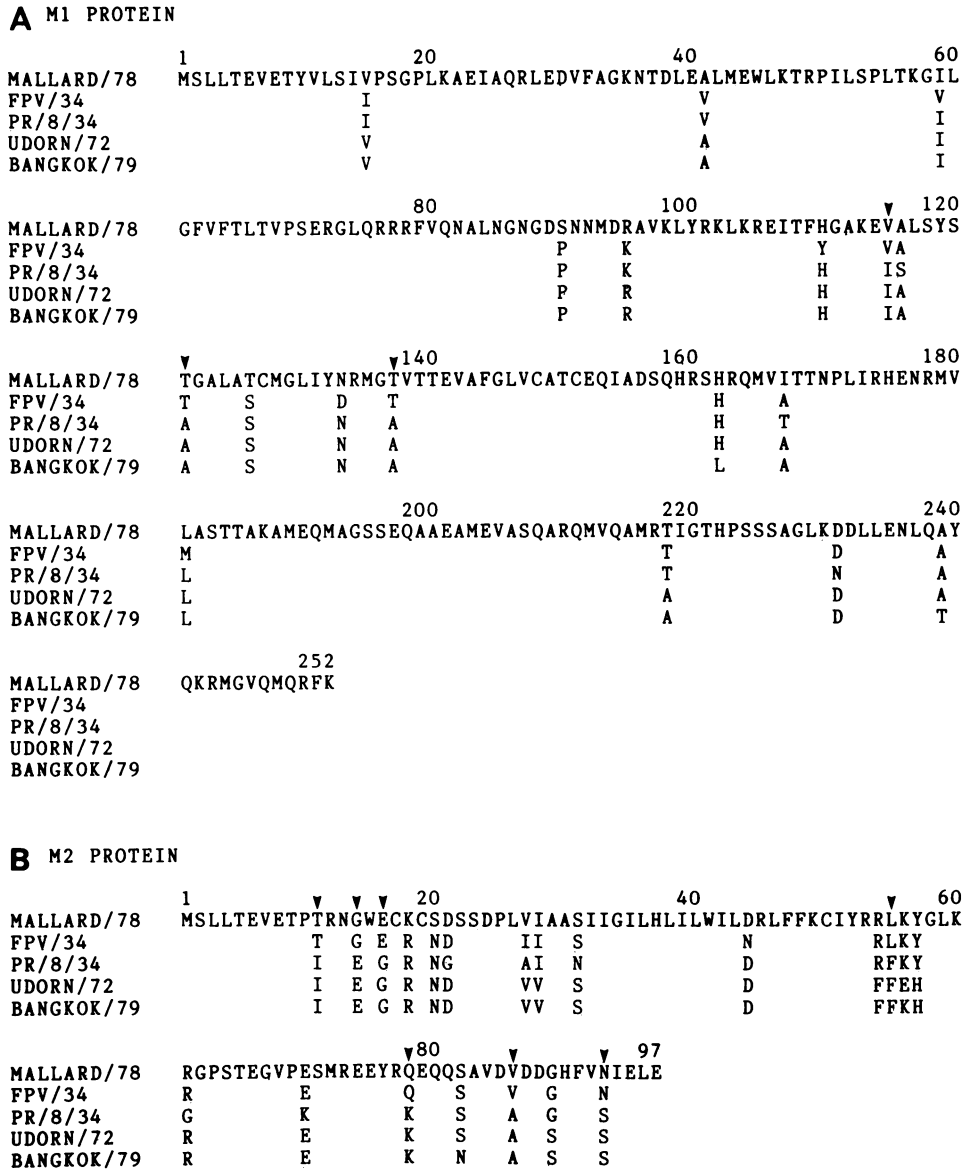


FIG. 2. Amino acid sequence of the M1 (A) and M2 (B) proteins of A/Mallard/NY/6750/78 as predicted from the nucleotide sequence of RNA segment 7. The sequence of A/Mallard/78 encoded proteins is indicated in the top line, with the sequences of A/FPV/Rostock/34 (5), A/PR/8/34 (1, 14), A/Udorn/307/72 (3), and A/Bangkok/1/79 (9) indicated at sites demonstrating heterogeneity. Sites at which conserved avian-influenza-specific and human-influenza-specific amino acids are located are indicated by arrows.

influenza virus strains, A/Pintail/Alberta/119/79 (H4N6) and A/Pintail/Alberta/121/79 (H7N8), support the conservation of avian-influenza-specific and human-influenza-specific amino acids at these sites. Analysis of the sequences of M proteins from additional avian and human influenza A virus strains may indicate which of these sites are important in defining classes of M proteins from different host species and could provide a method for evaluating the origin of genes other than HA or NA during pandemic change.

Although there is significant sequence divergence between the A/Mallard/78 and A/Udorn/72 strains in both the M1 and M2 proteins, the specific amino acid(s) responsible for the M protein-mediated host range restriction of the avian virus in monkeys remains to be identified. The attenuation phenotype may be specified by only one or two amino acids. Sequence analysis of phenotypic revertants isolated after passage of the attenuated virus in primates may indicate which of these sites are most important in determining the restriction of replication of the A/Mallard/78 virus and its avian-human influenza virus reassortants in primates. Additional evidence may come from analysis of M proteins of other avian influenza A viruses which replicate at different levels in squirrel monkeys.

## LITERATURE CITED

1. **Allen, H., J. McCauley, M. Waterfield, and M. J. Gething.** 1980. Influenza virus RNA segment 7 has the coding capacity for two polypeptides. Virology **107**:548–551.
2. **Bean, W. J.** 1984. Correlation of influenza A virus nucleoprotein genes with host species. Virology **133**:438–442.
3. **Lamb, R. A., and C. J. Lai.** 1981. Conservation of the influenza

virus membrane protein ($M_1$) amino acid sequence and an open reading frame of RNA segment 7 encoding a second protein ($M_2$) in H1N1 and H3N2 strains. Virology **112**:746–751.

4. **Massicot, J. G., B. R. Murphy, F. Thierry, L. Markoff, K.-Y. Huang, and R. M. Chanock.** 1980. Temperature-sensitive mutants of influenza virus. Identification of the loci of the two *ts* lesions in the Udorn-*ts*-1A2 donor virus and the correlation of the presence of these two *ts* lesions with a predictable level of attenuation. Virology **101**:242–249.
5. **McCauley, J. W., B. W. J. Mahy, and S. C. Inglis.** 1982. Nucleotide sequence of fowl plague virus RNA segment 7. J. Gen. Virol. **58**:211–215.
6. **Murphy, B. R., A. J. Buckler-White, W. T. London, J. Harper, E. L. Tierney, N. T. Miller, L. J. Reck, R. M. Chanock, and V. S. Hinshaw.** 1984. Avian-human reassortant influenza A viruses derived by mating avian and human influenza A viruses. J. Infect. Dis. **150**:841–850.
7. **Murphy, B. R., D. L. Sly, E. L. Tierney, N. T. Hosier, J. G. Massicot, W. T. London, R. M. Chanock, R. G. Webster, and V. S. Hinshaw.** 1982. Reassortant virus derived from avian and human influenza A viruses is attenuated and immunogenic in monkeys. Science **218**:1330–1332.
8. **Naeve, C. W., V. S. Hinshaw, and R. G. Webster.** 1984. Mutations in the hemagglutinin receptor-binding site can change the biological properties of an influenza virus. J. Virol. **51**: 567–569.
9. **Ortin, J., C. Martinez, L. del Rio, M. Davila, C. Lopez-Galindez, N. Villanueva, and E. Domingo.** 1983. Evolution of the nucleotide sequence of influenza virus RNA segment 7 during drift of the H3N2 subtype. Gene. **23**:233–239.
10. **Robertson, J. S.** 1979. 5′ and 3′ terminal nucleotide sequences of the RNA genome segments of influenza virus. Nucleic Acids Res. **6**:3745–3757.
11. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
12. **Skehel, J. J., and A. J. Hay.** 1978. Nucleotide sequences at the 5′ termini of influenza virus RNAs and their transcripts. Nucleic Acids Res. **5**:1207–1219.
13. **Tian, S.-F., A. J. Buckler-White, W. T. London, L. J. Reck, R. M. Chanock, and B. R. Murphy.** 1985. Nucleoprotein and membrane protein genes are associated with restriction of replication of influenza A/Mallard/NY/78 virus and its reassortants in squirrel monkey respiratory tract. J. Virol. **53**:771–775.
14. **Winter, G., and S. Fields.** 1980. Cloning of influenza cDNA into M13: the sequence of the RNA segment encoding the A/PR/8/34 matrix protein. Nucleic Acids Res. **8**:1965–1974.